Rajiv Khosla
Robert J. Howlett
Lakhmi C. Jain (Eds.)

Knowledge-Based Intelligent Information and Engineering Systems

9th International Conference, KES 2005
Melbourne, Australia, September 2005
Proceedings, Part II

2 **Part II**

Springer

Lecture Notes in Artificial Intelligence 3682

Rajiv Khosla   Robert J. Howlett
Lakhmi C. Jain (Eds.)

Knowledge-Based Intelligent Information and Engineering Systems

9th International Conference, KES 2005
Melbourne, Australia, September 14-16, 2005
Proceedings, Part II

Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Rajiv Khosla
La Trobe University
Business Systems and Knowledge Modelling Laboratory
School of Business, Victoria 3086, Australia
E-mail: R.Khosla@latrobe.edu.au

Robert J. Howlett
University of Brighton
School of Engineering, Engineering Research Centre
Moulsecoomb, Brighton, BN2 4GJ, UK
E-mail: r.j.howlett@bton.ac.uk

Lakhmi C. Jain
University of South Australia
School of Electrical and Information Engineering, KES Centre
Mawson Lakes Campus, Adelaide, South Australia SA 5095, Australia
E-mail: Lakhmi.Jain@unisa.edu.au

Preface

Dear delegates, friends and members of the growing KES professional community, welcome to the proceedings of the 9th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems hosted by La Trobe University in Melbourne Australia.

The KES conference series has been established for almost a decade, and it continues each year to attract participants from all geographical areas of the world, including Europe, the Americas, Australasia and the Pacific Rim. The KES conferences cover a wide range of intelligent systems topics. The broad focus of the conference series is the theory and applications of intelligent systems. rom a pure research field, intelligent systems have advanced to the point where their abilities have been incorporated into many business and engineering application areas. KES 2005 provided a valuable mechanism for delegates to obtain an e tensive view of the latest research into a range of intelligent-systems algorithms, tools and techni ues. The conference also gave delegates the chance to come into contact with those applying intelligent systems in diverse commercial areas. The combination of theory and practice represented a uni ue opportunity to gain an appreciation of the full spectrum of leading-edge intelligent-systems activity.

The papers for KES 2005 were either submitted to invited sessions, chaired and organi ed by respected e perts in their fields, or to a general session, managed by an e tensive International Program Committee, or to the Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP) orkshop, managed by an International orkshop Technical Committee. hichever route they came through, all papers for KES 2005 were thoroughly reviewed. The adoption by KES of the PROSE Publication Review and Organisation System software greatly helped to improve the transparency of the review process and aided uality control.

In total, 1382 papers were submitted for KES 2005, and a total of 688 papers were accepted, giving an acceptance rate of just under 50 . The proceedings, published this year by Springer, run to more than 5000 pages. The invited sessions are a valuable feature of KES conferences, enabling leading researchers to initiate sessions that focus on innovative new areas. A number of sessions in new emerging areas were introduced this year, including E perience Management, Emotional Intelligence, and Smart Systems. The diversity of the papers can be judged from the fact that there were about 100 technical sessions in the conference program. More than 400 universities worldwide participated in the conference making it one of the largest conferences in the area of intelligent systems. As would be e pected, there was good local support with the participation of 20 Australian universities. There was a significant business presence, provided by the involvement of a number of industry bodies, for e ample, CSIRO Australia, DSTO Australia, Daewoo South Korea and NTT Japan.

KES International gratefully acknowledges the support provided by La Trobe University in hosting this conference. e acknowledge the active interest and support from La Trobe University's Vice Chancellor and President, Prof. Michael Osborne, Dean of

the aculty of Law and Management, Prof. Raymond Harbridge, Dean of the aculty of Science and Technology, Prof. David inlay, and Head of the School of Business, Prof. Malcolm Rimmer. KES International also gratefully acknowledges the support provided by Emeritus Prof. Greg O'Brien.

A tremendous amount of time and effort goes into the organi ation of a conference of the si e of KES 2005. The KES community owes a considerable debt of gratitude to the General Chair Prof. Rajiv Khosla and the organi ing team at La Trobe University for their huge efforts this year in bringing the conference to a successful conclusion. As the conference increases in si e each year the organi ational effort needed increases and we would like to thank Prof. Khosla and his colleagues for coping efficiently with the largest KES conference to date.

 e would like to thank the Invited Session Chairs, under the leadership and guid- ance of Prof. Lakhmi Jain and Prof. Rajiv Khosla for producing high- uality sessions on leading-edge topics.   e would like to thank the KES 2005 International Program Committee for undertaking the considerable task of reviewing all of the papers submit- ted for the conference.   e e press our gratitude to the high-profile keynote speakers for providing talks on leading-edge topics to inform and enthuse our delegates. A con- ference cannot run without authors to write papers.   e thank the authors, presenters and delegates to KES 2005 without whom the conference could not have taken place. inally we thank the administrators, caterers, hoteliers, and the people of Melbourne for welcoming us and providing for the conference.

 e hope you found KES 2005 a worthwhile, informative and enjoyable e perience.

July 2005   Bob Howlett
                                                                        Rajiv Khosla
                                                                        Lakhmi Jain

KES 2005 Conference Organization

General Chair

Rajiv Khosla
Business Systems and Knowledge Modelling Laboratory
School of Business
La Trobe University
Melbourne, Victoria 3086
Australia

Conference Founder and Honorary Program Committee Chair

Lakhmi C. Jain
Knowledge-Based Intelligent Information and Engineering Systems Centre
University of South Australia, Australia

KES Executive Chair

Bob Howlett
Intelligent Systems and Signal Processing Laboratories/KTP Centre
University of Brighton, UK

KES Journal General Editor

Bogdan Gabrys
University of Bournemouth, UK

Local Organizing Committee

Malcolm Rimmer – Chair, School of Business, La Trobe University
Rajiv Khosla, Selena Lim, Brigitte Carrucan, Monica Hodgkinson, Marie enton, Maggie Van Tonder, and Stephen Muir
La Trobe University, Melbourne, Australia

KES 2005 Web Page Design Team

Joe Hayes, Anil Varkey Samuel, Mehul Bhatt, Rajiv Khosla
La Trobe University, Melbourne, Australia

KES 2005 Liaison and Registration Team

Rajiv Khosla, Selena Lim, Brigitte Carrucan, Jodie Kennedy, Maggie Van Tonder, Marie enton, Colleen Stoate, Diane Kraal, Cary Slater, Petrus Usmanij, Chris Lai, Rani Thanacoody, Elisabeth Tanusasmita, George Plocinski
La Trobe University, Melbourne, Australia

KES 2005 Proceedings Assembly Team

Rajiv Khosla
Selena Lim
Monica Hodgkinson
George Plocinski
Maggie Van Tonder
Colleen Stoate
Anil Varkey Samuel
Marie enton
Mehul Bhatt
Chris Lai
Petrus Usmanij

International Program Committee

Hussein Abbass, University of New South ales, Australia
Akira Asano, Hiroshima University, Japan
Robert Babuska, Delft University of Technology, The Netherlands
Patrick Bosc, IRISA/ENSSAT, rance
Pascal Bouvry, Lu embourg University of Applied Sciences, Lu embourg
Kr ys tof Cios, University of Colorado, USA
Carlos A. Coello, LANIA, Me ico
Ernesto Damiani, University of Milan, Italy
Da Deng, University of Otago, New ealand
Vladan Deved ic, University of Belgrade, Serbia and Montenegro
Vladimir Gorodetski, Russian Academy of Sciences, Russia
Manuel Grana, UPV/EHU, Spain
Lars Kai Hansen, Technical University of Denmark, Denmark
Yu o Hirai, University of Tsukuba, Japan
Daniel Howard, ineti , UK
T ung-Pei Hong, National University of Kaohsiung, Taiwan
Hisao Ishibuchi, Osaka Prefecture University, Japan
Naohiro Ishii, Aichi Institute of Technology, Japan
Seong-Joon Yoo, Sejong University, Korea
Nikos Karacapilidis, University of Patras, Greece
Rajiv Khosla, La Trobe University, Australia
Las lo T. Koc y, Budapest and S echenyi Istvan University, Hungary
Andrew Kusiak, University of Iowa, USA
 .K. Lai, MIMOS Berhad, Malaysia
Ian C. Parmee, University of the est of England, UK
Dong Hwa Kim, Hanbat National University, Korea
Jingli Cao, La Trobe University, Australia
Da Ruan, Belgian Nuclear Research Centre, Belgium
Ang Yang, University of New South ales, Australia
Adrian Stoica, NASA Jet Propulsion Laboratory, USA
Janus Kacpr yk, Polish Academy of Sciences, Poland
Vijayan Asari, Old Dominion University, USA
Raymond Lee, Hong Kong Polytechnic University, Hong Kong, China
Chee-Peng Lim, University of Science, Malaysia
Ignac Lovrek, University of agreb, Croatia
Bob McKay, University of NS , Australia
Dan C. Marinescu, University of Central lorida, USA
Radko Mesiar, Slovak Technical University, Slovakia
Hirofumi Nagashino, University of Tokushima, Japan
Ciprian Daniel Neagu, University of Bradford, UK
Ngoc Thanh Nguyen, roclaw University of Technology, Poland

Invited Session Chairs Committee

Krzys tof Cios, University of Denver, USA
Toyoaki Nishida, University of Tokyo, Japan
Ngoc Thanh Nguyen, roclaw University of Technology, Poland
T ung-Pei Hong, National University of Kaohsiung, Taiwan
Yoshinori Adachi, Chubu University, Japan
Nobuhiro Inu uka, Nagoya Institute of Technology, Japan
Naohiro Ishii, Aichii Institute of Technology, Japan
Yuji Iwahori, Chubu University, Japan
Tai-hoon Kim, Security Engineering Research Group (SERG), Korea
Daniel Howard, ineti , UK
Mircea Gh. Negoita, ellington Institute of Technology, New ealand
Akira Suyama, University of Tokyo, Japan
Dong Hwa Kim, Hanbat National University, Korea
Hussein A. Abbas, University of New South ales, Australia
illiam Grosky, University of Michigan-Dearborn, USA
Elisa Bertino, Purdue University, USA
Maria L. Damiani, University of Milan, Italy
Kosuke Sekiyama, University of ukui, Japan
Seong-Joon Yoo, Sejong Univwersity, Korea
Hiroka u Taki, akayama University, Japan
Satoshi Hori, Institute of Technologists, Japan
Hideyuki Sawada, Kagawa University, Japan
Vicenc Torra, Artificial Intelligence Research Institute, Spain
Maria Vamrell, Artificial Intelligence Research Institute, Spain
Manfred Schmitt, Technical University of Munich, Germany
Denis Helic, Gra University of Technology, Austria
Yoshiteru Ishida, Toyohashi University of Technology, Japan
Giuseppina Passiante, University of Lecce, Italy
Ernesto Damiani, University of Milan, Italy
Susumu Kunifuji, Japan Advanced Institute of Science and Technology, Japan
Motoki Miura, Japan Advanced Institute of Science and Technology, Japan
James Liu, Hong Kong Polytechnic University, Hong Kong, China
Honghua Dai, Deakin University, Australia
Pascal Bouvry, Lu embourg University, Lu embourg
Gregoire Danoy, Lu embourg University, Lu embourg
Ojoung Kwon, California State University, USA
Jun Munemori, akayama University, Japan
Takashi Yoshino, akayama University, Japan
Takaya Yui ono, Shimane University, Japan
Behrou Homayoun ar, University of Calgary, Canada
Toyohide atanabe, Nagoya University, Japan
Dharmendra Sharma, University of Canberra, Australia
Dat Tran, University of Canberra, Australia

Kim Le, University of Canberra, Australia
Takumi Ichimura, Hiroshima City University, Japan
K Yoshida, St. Marianna University, Japan
Phill Kyu Rhee, Inha University, Korea
Chong Ho Lee, Inha University, Korea
Mikhail Prokopenko, CSIRO ICT Centre, Australia
Daniel Polani, University of Hertfordshire, UK
Dong Chun Lee, Howon University, Korea
Dawn E. Holmes, University of California, USA
Kok-Leong Ong, Deakin University, Australia
Vincent Lee, Monash University, Australia
  ee-Keong Ng, Nanyang Technological University
Gwi-Tae Park, Korea University, Korea
Giles Oatley, University of Sunderland, UK
Sangkyun Kim, Korea Information Engineering Service, Korea
Hong Joo Lee, Daewoo Electronics Corporation, Korea
Ryohei Nakano, Nagoya Institute of Technology, Japan
Ka umi Saito, NTT Communication Science Laboratories, Japan
Ka uhiko Tsuda, University of Tsukuba, Japan
Torbjørn Nordgård, Norwegian University of Science and Technology, Norway
Øystein Nytrø, Norwegian University of Science and Technology, Norway
Amund Tveit, Norwegian University of Science and Technology, Norway
Thomas Bro Røst, Norwegian University of Science and Technology, Norway
Manuel Graña, Universidad Pais Vasco, Spain
Richard Duro, Universidad de A Coruña, Spain
Ka umi Nakamatsu, University of Hyogo, Japan
Jair Minoro Abe, University of Sao Paulo, Bra il
Hiroko Shoji, Chuo University, Japan
Yukio Ohsawa, University of Tsukuba, Japan
Da Deng, University of Otago, New ealand
Irena Koprinska, University of Sydney, Australia
Eiichiro Ta aki, University of Yokohama, Japan
Kenneth J. Mackin, Tokyo University of Information Sciences, Japan
Lakhmi Jain, University of South Australia, Australia
Tetsuo uchino, Tokyo Institute of Technology, Japan
Yoshiyuki Yamashita, Tohoku University, Japan
Martin Purvis, University of Otago, New ealand
Marius Nowostawski, University of Otago, New ealand
Bastin Tony Roy Savarimuthu, University of Otago, New ealand
Norio Baba, Osaka Kyoiku University, Japan
Jun o atada, aseda University, Japan
Petra Povalej, Laboratory for System Design, Slovenia
Peter Kokol, Laboratory for System Design, Slovenia
  oochun Jun, Seoul National University of Education, Korea
Andrew Kusiak, University of Iowa, USA
Hanh Pham, State University of New York, USA

IIHMSP Workshop Organization Committee

General Co-chairs

Jeng-Shyang Pan
National Kaohsiung University of Applied Sciences, Taiwan
Lakhmi C. Jain
University of South Australia, Australia

Program Committee Co-chairs

ai-Chi ang
California Institute of Technology, USA
Eiji Kawaguchi
Keio University, Japan

Finance Chair

Jui- ang Chang
National Kaohsiung University of Applied Sciences, Taiwan

Publicity Chair

Kang K. Yen
  lorida International University, USA

Registration Chair

Yan Shi
Tokai University, Japan

Electronic Media Chair

Bin-Yih Liao
National Kaohsiung University of Applied Sciences, Taiwan

Publication Chair

Hsiang-Cheh Huang
National Chiao Tung University, Taiwan

Local Organizing Chair

R. Khosla
La Trobe University, Australia

Asia Liaison

Yoiti Su uki
Tohoku University, Japan

North America Liaison

Yun   . Shi
New Jersey Institute of Technology, USA

Europe Liaison

R.J. Howlett
University of Brighton, UK

IIHMSP Workshop Technical Committee

Prof. Oscar Au, Hong Kong University of Science and Technology, Hong Kong, China
Prof. Chin-Chen Chang, National Chung Cheng University, Taiwan
Prof. Chang   en Chen,   lorida Institute of Technology, USA
Prof. Guanrong Chen, City University of Hong Kong, Hong Kong, China
Prof. Liang-Gee Chen, National Taiwan University, Taiwan
Prof. Tsuhan Chen, Carnegie Mellon University, USA
Prof. Yung-Chang Chen, National Tsing-Hua University, Taiwan
Prof. Yen-   ei Chen, Ritsumeikan University, Japan
Prof. Hsin-Chia   u, National Chiao Tung University, Taiwan
Prof. Hsueh-Ming Hang, National Chiao Tung University, Taiwan
Prof. Dimitrios Hat inakos, University of Toronto, Canada
Dr. Ichiro Kuroda, NEC Electronics Corporation, Japan

KES 2005 Reviewers

H. Abbass, University of New South ales, Australia
J.M. Abe, University of Sao Paulo, Bra il
Y. Adachi, Chubu University, Japan
 . Alpaslan, Middle East Technical University, Turkey
P. Andreae, Victoria University, New ealand
A. Asano, Hiroshima University, Japan
K.V. Asari, Old Dominion University, USA
N. Baba, Osaka-Kyoiku University, Japan
R. Babuska, Delft University of Technology, The Netherlands
P. Bajaj, G.H. Raisoni College of Engineering, India
A. Bargiela, Nottingham Trent University, UK
M. Ba u, Institute of Microtechnology, Romania
N. Berthou e, University of Ai u, Japan
E. Bertino, Purdue University, USA
Y. Bodyanskiy, Kharkiv National University of Radioelectronics, Ukraine
P. Bosc, IRISA/ENSSAT, rance
P. Bouvry, Lu embourg University, Lu embourg
P. Burrell, South Bank University, UK
J. Cao, La Trobe University, Australia
B. Chakraborty, Iwate Prefectural University, Japan
Y.- . Chen, Ryukyus University, Japan
Y.-H. Chen-Burger, University of Edinburgh, UK
V. Cherkassky, University of Minnesota, USA
K. Cios, University at Denver, USA
C.A. Coello, LANIA, Me ico
G. Coghill, University of Auckland, New ealand
D. Corbett, SAIC, USA
D. . Corne, University of E eter, UK
D. Cornforth, Charles Sturt University, Australia
 .S.C. da Silva, University of Sao Paulo, Bra il
H. Dai, Deakin University, Australia
E. Damiani, University of Milan, Italy
M.L. Damiani, University of Milan, Italy
G. Danoy, Lu embourg University, Lu embourg
K. Deep, Indian Institute of Technology Roorkee, India
D. Deng, University of Otago, New ealand
V. Deved ic, University of Belgrade, Serbia and Montenegro
D. Dubois, Université Paul Sabatier, rance
R. Duro, Universidad de A Coruña, Spain
D. Earl, Oak Ridge National Laboratory, USA
B. ar, University of Calgary, Canada

M.   athi, National Magnet Laboratory, USA
R.   lóre -Lópe , Campus de Vega ana, Spain
M.   rean, Victoria University of   ellington, New   ealand
R.   rias, University of Porto, Portugal
T.   uchino, Tokyo Institute of Technology, Japan
P.   unk, Mälardalen University, Sweden
B. Gabrys, University of Bournemouth, UK
B. Galitsky, University of London, UK
T. Gedeon, Murdoch University, Australia
M. Gen,   aseda University, Japan
A. Ghosh, ISICAI, India
V. Gorodetski, Russian Academy of Sciences, Russia
M. Grana,   acultad de Informatica UPV/EHU, Spain
   . Grosky, University of Michigan-Dearborn, USA
A. Gr ech,   roclaw University of Technology, Poland
D. Gwaltney, NASA George C. Marshall Space   light Center, USA
L.K. Hansen, Technical University of Denmark, Denmark
C.J. Harris, University of Southampton, UK
D. Helic, Gra   University of Technology Austria
L. Hildebrand, University of Dortmund, Germany
Y. Hirai, University of Tsukuba, Japan
D.E. Holmes, University of California, USA
B. Homayoun,   ar University of Calgary, Canada
T.-P. Hong, National University of Kaohsiung, Taiwan
S. Hori, Institute of Technologists, Japan
K. Horio, Kyushu Institute of Technology, Japan
D. Howard,   ineti , UK
B. Howlett, University of Brighton, UK
M.-P. Huget, University of Savoie,   rance
H. Iba, University of Tokyo, Japan
T. Ichimura, Hiroshima City University, Japan
N. Inu uka, Nagoya Institute of Technology, Japan
H. Ishibuchi, Osaka Prefecture University, Japan
Y. Ishida, Toyohashi University of Technology, Japan
N. Ishii, Aichi Institute of Technology, Japan
Y. Iwahori, Chubu University, Japan
L. Jain, University of South Australia, Australia
M.M. Jamshidi, University of New Me ico, USA
N. Jesse, Universität Dortmund, Germany
S. Joo, Sejong University, Korea
   . Jun, Seoul National University of Education, Korea
J. Kacpr yk, Polish Academy of Sciences, Poland
N. Karacapilidis, University of Patras, Greece
V. Kecman, Auckland University, New   ealand
R. Khosla, La Trobe University, Australia

D.H. Kim, Hanbat National University, Korea
S. Kim, Korea Information Engineering Service, Korea
T.-H. Kim, Security Engineering Research Group (SERG), Korea
L.T. Koc y, Budapest University of Technology and Economics, Hungary
P. Kokol, Laboratory for System Design, Slovenia
A. Konar, Jadavpur University, India
I. Koprinska, University of Sydney, Australia
H. Koshimi u, Chukyo University, Japan
S. Kunifuji, Japan Advanced Institute of Science and and Technology, Japan
A. Kusiak, University of Iowa, USA
O. Kwon, California State University, USA
 .K. Lai, MIMOS Berhad, Malaysia
P.L. Lan i, Polytechnic Institute, Italy
K. Le, University of Canberra, Australia
C.H. Lee, Inha University, Korea
D.C. Lee, Howon University, Korea
H.J. Lee, Daewoo Electronics Corporation, Korea
R. Lee, Hong Kong Polytechnic University, Hong Kong, China
V. Lee, Monash University, Australia
 . Li, La Trobe University, Australia
C.-P. Lim, University of Science, Malaysia
S. Lim, La Trobe University, Australia
J. Liu, Hong Kong Polytechnic University, Hong Kong, China
I. Lovrek, University of agreb, Croatia
H. Lu, La Trobe University, Australia
B. MacDonald, Auckland University, New ealand
K.J. Mackin, Tokyo University of Information Sciences, Japan
L. Magdalena-Layos, EUS LAT, Spain
D.C. Marinescu, University of Central lorida, USA
 . Masulli, University of Pisa, Italy
J. Ma umdar, University of South Australia, Australia
B. McKay, University of NS , Australia
S. McKinlay, ellington Institute of Technology, New ealand
R. Mesiar, Slovak Technical University, Slovakia
J. Mira, UNED, Spain
Y. Mitsukura, University of Okayama, Japan
M. Miura, Japan Advanced Institute of Science and Technology, Japan
J. Munemori, akayama University, Japan
H. Nagashino, University of Tokushima, Japan
N. Nagata, Chukyo University, Japan
K. Nakajima, Tohoku University, Japan
K. Nakamatsu, University of Hyogo, Japan
R. Nakano, Nagoya Institute, Japan
T. Nakashima, Osaka University, Japan
L. Narasimhan, University of Newcastle, Australia

V.E. Neagoe, Technical University, Romania
C.D. Neagu, University of Bradford, UK
M.G. Negoita,   elTec, New   ealand
   .-K. Ng, Nanyang Technological University, Singapore
C. Nguyen, Catholic University of America, USA
N.T. Nguyen,   roclaw University of Technology, Poland
T. Nishida, University of Tokyo, Japan
T. Nordgård, Norwegian University of Science and Technology, Norway
M. Nowostawski, University of Otago, New   ealand
Ø. Nytrø, Norwegian University of Science and Technology, Norway
G. Oatley, University of Sunderland, UK
Y. Ohsawa, University of Tsukuba, Japan
E. Oliveira, University of Porto, Portugal
K.-L. Ong, Deakin University, Australia
N.R. Pal, Indian Statistical Institute, India
V. Palade, O ford University, UK
G.-T. Park, Korea University, Korea
I.C. Parmee, University of the   est of England, UK
G. Passiante, University of Lecce, Italy
C.-A. Peña-Reyes, Swiss   ederal Institute of Technology - EP L, Swit erland
H. Pham, State University of New York, USA
D. Polani, University of Hertfordshire, UK
T. Popescu, National Institute for Research and Development Informatic, Italy
P. Povalej, Laboratory for System Design, Slovenia
M. Prokopenko, CSIRO ICT Centre, Australia
M. Purvis, University of Otago, New   ealand
G. Resconi, Catholic University, Italy
B. Reusch, University of Dortmund, Germany
P.K. Rhee, Inha University, Korea
J.A. Rose, University of Tokyo, Japan
T.B. Røst, Norwegian University of Science and Technology, Norway
E. Roventa, York University, Canada
R. Roy, Cranfield University, UK
D. Ruan, Belgian Nuclear Research Centre, Belgium
A. Saha, NCD, Papua New Guinea
K. Saito, NTT Communication Science Laboratories, Japan
T. Samatsu, Kyushu Tokai University, Japan
E. Sanche , Université de la Méditeranée,   rance
B.T.R. Savarimuthu, University of Otago, New   ealand
H. Sawada, Kagawa University, Japan
M. Schmitt, Technical University of Munich, Germany
M. Schoenauer, INRIA,   rance
U. Seiffert, Leibni Institute of Plant Genetics
    and Crop Plant Research Gatersleben, Germany
K. Sekiyama, University of   ukui, Japan

D. Sharma, University of Canberra, Australia
H. Shoji, Chuo University, Japan
A. Skabar, La Trobe University, Australia
B. Smyth, University College Dublin, Ireland
V.- . Soo, National Tsing Hua University, Taiwan
A. Stoica, NASA Propulsion Jet Laboratory, USA
M.R. Styt , Yamaguchi University, Japan
N. Suetake, Yamaguchi University, Japan
S. Sujitjorn, Suranaree University of Technology, Thailand
 . Sun, University of ollongong, Australia
A. Suyama, University of Tokyo, Japan
H. Taki, akayama University, Japan
T. Tanaka, ukuoka Institute of Technology, Japan
M. Tanaka-Yamawaki, Tottori University, Japan
E. Ta aki, University of Yokohama, Japan
S. Thatcher, University of South Australia, Australia
P. Theodor, National Institute for Research and Development Informatics, Romania
J. Timmis, University of Kent at Canterbury, UK
V. Torra, Artificial Intelligence Research Institute, Spain
J. Torresen, University of Oslo, Norway
D. Tran, University of Canberra, Australia
K. Tsuda, University of Tsukuba, Japan
C. Turchetti, Università Politecnica delle Marche, Italy
A. Tveit, Norwegian University of Science and Technology, Norway
J. Tweedale, Defence Science and Technology Organi ation, Australia
A.M. Tyrrell, University of York, UK
E. Uchino, University of Yamaguchi, Japan
A. Uncini, University of Rome, Italy
P. Urlings, Defence Science and Technology Organi ation, Australia
M. Vamrell, Artificial Intelligence Research Institute, Spain
J.L. Verdegay, University of Granada, Spain
M. Virvou, University of Piraeus, Greece
S. alters, University of Brighton, UK
D. ang, La Trobe University, Australia
L. ang, Nanyang Technical University, Singapore
P. ang, Temple University, USA
K. ard, University of ollongong, Australia
J. atada, aseda University, Japan
K. atanabe, Saga University, Japan
T. atanabe, Nagoya University, Japan
T. Yamakawa, Kyushu Institute of Technology, Japan
Y. Yamashita, Tohoku University, Japan
A. Yang, University of New South ales, Australia
 . Yao, University of Birmingham, UK
S.-J. Yoo, Sejong University, Korea

K. Yoshida, Kyushu Institute of Technology, Japan
T. Yoshino, akayama University, Japan
T. Yui ono, Shimane University, Japan
L. adeh, Berkeley University of Caliornia, USA
D. ambarbieri, Università di Pavia, Italy
 . ha, National Institute of Standards and Technology, USA

KES 2005 Keynote Speakers

1. Professor Jun Liu, Harvard University, MA, USA
    Topic: rom Se uence Information to Gene E pression

2. Professor Ron Sun, Rensselaer Polytechnic Institute, New York, USA
    Topic: rom Hybrid Systems to Hybrid Cognitive Architectures

3. Professor Jiming Liu, Hong Kong Baptist University, Hong Kong, China
    Topic: Towards Autonomy Oriented Computing (AOC):
                 ormulating Computational Systems with Autonomous Components

4. Professor Toyoaki Nishida, Kyoto University and Tokyo University, Japan
    Topic: Ac uiring, Accumulating, Transforming, Applying,
         and Understanding Conversational   uanta

5. Professor Marimuthu Palaniswami, University of Melbourne, Australia
    Topic: Convergence of Smart Sensors and Sensor Networks

Table of Contents, Part II

Machine Learning

Immunity-Based Systems

Medical Diagnosis

Intelligent Hybrid Systems and Control

Emotional Intelligence and Smart Systems

Context-Aware Evolvable Systems

Intelligant Fuzzy Systems and Control

Knowledge Representation and Its Practical Application in Today's Society

Approaches and Methods into Security Engineering III

Communicative Intelligent I

Intelligent Watermaking Algorithms and Applications

Intelligent Techniques and Control

e-Learning and ICT

Logic Based Intelligent Information Systems

Intelligent Agents and Their Applications I

Innovations in Intelligent Agents

Ontologies and the Semantic Web

Knowledge Discovery in Data Streams

Computational Intelligence Tools Techniques and Algorithms

Approaches and Methods to Security Engineering IV

Watermaking Applications I

Watermaking Applications II

Multimedia Retrieval I

Soft Computing Approach to Industrial Engineering

Experience Management and Information Systems

Table of Contents, Part III

Intelligent Agent Ontologies and Environments

Intelligent Multimedia Solutions and the Security in the Next Generation Mobile Networks

Intelligent E-Mail Analysis, News Extraction and Web Mining

Semantic Integration and Ontologies

Computer Vision, Image Processing and Retrieval

Communicative Intelligence II

Approaches and Methods to Security Engineering V

Multimedia Retrieval II

Multimedia Compression

Multimedia Signal Processing

Emergence and Self-organisation in Complex Systems

Soft Computing Techniques and Their Applications III

Information Engineering and Ubiquitous Computing

Location and Context-Based Systems

e-Based Systems in Education, Commerce and Health

Computational Biology and Bioinformatics

Complex Adaptive Systems

Communicative Intelligent III

Speech Processing and Robotics

Data Mining and Soft Computing Applications I

Multimedia Security and Stegnography

Stegnography

Soft Computing Approach to Industrial Engineering II

Medical Text Mining and Natural Language Processing

Knowledge Based Intelligent Systems for Health Care

Intelligent Learning Environment

Intelligent Data Analysis and Applications

Table of Contents, Part IV

Innovations in Intelligent Systems and Their Applications

Data Mining and Soft Computing Applications II

Skill Acquisition and Ubiquitous Human Computer Interaction

Soft Computing and Their Applications – IV

Agent-Based Workflows, Knowledge Sharing and Reuse

Multi-media Authentication and Watermarking Applications

Knowledge and Engineering Techniques
for Spatio-temporal Applications

Intelligent Data Analysis and Applications II

Creativitiy Support Environment and Its Social Applications

Collective Intelligence

Computational Methods for Intelligent Neuro-fuzzy Applications

Evolutionary and Self-organizing Sensors, Actuators and Processing Hardware

Knowledge Based Systems for e-Business and e-Learning I

Multi-agent Systems and Evolutionary Computing

Ubiquitous Pattern Recognition

Neural Networks for Data Mining

Intelligent Systems for e-Business and e-Learning II

Knowledge-Based Technology in Crime Matching, Modelling and Prediction

Soft Computing Applications

Table of Contents, Part I

Soft Computing Techniques in Stock Markets

Intelligent Network Based Education

Maintenance and Customization of Business Knowledge

Intelligent Data Processing in Process Systems and Plants

Intelligent Agent Technology and Applications I

Intelligent Design Support Systems

Data Engineering, Knowledge Engineering and Ontologies

Knowledge Discovery and Data Mining

Advanced Network Application

Approaches and Methods of Security Engineering I

Chance Discovery I

Information Hiding and Multimedia Signal Processing

Soft Computing Techniques and Their Applications I

Intelligent Agent Technology and Applications II

Smart Systems

Knowledge – Based Interface Systems

Intelligent Information Processing for Remote Sensing

Intelligent Human Computer Interaction Systems

Experience Management and Knowledge Management

Network (Security) Real-Time and Fault Tolerant Systems

Advanced Network Application and Real-Time Systems

Approaches and Methods of Security Engineering II

Soft Computing Techniques and Their Applications II

POISE – Achieving Content-Based Picture Organisation for Image Search Engines

Da Deng and Heiko Wolf

Department of Information Science, University of Otago
PO Box 56, Dunedin, New Zealand
{ddeng,hwolf1}@infoscience.otago.ac.nz

Abstract. To overcome the drawbacks of both keywords-based and content-based image retrieval approaches, we propose a hybrid solution for image retrieval on WWW. A content organisation system is introduced to assess content-based similarity of results returned from contemporary image search engines so that the results are organised with content similarity, and alternative search modes such as query by example are also enabled.

1 Introduction

The ever-growing World Wide Web has become the largest information corpus that challenges our traditional ways of information organisation and retrieval. Our increased perhaps daily experience in using search engines should have manifested this. While the retrieval of textual information has been the major focus of commercial search engines so far, there is a growing demand in effectively accessing multimedia data stored on the Web, such as images, videos and sound clips . There are a few search engines (for example, the Google Image Search [1]) offering image and music search capabilities. However, almost all commercially available web image search engines today are keyword-based. In most cases, the images are indexed using text near to the image, such as filenames, captions or HTML tags (e.g. IMG and ALT) [2]. A search engine returns a range of images according to a query entered. It may also employ techniques to rank the images according to some relevancy criteria. Because of the limit of this approach, results returned from these search engines are often irrelevant, random and unorganised [3]. Little improvement is observed on Meta-search engines which process user queries through a number of general engines to retrieve a broader result set.

On the other hand, content-based image retrieval (CBIR) has been a major research topic since the early 1990s. CBIR made some significant achievements in providing new ways of image (and multimedia) querying, such as query by example and query by sketch etc. However, its reliance on low level features also keeps itself from reaching further, despite the use of techniques such as relevance feedback. Consequently, result sets obtained through CBIR are based on similarities defined on low visual features of colour, shape and textual, but little is guaranteed on the semantic level. To our knowledge, there are a number

of experimental CBIR image search engines [4], but few are available on the Web today. For example, WebSEEK system [5] is a semi-automatic image search engine that collects information which then is catalogued and indexed for fast search and retrieval. Both textual and content-based queries are possible. It only indexes and categorises about 600,000 images, nothing comparable with the commercial text-based search engines. Another drawback of CBIR is its queries need much longer response time.

In this paper, we introduce POISE (Picture Organisation Interface for Search Engines), a system that uses visual features to organise the picture results obtained from image search engines and thus enables a better retrieval of relevant images. Functioning as an interface between textual and content-based image retrieval, it strengthens the use of image search engines with CBIR capabilities, such as automatic content organisation and query-by-examples.

2   Content-Based Picture Organisation

2.1   System Overview

The goal of our work is to take advantage of the fast response time of powerful search engines and their huge amounts of indexed images and improve the search outcome with automated content organisation catered for easier browsing, as well as to provide query-by-example capabilities to enable the user to find the most relevant images.

The building blocks of the system are depicted in Figure 1. The user enters a keyword, after which an image search engine is queried. The retrieved images are then displayed as thumbnails in the application. To sort the retrieved images based on their visual similarity, visual features are generated, clustered and visualised in an image map. After retrieving the result set from the search engine, the user also can invoke a "query-by-example" action, whereupon the system presents the closest matches from the result set. This step can be combined with the organised image map to further facilitate the search process.
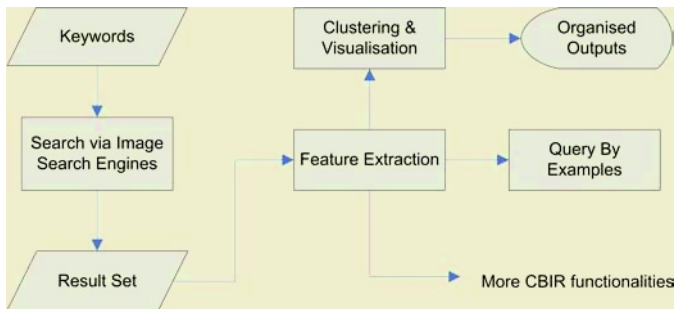
Fig. 1. POISE system overview

2.2   Visual Features

Colour is found to be the most expressive visual feature employed in CBIR systems. The MPEG-7 core experiments (XM) include a number of rigorously evaluated colour descriptors, such as the colour layout (CLD) and the scalable colour (SCD) [6]. The CLD captures the spatial distribution of colour in an image and is defined as representative colours in YCrCb space on an 8×8 grid followed by DCT. The SCD on the other hand is a colour histogram feature quantised in 256 bins in the HSV colour space. XM has also defined a number of texture and shape descriptors.

To achieve an organised interface for browsing the image results, a feature space of relatively low embedding dimensionality would be desirable. Following the study made in [7] on chromaticity histograms, the embedding dimension of CLD is found to be rather low in our experiments. The residual variance of dimensions becomes lower than 0.04 since since the third dimension for Isomap [8], or since the fourth dimension in principal component analysis (PCA). The test image set of more than 1000 images in [9] is used to get these results. Based on this analysis, and regarding the computational efficiency, CLD is chosen as the main feature for image content organisation.

2.3   Automatic Content Organisation

To organise the almost random search results returned by image search engines, a straightforward approach is to carry out clustering analysis on the visual features of the returned images and self-organise them into some feature maps that are easy to browse. As the results vary greatly from query to query, and a large amount of image data are to be processed, the overall computational process needs to be fast and effective.

There are a number of different methods to organise and visualise high-dimensional data, for example dimension reduction using PCA or Isomap. A method often used in current image retrieval research is the self-organising map (SOM) [10], which conducts data clustering and multi-dimension scaling at the same time. Recent research efforts using the SOM include a CBIR system utilising a hierarchical SOM implementation [11], and the visualisation and comparison of image collections [12]. We adopt the SOM as the clustering algorithm to organise the image set based on two reasons: First, SOMs map high-dimensional data to a two- or three-dimensional space, which makes visualisation and human interpretation of this data easier; and second, SOMs are topology preserving, which means that similar inputs are mapped to the same node or nodes in the neighbourhood, a very good property for a browsing interface.

As mentioned before, the CLD feature space has a low embedding dimension, so a dimension reduction process will help to reduce the input dimension and thence the training time of the SOMs, which is verified in our experiments. We use PCA for dimension reduction since it works much faster than Isomap.

2.4   Other Functionalities

Because of the use of CBIR techniques, more functionality can be introduced into the system to improve the search quality. If the organised image map does not reveal an appropriate result, another search option is made available: "query-by-example". In this mode, the user selects an image from the sorted result set and is then presented with a subset with the best matched images. Another MPEG-7 descriptor SCD is also employed to improve the query quality. The Euclidean distance is used on the Haar transform coefficients of the SCD.

More visual features can be used, for instance the homogeneous texture descriptor, but this may slow down the query process.

3   The POISE System

The POISE system is implemented in Java and can run as an applet, hence accessible from any Java-enabled web browser. The Google image search is used as the search engine for testing purposes. From the user interface one can specify keywords and number of images of the search results to fetch from the search engine. For each result set retrieved from the search engine, a two-dimensional feature map is created using the CLD features from each image. The size of the map is chosen to be 10% of that of the result set, small enough to allow easy navigation and search. For the current size of result sets (up to 2000 images), flat rather than hiearchical SOMs are used for computational simplicity. For each node, the closest five images are mapped to by sorting the Euclidean distance of the feature vectors. The results are then presented in the search window, with the best match for each node displayed magnified and the other four images displayed as smaller thumbnails around it. By clicking on the thumbnails, they will be magnified. This allows easier finding of similar images, one of the major design goals of our application. An organised picture set is shown in Figure 2. The ordered display of 120 images in our system is an advance from the unordered display of 20 images on one Google image search result page.

The system works particularly well for specific searches, for example a search for a movie poster. Figure 2 shows the SOM-generated image grouping for the query "Finding Nemo". While the search engine returns an unordered set with images of sushi, people and other movies all mixed together, our system groups similar images together and therefore allows a faster search for a relevant group. If more similar images are needed, for example a movie poster in different resolutions, the query-by-example can be used. In Figure 3, the result for a query for the movie poster showing the shark can be seen.

A very important issue for any web search engine is the retrieval speed. With the low computational cost for the feature generation, our system performs particularly well. The times spent for the query "Finding Nemo" with different result set sizes are shown in Figure 4. It can be seen that most time is used to retrieve the images from the search engine (an average of 19.67 s), while the actual feature generation, clustering and frame reorganisation are less expensive. The SOM clustering was further sped up by reducing the feature vector to
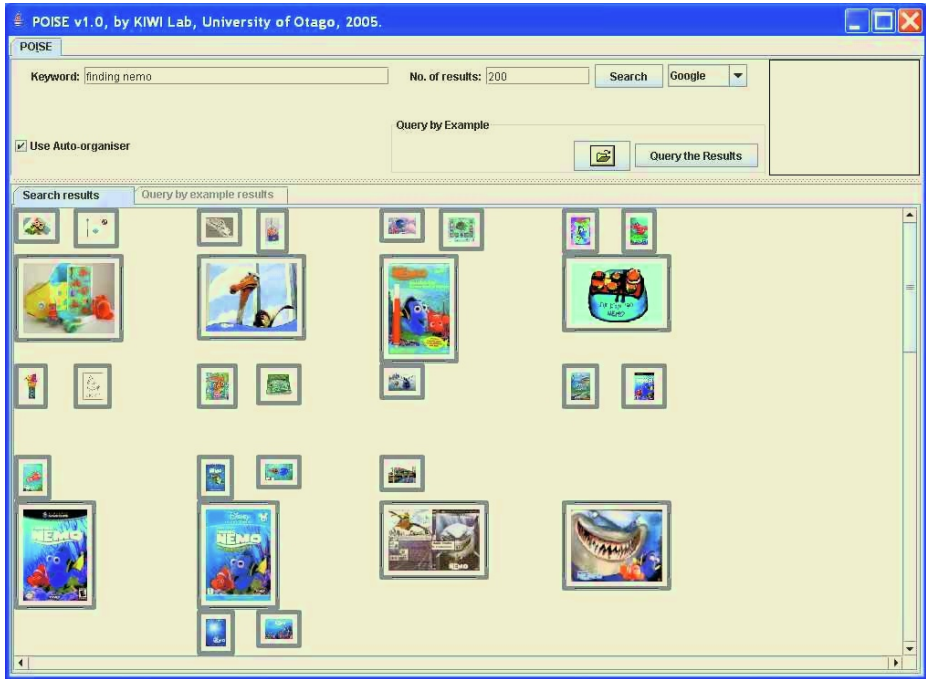
Fig. 2. Result set organised by SOM using CLD features

the first two principal components. This step enhances the speed of the SOM clustering by up to 24%, which will become more important if the result sets grow. The retrieval and feature generation parts of our system can be easily parallelised, which would lead to a further speed-up of the process.

All results have been obtained running POISE on a 3.2GHz Pentium 4 with 1GB RAM. Since the measured retrieval times depend heavily on the fluctuation of network traffic and the current load of the search engine servers, validity of measured times was achieved by using the average of 100 queries.

4   Conclusions

The POISE system we developed has shown that automatic content organisation can improve the efficiency and usefulness of web image searches. Especially searches with big result sets of hundreds or thousands of images will benefit from the clear presentation organised in visual similarity. The query-by-example capability can lead to an even more refined result set, enabling the user to effectively browse through similar images to find the target. Also, the retrieval and processing works very fast, which is important for a successful web search engine. As revealed in Figure 4, a bottleneck emerges when larger result sets are used, since hauling many images from the search engine is very time-consuming using

Fig. 3. Query-by-example result set
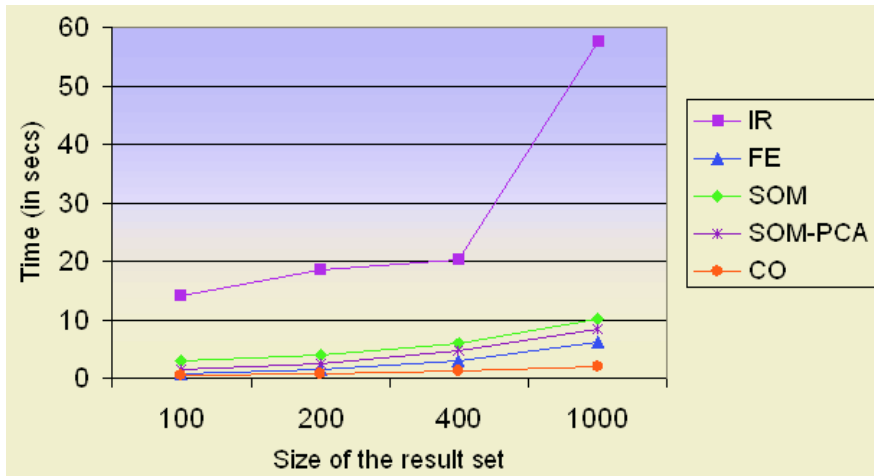
Fig. 4. Average time for the computing steps: IR - image retrieval, FE - feature extraction, SOM - mapping, SOM-PCA - mapping with 2-D PCA, CO - content organisation

a single computer. This limit can be overcome by parallelising the retrieval and image processing tasks and dispatching them to a dedicated computer cluster.

There are many open issues for future work. The automatic content organisation could be improved by combining different MPEG-7 features, for example colour layout, dominant colour, edge histogram and texture browsing. The

global approach used in the current system is limiting the accuracy, a problem that could be overcome by segmenting the images and including texture and shape descriptors so as to support region-based image retrieval.

Also, the SOM matching could be improved by using multiple SOMs and choosing the best fitting by a relevance feedback system such as the one used in the MetaSEEK system [13]. Hierarchical clustering such as a tree-structured SOM might enable easier navigation, but both approaches are also computationally more expensive than the single SOM used in our system, which can accommodate about 1000-2000 images in a result set. In general, a "flat" visualisation like the one used in our system can only display a limited number of results, for larger sets of several of thousands of images other approaches such as hierarchical structures could be employed.

To improve the system, we are experimenting with more feature descriptors and implementing a cluster-based system with parallelised processes of image acquisition, feature extraction and hierarchical clustering.

References

1. Google: Image search. URL http://images.google.com/ (2005)
2. Entlich, R.: FAQ image search engines. RLG DigiNews **5** (2001)
3. TASI: A review of image search engines (2004)
4. Veltkamp, R., Tanase, M.: Content-based image retrieval systems: a survey. Technical Report UU-CS-2004-34, Department of Computing Science, Utrecht University (2002)
5. Smith, J.R., Chang, S.F.: WebSEEK - a content-based image and video search and catalog tool for the web (2005)
6. Manjunath, B., Ohm, J., Vasudevan, V., Yamada, A.: Colour and texture descriptors. IEEE Trans on Circuits and Systems for Video Technology **11** (2001) 703–715
7. Funt, B., Kulpinski, D., Cardei, V.: Non-linear embeddings and the underlying dimensionality of reflectance spectra and chromaticity histograms. In: Proc. IS&T/SID Color Imaging Conference 2001. (2001) 126–129
8. Tenenbaum, J., De Silva, V., Langford, J.: A global geometric framework for non-linear dimensionality reduction. Science **290** (2000) 2319–2323
9. Wang, J., Li, J., Wiederhold, G.: Simplicity: Semantics-sensitive integrated matching for picture libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence **23** (2001) 947–963
10. Kohonen, T.: Self-organizing Maps. Second edn. Springer-Verlag (1997)
11. Laaksonen, J., Koskela, M., Oja, E.: Content-based image retrieval using self-organizing maps. In: Visual Information and Information Systems. (1999) 541–548
12. Deng, D.: Content-based profiling of image collections: a som-based approach. International Journal of Computers, Systems and Signals **5** (2004) 44–52
13. Benitez, A.B., Beigi, M., Chang, S.F.: Using relevance feedback in content-based image metasearch. IEEE Internet Computing **2** (1998) 58–69

Estimation of the Hierarchical Structure of a Video Sequence Using MPEG-7 Descriptors and GCS

Masoumeh D. Saberi[1,2], Sergio Carrato[1], Irena Koprinska[3], and James Clark[3]

[1] DEEI, University of Trieste, Trieste, Italy
[2] ICTP, International Center for Theoretical Physics, Trieste, Italy
[3] School of Information Technologies, University of Sydney, Sydney, Australia

Abstract. We present a clustering approach for video summarization and browsing based on the Growing Cell Structures (GCS) neural algorithms and MPEG-7 video descriptors. Each I frame is represented as a feature vector of MPEG-7 descriptors. These vectors are clustered using GCS to select one or more keyframes for each shot. The extracted keyframes are then grouped together by TreeGCS to form a hierarchical view of the video for efficient browsing. We evaluate the effectiveness of different combinations of MPEG-7 descriptors and also compare them with the performance of color histograms.

Introduction

Digital video is becoming widely available and used in education, entertainment and commerce. It is a sequential and information-rich medium, with long temporal contextual relationships between shots and scenes. Consequently, there is a great interest in designing and building systems that organize and search video data based on its content or other important video features. In addition to search capabilities, such systems should be able to derive intuitive and compact data representations so that users can easily and quickly browse through the whole database or through the results of a query. There has been a growing interest in developing content-based techniques for video organization that provide fast and meaningful non-linear access to the relevant material. Most of these systems first segment the video into shots, then select key frames to represent the content of each shot, and finally organise the shots into a hierarchical structure to allow faster browsing and retrieval. In this paper we present an approach for key frame selection and video browsing based on the MPEG-7 descriptors and the Growing Cell Structures (GCS) neural algorithm. After temporal video segmentation into shots, our system clusters the frames based on their MPEG-7 descriptors using the GCS algorithm. A hierarchical clustering algorithm (TreeGCS) is then used to cluster the shots and provide a high-level hierarchical video representation. This work extends the approach presented in [11], where simple feature vectors based on color histograms were used as frames descriptors. We also evaluate the effectiveness of different combinations of MPEG-7 descriptors for our application.

The motivation behind our system is as follows:
- MPEG-7 based features may provide a rich and accurate description of the frames,
- GCS may be able to accurately find clusters of similar frames within each shot in unsupervised fashion, and, similarly,

- using the extracted keyframes, TreeGCS may be able to provide a good high-level video hierarchical representation thanks to its self-organization and ability to form a flexible hierarchy depending on the video content.

1 Previous Work

As already mentioned, video summarization is an important technique for efficient video browsing and management. A number of techniques have been proposed and developed to automatically create video summaries. In this section, some of the most common are briefly reviewed.

In [2], the video skimming generation problem is formulated as a two-stage optimization problem. Some approaches use MPEG-7 descriptions of the multimedia content in order to develop integrated search mechanisms that assist users in identifying additional material that is specifically of interest to them. For example, MPEG-7 motion activity descriptor are used in [3] to establish video summarization and indexing techniques. Mosaic images, due to the fact that they summarize the static scene content of entire clips, can facilitate browsing and retrieval on the basis of scene appearance, while object tracks can facilitate access on the basis of activity; in [4] a scheme is developed for summarizing video content using mosaics and moving object trajectories, and an approach exploiting the scheme for content-based retrieval and scene visualization is proposed. The problem of analyzing and managing complex dynamic scenes captured in a video sequence is addressed in [5], where the approach followed to summarize video datasets is based on the analysis of the trajectories of objects within them. In order to highlight the importance of visual material for exploiting foreign broadcast news and illustrating the utility of video browsing and summarization interfaces for comparing and contrasting viewpoints across cultural and language boundaries, in [6] systems and tools are developed to automatically detect, extract, and report high interest people, patterns, and trends in visual content from foreign news. A genetic algorithm for video segmentation and summarization was proposed in [7]; it partitions the video into a pre-specified number of segments and generates the boundary images that can be used for indexing and summarization. This approach is particularly useful for applications requiring incremental segmentation (e.g. streaming video over web) and adaptation to user access patterns.

2 Tools Used

In this section, a brief review of the tools we used (GCS and MPEG-7 descriptors) is provided.

- **GCS and TreeGCS** The GCS model [1] is an incremental self-organising clustering algorithm, an extension of SOM [13]. It is very useful tool for data visualization as it generates a mapping from a high dimensional input data to a lower (typically two-dimensional) space. An important property of this mapping is that it is topology preserving, i.e. similar input vectors are mapped on the same node or neighboring nodes. An advantage over SOM and most of the classical clustering algorithms (e.g. k-means) is that GCS does not require the number of clusters to be pre-specified in

advance but finds them automatically through a process of controlled growing and removal of nodes. Unlike SOM, where the clusters remain connected, GCS is able to form discrete clusters.

TreeGCS [14] is a hierarchical clustering algorithm that is based on GCS. It maps high dimensional input vectors onto a multi-depth two dimensional hierarchy that preserves the topological ordering of the input space. The tree is generated dynamically and adapts to the underlying GCS structure.

- **MPEG-7 descriptors** The MPEG-7 standard specifies a rich set of description structures for audio-visual (AV) content, which can be instantiated by any application to describe various features and information related to the AV content. The set of description structures that MPEG-7 standardises is very broad, and includes both general visual descriptors and domain specific ones; each application is responsible for selecting an appropriate subset to instantiate, according to the application functionality requirements.

  In our system, we used a subset of general visual descriptors [8], which are briefly reviewed below.

  - **Color descriptors** Color is one of the most widely used image and video retrieval features. The MPEG-7 standard includes five color descriptors which represent different aspects of the color and includes color distribution, spatial layout, and spatial structure of the color. The histogram descriptors capture the global distribution of colors. The dominant color descriptor represents the dominant colors used. The color layout descriptor captures the spatial distribution or layout of the colors in a compact representation. In our experiments we used *Dominant Color*, *Scalable Color*, *Color layout*, and *Color structure*.
  - **Texture descriptors** The image texture is one of the most important image characteristic in both human and computer image analysis and object recognition [8]. Visual texture is a property of a region in an image. There are two texture descriptors in MPEG-7: a homogeneous texture descriptor, and edge histogram descriptor. We used the *Edge histogram* descriptor.
  - **Shape descriptors** MPEG-7 supports region and contour shape descriptors. Object shape features are very powerful when used in similarity retrieval. We used *Region Shape* and *Contour Shape*.

3   The Proposed Architecture

As already mentioned, our approach for keyframes selection and video summarization is based on the MPEG-7 descriptors and the GCS neural algorithm. In the present version, the system is applied to MPEG-2 coded sequences, which are decoded in order to use the MPEG-7 eXperimentation Model (XM) [10].

A block diagram of the proposed system is shown in Fig. 1. After shot boundary detection, MPEG-7 descriptors are computed for all the I frames of the sequence, and the features obtained by the different descriptors for each frame are concatenated in a corresponding vector. GCS is then used to cluster the frames in each shot based on their feature vectors. Depending on the content of the shot, GCS forms different number of discrete clusters. For each of them, the frame closest to the centroid is selected as a

Fig. 1. Block diagram of the proposed system

| (a) | (b) |
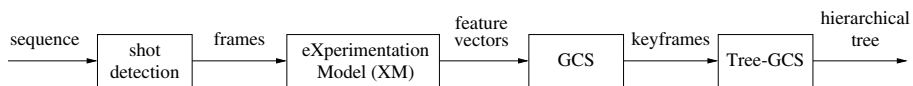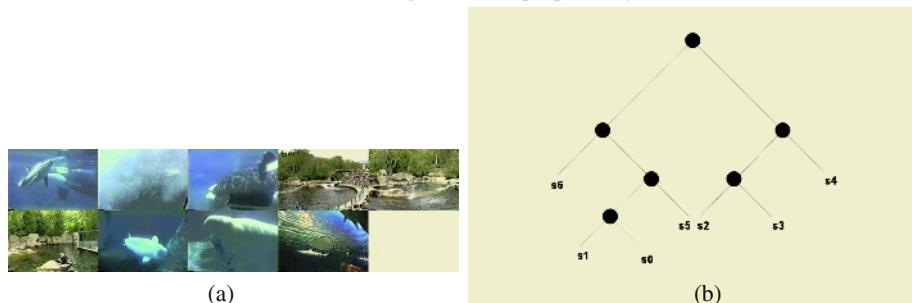
Fig. 2. (a)"True" keyframes for the *acqua* sequence ($s0$ and $s5$ have two keyframes); (b) a possible desired hierarchical description for the *acqua* sequence

keyframe. Because GCS preserves the topology, similar frames are mapped to neighboring neurons.

The selected keyframes are further clustered using TreeGCS to create a hierarchical view of the video sequence allowing the user to browse at different levels of content. The depth of the hierarchy and number of nodes in each level depend on the video content. Each node corresponds to a cluster of similar shots, and can be represented by one single keyframe chosen as described above. The bottom level nodes are associated with shots, that are mapped on a 2-dimensional GCS grid allowing efficient visualization and browsing.

4   Experimental Results

In this section we present some preliminary results which have been obtained applying our system to the *acqua* sequence [12]. It consists of 1281 frames taken in an aquarium and is organized in 7 shots.

The first two shots $s0$ and $s1$ are underwater scenes showing dolphins and are relatively similar to each other. The following three shots $s2$, $s3$ and $s4$ are taken outside and show a pool with dolphins jumping out of the water. The two final shots $s5$ and $s6$ are again underwater scenes; while $s5$ is rather similar to $s0$ and $s1$, $s6$ is different and shows sharks, ray fish and other fish. A good set of keyframes could be the one shown in Fig. 2a that was extracted manually by a human observer. A reasonable hierarchical description of this sequence could be the one shown in Fig. 2b. At the top level, a distinction should be made between underwater shots and the outdoor scenes. It would be nice to have shots $s0$, $s1$, and perhaps $s5$, grouped together, with shot $s6$ apart. The outdoor scenes are rather similar, with perhaps shot $s4$ slightly different from the other two as there is also a seal in the middle of the frame.

Before presenting the experimental results, it has to be noted that a quantitative performance evalutation, even if highly desirable, is very difficult. For example, the

Table 1. Experimental results using several MPEG-7 descriptors separately

descriptor	keyframes extraction	tree generation	descriptor	keyframes extraction	tree generation
Dominant Color	fair	good	*Edge Histogram*	good	fair
Scalable Color	fair	good	*Contour Shape*	poor	poor
Color Layout	fair	fair/good	*Region Shape*	fair	poor
Color Structure	good	fair			
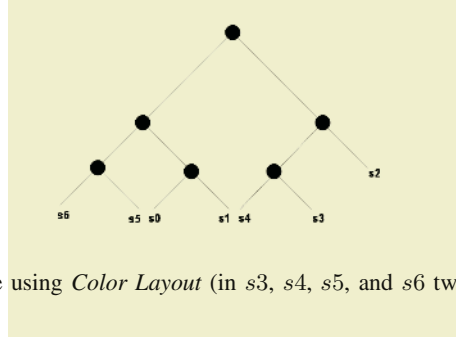
Fig. 3. Extracted keyframes and generated tree using *Color Layout* (in $s3$, $s4$, $s5$, and $s6$ two keyframes have been found)

"true" number of the keyframes for each shot and their exact position is subjective; indeed, in sequences with limited motion (as in our example) a temporal displacement of a few tens of frames may be widely acceptable. To evaluate the performance of the keyframe selection we compare with the ground truth sequence from Fig. 2a. The proposed ground truth tree in Fig. 2b is also subjective. We use it as a benchmark and report the quality of the generated tree as poor, fair, and good. A better approach would be to use metrics for tree comparison. However, the choice of these metrics is not trivial (see e.g. [9]) and has not been done at this preliminary level. Finally, the results may be slightly different at the different runs of our program. This is because GCS is sensitive to the order of input vectors and in our current implementation we apply the input vectors (corresponding to the frames) in random order. We report typical results. It should be noted that for the *aqua* video sequence the results are relatively stable across the various runs.

We first performed some tests using the MPEG-7 descriptors separately, in order to obtain a preliminary indication of their suitability for our classification task. The results are shown in Table 1. As it may be seen the color and edge descriptors were the best performing features while the shape descriptors were the worse. As an example, the keyframes and the tree obtained with *Color Layout* are shown in Fig. 3.

We also tested several combinations of descriptors: the five "best" (according to Table 1), all four color, and all seven descriptors, see Table 2. The results were compared with three other approaches: *VideoGCS, HistInt* and *Signatures*. *VideoGCS* [11] is our previous system where we used GCS and TreeGCS in a similar manner but the feature vectors were color histograms (64-bin normalized histograms in the YCbCr color space) computed for the dc-images of the I frames. *HistInt* [15] is a clustering based approach for keyframe selection that processes uncompressed video. The frames within

Table 2. Experimental results using selected combinations of MPEG-7 descriptors and comparison with three other approaches

descriptors	keyframes				tree
	generated	correct	missed	redundant	generation
all except *Contour Shape* and *Region Shape*	9.0	7.4	2.2	1.6	fair
all color descriptors	9.2	7.8	2.1	1.4	fair/good
all descriptors	9.4	8.1	1.7	1.4	fair/good
VideoGCS	8.6	8.0	0.9	0.6	poor
HistInt	4.0	4.0	5.0	0.0	n.a.
signatures	36.0	9.0	0.0	27.0	n.a.

Fig. 4. Extracted keyframes using all the descriptors (in $s3$, two keyframes have been found). The generated tree is not reported here, as it coincides with the "desired" one, shown in Fig. 2b

each shot are first clustered using a k-means like algorithm. Two keyframes are extracted for clusters with high intercluster variance (the closest and the farthest to the centroid); frames with large deviations from the average shot luminance are selected as keyframes as well. *Signatures* [16] is also a color histogram-based approach processing uncompressed video. It defines a new feature based on color histograms and then applies an agglomerative hierarchical clustering that merges clusters depending on cluster variance and temporal distance. The results for *HistInt* and *Signatures* are taken from [16]. We report the number of generated, correctly generated (*correct*), missed, and redundant keyframes. Fig. 4 and 5 show some sets of keyframes and generated trees using all descriptors and all color descriptors. The three combinations of the MPEG-7 descriptors perform almost equally well and outperform *HistInt* and *Signatures* in the keyframe selection. *HistInt* tends to generate too many keyframes while *Signatures* generates a very compact summarization but there are many misses and redundancies. VideoGCS was found to be the best approach for keyframe selection, i.e. the use of the MPEG-7 descriptors did not improve performance in comparison to color histograms. The best hierarchical representation, however, was achieved using MPEG-7 descriptors - the use of all three combinations of MPEG-7 descriptors resulted in better trees than the ones generated using color histograms.

5 Conclusions

We presented an approach for key frame selection and video browsing based on the MPEG-7 descriptors and the Growing Cell Structures (GCS) neural algorithm. According to the first preliminary tests, it seems to be able to generate a hierarchical structure which closely resembles the semantic content of the video, probably thanks to the richness and effectiveness of the MPEG-7 descriptors, and the adaptability and self-learning capability of the GCS.
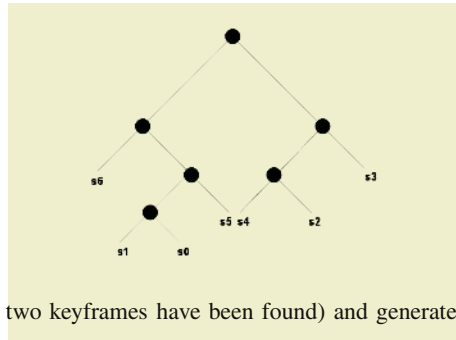
Fig. 5. Extracted keyframes (in $s4$, $s5$ and $s6$, two keyframes have been found) and generated tree using the 4 color descriptors

However, a lot of work is still to be done. First of all, much more extended tests have still to be performed, with reference both to the number and type of sequences, and to other systems, in order to be able to compare the performances. Moreover, a less subjective method should be found both to define the "ground truth", and to evaluate the quality of the results. Indeed, a metrics for quantifying the distance between desired and actual keyframes, and between desired and actual tree, would be useful to objectively evalute the performances of the system for what concerns both keyframes extraction and tree generation.

References

1. B. Fritzke, "Growing Cell Structures - a Self-Organizing Network for Unsupervised and Supervised Learning,", Neural Networks, vol.7(9), pp.1441-1460, 1994.
2. S. Lu, I. King and M. R. Lyu, "Video Summarization by Video Structure Analysis and Graph Optimization", In Proc. IEEE International Conference on Multimedia and Expo (ICME'04), PP. 27-30, June 2004, Taiwan.
3. A. Divakaran, K. Peker, R. Radhakrishnan, Z. Xiong and R. Cabasson, "Video Summarization using MPEG-7 Motion Activity and Audio Descriptors," Video Mining, eds. A. Rosenfeld, D. Doermann and D. DeMenthon, Kluwer Academic Publishers, 2003.
4. Pope, Rakesh Kumar, Harpreet Sawhney, and Charles Wan Sarnoff Corporation, "Summarizing Video Content for Retrieval and Visualization", In Proc. Thirty-Second Asilomar Conference on Signals, Systems & Computers, pp. 915-919, 1998.
5. A. Stefanidis, P. Partsinevelos, P. Agouris, P. Doucette, "Summarizing Video Datasets in the Spatiotemporal Domain", IEEE Press, pp. 906-912, Sept. 2000, London, UK.
6. M. Christel, P. Duygulu, A. Hauptmann, D. Ng, J. Yang, "Extensible News Video Information Exploitation", http://www.informedia.cs.cmu.edu/arda/vaceII.html
7. P. Chiu, A. Girgensohn, W. Polak, E. Rieffel, L. Wilcox, "A Genetic Algorithm for Video Segmentation and Summarization", Proceedings of ICME (IEEE International Conference on Multimedia and Expo), August 2000, New York City, USA.
8. Manjunath B., Ohm J., Vasudevan V., Yamada A., "Color and Texture Descriptors", IEEE Transactions on circuits and systems for video technology, V. 11, No. 6, June 2001, 703-715.
9. D. Hillis, C. Moritz, and B. Mable, "Molecular Systematics", Sinauer Pub., Boston, 1996.
10. http://phalanx.lis.e-technik.tu-muenchen.de/research/bv/topics/mmdb/e_mpeg7.html
11. I. Koprinska, J. Clark, and S. Carrato, "Video GCS - A clustering-based system for video summarization and browsing," in *Proc. 6th COST 276 Workshop*, (Thessaloniki (Greece)), pp. 34–40, May 2004.

12. http://www.cs.sfu.ca/ mark/ftp/AcmMM00/
13. T. Kohonen, "Self-Organizing Maps", 2nd ed., Springer 1997.
14. V.J. Hodge and J. Austin, "Hierarchical Growing Cell Structures: TreeGCS," IEEE Trans Know & Data Eng, v. 13(2), pp. 207-218, 2001.
15. A.M. Ferman and A.M. Tekalp, "Efficient Filtering and Clustering Methods for Temporal Video Representation and Visual Summarization," J. Visual Com. & Image Rep., vol. 9, pp.336-351, 1998.
16. M. S. Drew and J. Au, "Video Keyframe Production by Efficient Clustering of Compressed Chromaticity Signa-tures," ACM Multimedia, 2000.

Using Relevance Feedback to Learn Both the Distance Measure and the Query in Multimedia Databases

Chotirat Ann Ratanamahatana and Eamonn Keogh

Dept. of Computer Science & Engineering, Univ. of California, Riverside, CA 92521 USA
{ratana,eamonn}@cs.ucr.edu

Abstract. Much of the world's data is in the form of time series, and many other types of data, such as video, image, and handwriting, can easily be transformed into time series. This fact has fueled enormous interest in time series retrieval in the database and data mining community. However, much of this work's narrow focus on efficiency and scalability has come at the cost of usability and effectiveness. Here, we introduce a general framework that learns a distance measure with arbitrary constraints on the warping path of the Dynamic Time Warping calculation. We demonstrate utility of our approach on both classification and query retrieval tasks for time series and other types of multimedia data, then show that its incorporating into the relevance feedback system and query refinement can further improve the precision/recall by a wide margin.

1 Introduction

Much of the world's data is in the form of time series, leading to enormous interest in time series retrieval in the database and data mining community. However, most of previous work has utilized the Euclidean distance metric because it is very amenable to indexing [1], [2]. However, there is increasing evidence that the Euclidean metric's sensitivity to discrepancies in the time axis makes it unsuitable for most real world problems [3], [4], [5], [6]. This fact appears to be almost unnoticed because, unlike their counterparts in information retrieval, many researchers in the database/data mining community evaluate algorithms without considering precision/recall or accuracy [7]. In this work, we introduce a distance measure and show its utility with comprehensive experiments. Despite its potential weakness that it requires some training or human intervention to achieve its superior results, we can use the classic relevance feedback technique to achieve this end.

1.1 The Ubiquity of Time Series Data

In this section, we wish to expand the readers' appreciation for the ubiquity of time series data. Rather than simply list the traditional application domains, we will consider some less obvious applications that can benefit from efficient and effective retrieval.

Video Retrieval: Video retrieval is one of the most important issues in multimedia database management systems. Generally, research on content-based video retrieval

represents the content of the video as a set of frames, leaving out the temporal features of frames in the shot. However, for some domains, it may be fruitful to extract time series from the video, and index just the time series (with pointers back to the original video). Figure 1 shows an example of a video sequence transformed into a time series. One obvious advantage of this representation is the huge reduction in dimensionality, enhancing the ease of storage, transmission, analysis, and indexing; it is also much easier to make the time series representation invariant to distortions in the data.
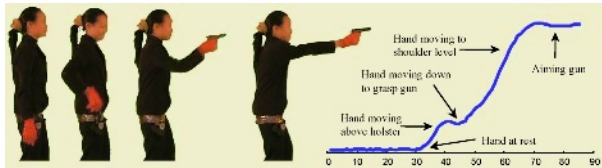
Fig. 1. Stills from a video sequence; the right hand is tracked, and converted into a time series

Image Retrieval: Image Retrieval has become increasingly crucial in our information-based community. For some specialized domains, it can be useful to convert images into "pseudo time series." Consider Figure 2, where we convert an image of a leaf into a time series by measuring the local angle of its perimeter's trace. The utility of such a transform is similar to that for video retrieval, especially in natural images, where we can more easily handle the natural variability of living creatures (scale, offset, and rotation invariance) in the time domain with dynamic time warping [8].
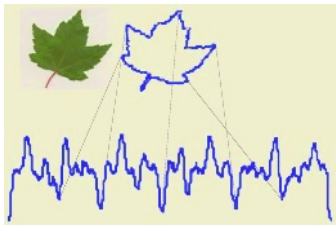
Fig. 2. Many image indexing/classification/clustering can be solved more effectively and efficiently after converting the image into a "pseudo time series"

Handwriting Retrieval: While the recognition of online handwriting [9] may be largely regarded as a solved problem, the problem of transcribing and indexing existing historical archives remains a challenge.
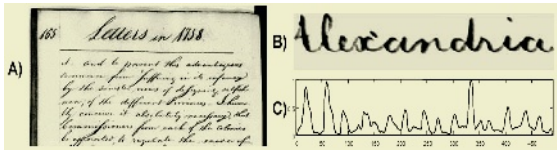
Fig. 3. A) An example of handwritten text by George Washington. B) A zoom-in on "Alexandria", after slant removal. C) A variety of techniques exist to convert 2-D handwriting into a time series; in this case, the projection profile is used

The problem of indexing historical archives is difficult, because unlike the online handwriting problem, there is no pen-acceleration information. In addition, the archives may be degraded and stained. For example, Figure 3.A) shows an example of text written by George Washington, which is all but illegible to modern readers with little experience with cursive writing. Many off-line handwritten document image-processing algorithms have recently been proposed in the interest of word recognition and indexing [10], [11], mostly without associating with problem with time series. While handwriting is not a time series, there exist several techniques to convert handwriting to time series. Recent work suggests that this representation may still allow the high precision in indexing historical archives while simplifying the problem from 2-dimentional to 1-dimentional domain [12]

1.2 Existing Work on Time Series Retrieval

The explosion of interest in time series indexing in the last decade has been extraordinary, but the majority of the work has focused on the Euclidean distance, which assumes linear mappings between the query and the candidate time series. However, recent work has demonstrated that this similarity model generally does not work well for many real-world problems, where variability in the time axis is always present. This problem of distortion in the time axis can be addressed by Dynamic Time Warping (DTW), a distance measure that has long been known to the speech processing community [13], [14]. This method allows non-linear alignments between the two time series to accommodate sequences that are similar but out of phase. The superiority of DTW over Euclidean distance for classification and indexing has been demonstrated by several authors on a diverse set of domains, including bioinformatics, music retrieval, handwritten document archives, biometrics, chemical engineering, industry, and robotics [3], [4]. Our approach takes this recent work on DTW as its starting point, then fine-tune the algorithm, for a particular domain, and even a particular query, by selectively limiting the amount of warping we allow along various parts of the query. As we will demonstrate, by selectively limiting the amount of warping allowed, we can actually improve the accuracy of DTW, as well as the indexing performance. Due to space limitations, please refer to the DTW background and related work in [15].

2   The Ratanamahatana-Keogh Band (*R-K Band*)

The 'global constraint' of DTW has been almost universally applied to DTW, primarily to prevent unreasonable warping and to speed up its computation. However, surprisingly little research has looked at discovering the best shape and size of the window. Most practitioners simply use one of the well-known bands, e.g., Sakoe-Chiba Band [14] or Itakura Parallelogram [16], proposed in the context of speech recognition several decades ago. In addition, there are several widespread beliefs about DTW, which have been proven to be false by our extensive experiments on wide variety of datasets [17]. The motivation for our work has sparked from this finding that the effect of the window shape/size on accuracy is very substantial, and is strongly domain dependent. Our ideal solution is then to find an optimal band for a

given problem that will potentially improve the accuracy. We recently proposed a new representation, the *Ratanamahatana-Keogh Band (R-K Band)* [15], which allows arbitrary shaped constraints.

2.1   A General Model of Global Constraints

To represent a warping window, we can define it in terms of a vector R, $R_i = d$, where $0 \leq d \leq m$, $1 \leq i \leq m$ and R_i is the height above the diagonal in the y direction, as well as the width to the right of the diagonal in the x direction. The classic Euclidean distance can be defined in terms of $R_i = 0$; $1 \leq i \leq m$; only the diagonal path is allowed. More generally, we can define any arbitrary constraint with a vector R. Figure 4 illustrates some examples of our *R-K Bands*. An interesting and useful property of our representation is that it also includes the Euclidean distance and classic DTW as special cases.
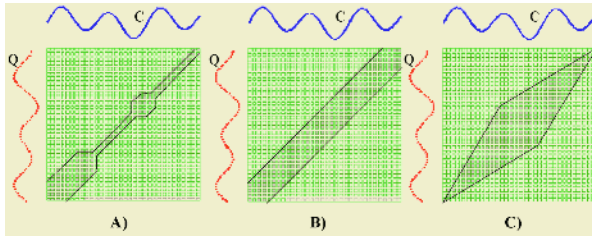
Fig. 4. We can use R to create arbitrary global constraints. A) We can use R to specify all existing global constraints, including the Sakoe-Chiba Band B) and the Itakura Parallelogram C)

We can exploit the *R-K Bands* for both classification and indexing (query retrieval) problems, depending on the task at hand. In particular,

- for classification, we can use a different *R-K Band* for each class; we denote the band learned for the cth class, as the *R-K$_c$ Band.*
- for indexing, we can use *one R-K Band* that maximizes the trade off between efficiency and precision/recall.

Having introduced an *R-K Band*, we can represent any arbitrary warping windows. However, *discovering* the optimal *R-K Band* still remains a question. In some cases, it maybe is possible to manually construct the bands, based on domain knowledge. Unfortunately, our preliminary attempts to manually construct *R-K Bands* met with limited success, even for simple toy problems. Furthermore, since the number of possible *R-K Bands* is exponential, exhaustive search over all possibilities is clearly not an option. We will show how we can *learn* the high-quality bands automatically from data.

2.2   Learning Multiple *R-K$_c$ Bands* for Classification

While it is generally not possible to handcraft accurate *R-K Bands,* it *is* possible to pose the problem as a search problem. Using generic heuristic hill-climbing search

techniques, we can perform both forward and backward searches. The forward search starts with the initial Sakoe-Chiba band (uniform) of width 0 (Euclidean), and the backward search starts from the uniform band of the maximum width m, above and to the right of the diagonal. The searches are complete when one of the following is true:-No improvement in accuracy can be made; the width of the band reaches m for the forward search and 0 (Euclidean) for the backward search or; each section of the band (after recursively cut the portion in half) reaches some threshold. The utility of our *R-K_c Bands* for classification has been extensively shown in [18].

2.3   Learning One *R-K Band* for Indexing

In addition to creating *R-K_c Bands* for classification, we can learn one single *R-K Band* for indexing or query retrieval. The one-band learning algorithm is very similar to the multiple-band learning in the previous section, except that we only maintain one single band representing the whole problem and that we measure the precision/recall instead of the accuracy. Utilizing an *R-K Band* for an indexing problem has been shown in [18] to improve both precision and recall by a wide margin. However, an *R-K Band* needs to be learned from a training data, which may not be practical or available in many circumstances. To resolve this problem, we can build a training data through relevance feedback system, with some help from a user in identifying the positive and negative examples to the system, which is demonstrated in the next section.

3   Relevance Feedback

In text-mining community, relevance feedback is well known to be effective method to improve the query performance [4], [19], [20]. However, there has been much less research in non-text domains. In section 1.1, we have introduced time series as an alternative representation of some multimedia data. This section shows how to incorporate that technique into relevance feedback system using our proposed *R-K Band*.

3.1   Query Refinement

Relevance feedback methods attempt to improve performance for a particular informational need by refining the query, based on the user's reaction to the initial retrieved documents/objects. Working with time series retrieval is rather similar to the text retrieval; a user can draw or provide an example of a query and retrieve the set of best matches' retrieval of data. Once the user ranks each of the results, a query refinement is performed such that a better-quality query is produced for the next retrieval round. In our system, the user is asked to rank each result in a 4-point scale, based on relevance to their informational needs, which are converted into appropriate weights being used in query refinement processes (averaging the positive results with current query). However, averaging a collection of time series that are not perfectly

time-aligned is non-trivial and DTW is needed [21]; Hierarchically, each pair of time series are averaged based on weights and warped alignments.

3.2 *R-K Band* Learning in Relevance Feedback

We will empirically demonstrate that our proposed *R-K* Band combined with the query refinement can improve precision and recall of retrieval. Table 1 shows our algorithm.

Table 1. *R-K Band* learning with Relevance Feedback algorithm

```
Algorithm RelFeedback(initial_query)
```

1.   **Repeat** until all rankings are positive.
2.   Show the 10 best matches to the current query to the user.
3.   Let the user rank how relevant each result is.
4.   According to the ranking, accumulatively build the training set; positive result → class 1, negative result → class 2.
5.   Learn a single envelope that represents the given training data.
6.   Generate a new query, by averaging (with DTW) the positive results with the current query according to their weights.
7.   **end;**

In the first iteration, given a query, the system uses the initial *R-K Band* (size 0), retrieves the 10 nearest neighbors, and then shows them to the user (line 1). After ranking, the positive and negative responses are noted and collected as a training data (lines 3-4). The algorithm uses this training data to learn an *R-K Band* that best represents the positive objects in the training set while being able to correctly differentiate the positive from the negative instances (line 5). The training data will be accumulated during each round, thus producing progressively finer results. The process is complete when only positive feedbacks are returned or the user abandons the task. In our experiments, we consider 3 multimedia datasets, shown in Table 2, in our relevance feedback technique (complete datasets details are available in [18]). To evaluate our framework, we measure the precision and recall for each round of the relevance feedback retrieval. Since we only return the 10 best matches to the user and we would like to measure the precision at all recall levels, we purposely leave only 10 relevance objects of interest in all the databases.

Table 2. Datasets used in the relevant feedback experiment

Dataset	Number of Classes	Time Series Length	Number of items
Gun	2	150	200
Leaf	6	350	442
Wordspotting	4	100	272

We then measure the performance of our relevance feedback system with the precision-recall plot from each round of iteration. Figure 5 shows the precision-recall curves of the three datasets for the first five iterations of relevance feedback. Our experiments illustrates that each iteration results in significant improvement in both precision and recall.
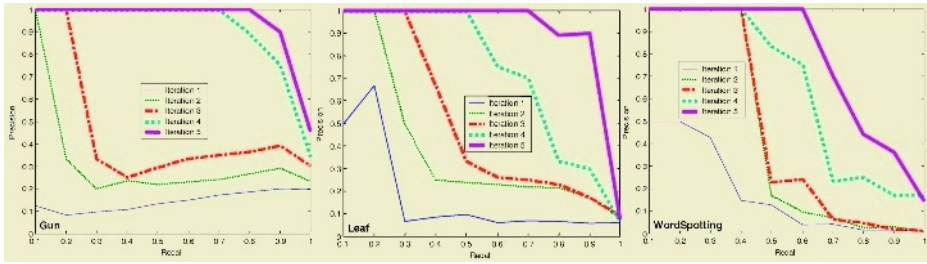
Fig. 5. The precision-recall plot for the 3 datasets with 5 iterations of relevance feedback

4   Conclusion

We have introduced a framework for both classification and time series retrieval. The Ratanamahatana-Keogh Band (R-K Band) allows for any arbitrary shape of the warping window in DTW calculation. With our extensive evaluation, we have shown that our framework incorporated into relevance feedback can reduce the error rate in classification, and improve the precision at all recall levels in video and image retrieval.

Acknowledgments

References

1. Agrawal, R., Lin, K. I., Sawhney, H. S., & Shim, K. (1995). Fast similarity search in the presence of noise, scaling, and translation in times-series databases. VLDB, pp. 490-501.
2. Faloutsos, C., Ranganathan, M., & Manolopoulos, Y. (1994). Fast subsequence matching in time-series databases. In Proc. ACM SIGMOD Conf., Minneapolis. pp. 419-429.
3. Aach, J. & Church, G. (2001). Aligning gene expression time series with time warping algs. Bioinformatics. 17: 495-508.
4. Bar-Joseph. Z., Gerber, G., Gifford, D., Jaakkola T & Simon. I. (2002). A new approach to analyzing gene expression time series data. RECOMB, pp. 39-48.
5. Kadous, M.W. (1999). Learning comprehensible descriptions of multivariate time series. ICML: 454-463.
6. Schmill, M., Oates, T., & Cohen, P. (1999). Learned models for continuous planning. In 7th International Workshop on Artificial Intelligence and Statistics.
7. Keogh, E. and Kasetty, S. (2002). On the Need for Time Series Data Mining Benchmarks: A Survey and Empirical Demonstration. In the 8th ACM SIGKDD. pp. 102-111.
8. Gandhi, A. (2002). Content-based Image Retrieval: Plant species Identification. MS Thesis, Oregon State U.
9. Jain, A.K. & Namboodiri, A.M. (2003). Indexing and Retrie-val of On-line Handwritten Documents. ICDAR: 655-659.
10. Kavallieratou, E., Dromazou, N., Fakotakis, N., & Kokkinakis, G. (2003). An Integrated System for Handwritten Document Image Processing. IJPRAI, (17)4, pp. 617-36.

11. Tomai, C.I., Zhang, B., & Govindaraju, V. (2002). Transcript Mapping for Historic Handwritten Document Images. IWFHR'02.

12. Rath, T. & Manmatha, R. (2003). Word image matching using dynamic time warping. CVPR: II, 521-27.

13. Kruskall, J. B. & Liberman, M. (1983). The symmetric time warping algorithm: From continuous to discrete. In Time Warps, String Edits and Macromolecules.

14. Sakoe, H. & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acous., Speech, and Signal Proc., ASSP-26: 43-49.

15. Ratanamahatana, C.A. & Keogh, E. (2004). Making time-series Classification More Accurate Using Learned Constraints. SDM International Conference, pp. 11-22.

16. Itakura, F. (1975). Minimum prediction residual principle applied to speech recognition. IEEE Trans. Acoustics, Speech, and Signal Proc., Vol. ASSP-23, pp. 52-72

17. Ratanamahatana, C.A. & Keogh, E. Three Myths about Dynamic Time Warping Data Mining. SIAM Int'l Conference on Data Mining (SDM), Newport Beach, CA, (2005).

18. Ratanamahatana, C.A. (2005). Improving Efficiency and Effectiveness of Dynamic Time Warping in Large Time Series Databases. Ph.D. Dissertation, Univ. of Calif., Riverside.

19. Rocchio, J.J. (1971). Relevance Feedback in Information Retrieval. The SMART Retrieval System: Experiments in Automatic Document Processing. pp. 313-323.

20. Salton, G., & Buckley, C. (1990). Improving retrieval performance by relevance feedback. JASIS41(4),.288-297.

21. Gupta, L., Molfese, D., Tammana, R., and Simos, P. (1996). Nonlinear Alignment and Averaging for Estimating the Evoked Potential. IEEE Trans. on Biomed Eng(43)4.

Multi-level Semantic Analysis for Sports Video

Dian W. Tjondronegoro[1] and Yi-Ping Phoebe Chen[2]

[1] School of Information Systems, Queensland University of Technology
dian@qut.edu.au
[2] School of Information Technology, Deakin University
Melbourne, Australia
phoebe@deakin.edu.au

Abstract. There has been a huge increase in the utilization of video as one of the most preferred type of media due to its content richness for many significant applications including sports. To sustain an ongoing rapid growth of sports video, there is an emerging demand for a sophisticated content-based indexing system. Users recall video contents in a high-level abstraction while video is generally stored as an arbitrary sequence of audio-visual tracks. To bridge this gap, this paper will demonstrate the use of domain knowledge and characteristics to design the extraction of high-level concepts directly from audio-visual features. In particular, we propose a multi-level semantic analysis framework to optimize the sharing of domain characteristics.

1 Introduction

Sports video indexing approaches can be categorised based on the two main levels of video content: low-level (perceptual) *features* and high-level *semantic* annotation [1-4]. The main benefits of *feature-based indexing* techniques: 1) they can be fully automated using feature extraction techniques such as image and sound analysis and 2) users can do similarity search using certain feature characteristics such as the shape and colour of the objects on a frame or the volume of the sound track. However, feature-based indexing tends to ignore the semantic contents whereas users mostly want to search video based on the semantic rather than the low-level characteristics. Moreover, there are some elements beyond perceptual level (widely known as the *semantic gap*) which can make low-level feature based indexing tedious and inaccurate. For example, users cannot always describe the visual characteristics of certain objects they want to view for each query. In contrast, the main benefit of *semantic-based indexing* is the ability to support more natural queries. However, semantic annotation is still generally time-consuming, and often incomplete due to limitations of manual supervision and the currently available techniques for automatic semantic extraction.

In order to bridge the semantic gap between low-level features and high-level semantic, a mid-level feature layer can be introduced [5]. The main purpose is to separate (sports) specific knowledge and rules from the low-level feature extraction; thus making it less domain-specific and robust for different sports. However, the use of mid-level features has not been optimized since most of the current semantic detection schemes still rely heavily on domain knowledge (see [6] as example).

It should be noted that some sport events can be detected using a more generalized knowledge while some have to use very specific rules. Thus, rather than immediately extracting domain-specific events, we can firstly detect generic key events based on, for example, replay scene, excited crowd/commentator, and whistle. Thus, stretching the semantic layer will allow us to generalize common properties of sports video, while minimizing the use of domain knowledge as far as possible. The longer that we can stay away from committing into a particular sport domain, the more robust that the designed algorithms and framework would be.
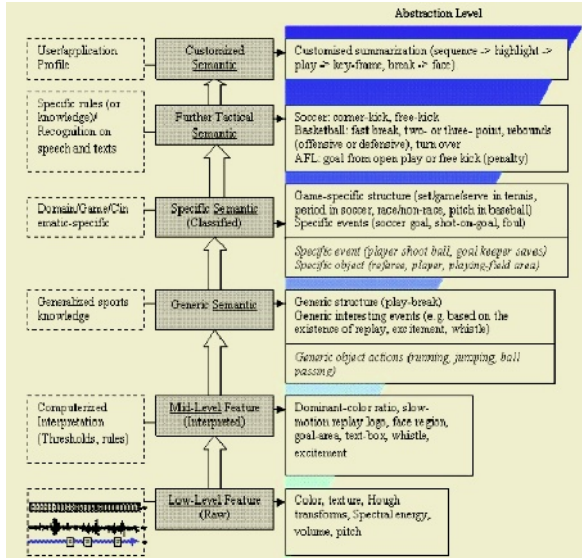
Fig. 1. Multi-level Semantic Analysis in Sports Video Framework

Hence, we propose to extend semantic layer (illustrated in Figure 1) by dividing *semantic features* into:

- *Generic semantic*. This layer can be detected using generalized knowledge on the common features shared by most of sports videos. For example, *interesting events* in sport videos are generally detectable using generic mid-level features such as whistle, excited crowd/commentator, replay scene, and text display.
- *Specific/classified semantic*. This layer is usually detected using a set of more specific characteristics in different sports based on domain-knowledge. The main purpose is to classify interesting events into domain-specific events such as soccer goal in order to achieve a more sophisticated retrievals and browsing.
- *Further tactical semantic*. This layer can be detected by focusing on further-specific features and characteristics of a particular sport. Thus, tactical semantic need to be separated from specific semantic due to its requirement to use more complex and less-generic audio-visual features. For example, *corner kick* and *free-kick* in soccer needs to be detected using specific analysis rules on soccer-field and player-ball motion interpretation (such as in [7]). Playing-field analysis and motion interpretation are less generic than excitement and break shot detection. For instance, excitement detection algorithm can be shared for different sports by adjust-

ing the value of some parameters (or thresholds), whereas the algorithm to detect corner (playing-field) in soccer will not be applicable for tennis due to the difference between soccer and tennis fields.

− *Customized semantic*. This layer can be formed by selecting particular semantics using user or application usage-log and preferences. For example, a sport summary can be structured using an integrated highlights and play-breaks using some text-alternative annotations [8].

The following describes an example of how the semantic layers are connected. During the extraction of generic events, play-breaks which can contain interesting events are detected. Statistics of mid-level features in each play-break can be used to construct heuristic-rules that classify the events contained into specific domain such as soccer *goal* (i.e. each play-break contain zero/more highlight events). When soccer goal is detected; if the play scene switch from left-goal-area to right-goal-area in less than 3-4 seconds, then further-tactical event of *counter-attack* is detected. Customized semantic is constructed by combining particular events based on user/application requirements.

One of the main benefits from using the proposed extended semantic layer is to achieve clearer boundaries between semantic analysis approaches; thus enabling a more systematic framework for designing event detection. For example, we can categorize goal detection in soccer [6] and basketball [9] into 'specific-semantic analysis', whereas the work presented to detect corner kick and free-kick in soccer [7] should be categorized under 'further-tactical events analysis'. Using this framework, we can achieve more organized reviews, benchmark, and extensions from current solutions. This paper focuses on the extraction of generic and specific semantic.

2 Generic Semantic Extraction

Play-Break

A sport video is usually started with a pre-game scene which often contains introductions, commentaries and predictions. After a game is started, the game is played until there is an event which causes break. After a play being stopped, it will be resumed until being stopped again. This play-break phase is iterative until the end of the game which is then followed by a post-game scene which has similar content to pre-game.

Duration of camera-views have been used successfully for play-break segmentation (such as in [6]) since a long global shot (interleaved shortly by other shots) is usually used by broadcasters to describe attacking play which could end because of a goal. After a goal is scored, zoom-in and close-up shots will be dominantly used to capture players and supporters celebration. We have extended this approach by adding replay-based correction to design the algorithm for play-break segmentation which is effective and reliable for soccer, AFL and basketball video. Figure 2 shows the various scenarios on how a replay scene can fix the boundaries of play-break sequences. In addition to these scenarios, it is important to note that most broadcasters would play a long replay scene(s) during a long break, such as goal celebration in soccer or timeout in basketball, in order to keep viewers' attention. However, some broadcasters insert some advertisements (*ads*) in-between or during the replay. Thus, to obtain the correct length of the total break, these types of ads should not be removed; at least the total length of the ads has to be taken into account.
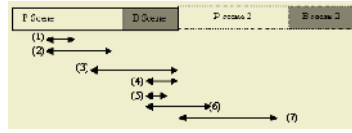
Fig.2. Locations of Replays in Play-breaks

Based on these scenarios, an algorithm to perform replay-scene based play-break segmentation has been developed.

---

(1) If [R strict_during P] & [(R.e – P.e) >= dur_thres]
    B.s = R.s; B.e = R.e; Create a new sequence where [P_2.s = R.e+1] & [P_2.e P.e]

(2) If [R strict_during P] & [(R.e – P.e) <= dur_thres]
    P.e = R.e; B.s = R.e+1
(3) If [R meets B] & [R.s < P.e]
    P.e = R.s
(4-5) If [R during B] & [R meets B]) OR (If [R strict_during B])
    No processing required
(6) If [R during B] & [(R.e – P_2.s) >= dur_thres]
    B.e = R.e; Amend the neighbor sequence: [P_2.s = R.e+1]
(7) If [R during P_2] & [(R.e – P_2.s) >= dur_thres]
    Attach sequence 2 to sequence 1 (i.e. combine seq 1 and seq 2 into 1 sequence)

---

where, If A *strict_during* B, (A.s > B.s) & (A.e < B.e); if A *during* B, (A.s > B.s & A.e <= B.e) OR (A.s >= B.s & A.e < B.e); If A *meets* B, A.e = B.e.

Table 1 depicts the performance of the play-break segmentation algorithm on 2 soccer videos which were recorded from Champions League: Milan-Depor (S1)**,** Madrid-Bayern (S2). S1-1 is video1 first half while S1-2 is video2 second half. In this table, *RC*=Replay-based Correction for play-break segmentation, *PD*=perfectly detected, *D*=detected, *M*=missed, *F*=false detection, *Tr* =Total number in Truth, *Det* =Total Detected, *RR*=Recall Rate, *PR*=Precision Rate, and *PDdecr*=perfectly detected decrease rate if *RC* is not used; where: *Tru=PD+D+M*, *Det=PD+D+F*, *RR=* (PD+D+M)/Tru*100%, PR=(PD+D)/Det*100%, and PD_Decr=(PD-D)/PD*100%.

The results have confirmed that *RC* improves the play-break segmentation performance. It is due to the fact that many (if not most) replay scenes use global (i.e. play) shots. This is shown by all *PD_decr, RR,* and *PR* since *RC* always improves all of these performance statistics. In particular, the *RR* and *PR* for soccer 1-1 with *RC* are 100% each but they are reduced to below 50% without *RC*. In soccer 1-1 without *RC*, the *PD* dropped from 49 to 12 (i.e. 75% worse) whereas *M* increases from 0 to 25 and *F* increases from 0 to 5. This is due to the fact that soccer1 video contains many replay scenes which are played abruptly during a play, thereby causing a too-long play scene and missing a break.

Generic Key Events

In Table 2, we have demonstrated the benefits of using three features: whistle, excitement and text, in order to detect generic key events (see [10] for more details). Based on this table, it is clear that whistle detection is very fast, but it can only localize 20 to 30% of the total highlights which are mostly caused by foul and offside. By

combining whistle and excitement, users only need to wait slightly longer to detect 80% to 90% of the highlights since excitement can locate most of exciting events, such as *good display of attacking/defending, goal,* and *free kick.* Excitement detection is particularly effective to localize goal highlights due to the massive amount of excitement during the start of a good attack which often leads to the goal itself. When whistle and text detection are combined, the number of highlights detected will only slightly increase while the waiting-period is longer than using excitement. This is due to the fact that visual features are generally more expensive computationally than audio features. Text detection is needed to localize *start of a match*, *goal* and *shot on goal*, as well as confirming offside and foul events.

3   Classified Semantic

While generic key events are good for casual video skimming, domain-specific (or classified) highlights will support more useful browsing and query applications. Unlike the current heuristic rules scheme for highlight detection, our system does not use manual knowledge. This is achieved by using a semi-supervised training on 20 samples from different broadcasters and matches for each highlight to determine the characteristics of play-breaks containing different highlights and no highlights. It is semi-supervised as we manually classify the specific highlight that each play-break sequence (for training) contains. Moreover, the automatically detected play-break boundaries and mid-level features locations within each play-break (such as excitement) are manually checked to ensure the accuracy of training.

We have demonstrated that the statistics of mid-level features (e.g. close-up and excitement ratio) within a play-break sequence can be used for classification of sport domain events. At this stage we have successfully detected highlights and classify them into: goal, shoot, foul, or non in soccer; goal, behind, mark, tackle, or non in AFL (see [11] for soccer and AFL experiment). In this section, we will discuss the experiment to detect *goal*, *free-throw*, *foul*, or *timeout* in basketball.

Using the trained statistics, heuristic rules for basketball events can be identified:
- *Timeout*: compared to any other highlight, this event has the longest possible duration (of play-break), mostly containing break shots. Consequently, there are many slow motion replays and close-up shots to keep viewers' interest. This event usually contains the least portion of near-goal play.
- *Goal/Free throw*: compared to foul, goal and free throw contain a large amount of near goal and replay scene is never played. To differentiate goal from free throw, the statistics show that goal contains less close-up shots; thus play shot is more dominant. However, free throw usually contains longer near-goal and duration.
- *Foul:* when a sequence is less likely to contain goal, free throw or goal, foul can usually be detected. Moreover, the close-up is more than goal and free throw but less than timeout; play and break ratio is almost equal; at least one replay scene; and the excitement is less than goal but more than free throw.

The statistical-driven heuristic rules have been tested to show its effectiveness and robustness using a large dataset of basketball videos (as described in Table 3) where each sport is recorded from different broadcasters, competition, match, and/or stage of

competition. Table 4 depicts the performance of highlight events classification in basketball videos. The experimental results show high *RR* and reasonable *PR*.

Table 1. Play-break Segmentation Performance in Soccer Videos

Video	Soccer Play-break detection								
	PD	D	M	F	Tru	Det	RR	PR	PD_decr
S1-1 (*RC*)	49	0	0	0	49	49	100.00	100.00	
S1-1	12	12	25	5	49	54	48.98	44.44	75.51
S1-2(*RC*)	53	0	0	1	53	54	100.00	98.15	
S1-2	36	10	7	1	53	54	86.79	85.19	32.08
S2-1(*RC*)	54	1	1	12	56	68	98.21	80.88	
S2-1	53	2	1	12	56	68	98.21	80.88	1.85
S2-2(*RC*)	58	1	0	7	59	66	100.00	89.39	
S2-2	55	4	0	7	59	66	100.00	89.39	5.17

Table 2. Generic Events Detection Results (#H: number of highlights, Time is in minute)

Sample Video	Total Highlights	Automatically Detected Highlights							
		Whistle Only (W)		Whistle + Text (W+T)		Whistle + Excitement (W+E)		W+T+E	
		#H	Time	#H	Time	#H	Time	#H	Time
Soccer1 (40mins)	**62**	**11**	1.7	**13**	37.1	**54**	22.9	**56**	58.2
Soccer2 (20mins)	**24**	**7**	0.7	**8**	24.8	**22**	10.6	**23**	35.4
Soccer3 (20mins)	**40**	**11**	0.7	**11**	26.7	**39**	8.8	**39**	35.5
Soccer4 (20 mins)	**22**	**2**	0.9	**3**	18.1	**21**	8.9	**22**	19
Rugby1 (20 mins)	**34**	**18**	0.9	**20**	20.6	**25**	10.9	**27**	29.9
Rugby2 (17 mins)	**21**	**8**	0.7	**9**	16.9	**18**	9.6	**19**	18.5
Basketball (15 m)	**37**	**7**	0.8	**12**	14.6	**30**	7.9	**35**	21.9
Tennis (20 mins)	**40**	**0**	0	**0**	0	**33**	9.9	**33**	28.8
Netball (19 mins)	**43**	**36**	0.4	**39**	8.8	**38**	4.9	**41**	13.4
Average time spent for 1 min segment		**0.04**		**1.06**		**0.52**		**1.49**	

Table 3. Basketball Video Samples

Group (broadcaster)	Basketball Videos "team1-teams2_period-[duration]"
Athens 2004 Olympics (Seven)	Women: AusBrazil_ 1,2,3-[19:50,19:41,4:20] Women: Russia-USA_3-[19:58] Men: Australia-USA_1,2-[29:51,6:15]
Athens 2004 Olympics (SBS)	Men: USA-Angola_2,3-[22:25,15:01] Women: Australia-USA_1,2-[24:04-11:11]

Table 4. Highlight Classification in Basketball Videos

Ground truth	Highlight classification of 5 basketball videos				
	Goal	Free throw	Foul	Timeout	Truth
Goal	**56**	0	0	2	58
Free throw	4	**14**	0	0	18
Foul	21	2	**28**	3	54
Timeout	0	0	0	**13**	13
Total Detected	81	16	28	18	

4   Conclusion and Future Work

In this paper, we have presented a multi-level semantic analysis framework for sports video by extending semantic layer into: generic-, classified-, further-tactical- and customized-semantic. This paper focuses on the analysis methods and performance results for extraction of generic and classified semantic. In future work, we aim to enhance the performance of the semantic analysis while extending the scope of detectable mid-level features and semantic for various sports genre. We also aim to demonstrate the extraction of further-tactical semantic such as soccer free kick.

References

1. Lu, G.J.: Multimedia database management systems. Artech House. Boston London (1999).
2. Tusch, R., Kosch H., Böszörményi, L.: VIDEX: an integrated generic video indexing approach. Proceedings of the 8th ACM international conference on Multimedia. ACM Press. Marina del Rey California United States (2000) 448-451.
3. Elgmagarmid, A.K. (ed): Video database systems: issues, products, and applications. Kluwer Academic Publishers. Boston London (1997).
4. Djeraba, C.: Content-based multimedia indexing and retrieval. IEEE Multimedia. Vol. 9(2) (2002) 18-22.
5. Duan, L.-Y., Min, X., Chua, T-S., Qi, T., Xu, C-S: A mid-level representation framework for semantic sports video analysis. Proceedings of the 11th ACM international conference on Multimedia. ACM Press. Berkeley USA (2003) 33-44.
6. Ekin, A., Tekalp, M.: Automatic Soccer Video Analysis and Summarization. IEEE Transaction on Image Processing. Vol. 12(7) (2003) 796-807.
7. Han, M., Hua, W., Chen, T., Gong, Y.: Feature design in soccer video indexing. Proceedings of the 4th Conference on Information, Communications and Signal Processing. IEEE. Singapore (2003) 950-954.
8. Tjondronegoro, D., Chen, Y-P.P., Pham, B.: The Power of Play-Break for Automatic Detection and Browsing of Self-consumable Sport Video Highlights. The Proceedings of the 6th International Multimedia Information Retrieval Workshop. ACM Press. New York USA (2004) 267-274.
9. Nepal, S., Srinivasan, U., Reynolds, G.: Automatic detection of goal segments in basketball videos. Proceedings of the 9th ACM International Conference on Multimedia. ACM Press. Ottawa Canada (2001) 261-269.
10. Tjondronegoro, D., Chen, Y-P.P., Pham, B.: Integrating Highlights to Play-break Sequences for More Complete Sport Video Summarization. IEEE Multimedia. Vol.11(4) (2004) 22-37.
11. Tjondronegoro, D., Chen, Y-P.P., Pham, B.: A Statistical-driven Approach for Automatic Classification of Events in AFL Video Highlights. Proceedings of the 28th Australasian Computer Science Conference. ACS. Newcastle Australia (2005) 209-218.

Aerial Photograph Image Retrieval
Using the MPEG-7 Texture Descriptors

Sang Kim[1], Sung Baik[1,*], Yung Jo[1], Seungbin Moon[1], and Daewoong Rhee[2]

[1] College of Electronics and Information Engineering, Sejong University
Seoul 143-747, Korea
{sskim,sbaik,joyungki,sbmoon}@sejong.ac.kr
[2] College of Computer Software & Media Technology, Sangmyung University
Seoul 110-743, Korea
rhee219@smu.ac.kr

Abstract. The MPEG-7 texture description tools are applied to the retrieval of the historical aerial photographs. These aerial photographs had been digitized into gray-scaled and low-resolution images to build a digital library. The aerial image retrieval over the digital library is achieved on the basis of 1) texture feature extraction methods, which are homogeneous texture, texture browsing and edge histogram in the MPEG-7 texture descriptors, and 2) their corresponding similarity measurements. The image retrieval methods are evaluated over a collection of historical aerial photographs. The paper presents the retrieval results of these texture feature extraction methods.

1 Introduction

Content-based image retrieval requires the integration of image processing and information retrieval technologies. This is the retrieval of images on the basis of features automatically derived from the images themselves. Texture, color and shape are used the most widely in most researches to describe features in the image. However, the retrieval of the historical aerial image photographs is based only on texture features because they are gray-scaled and low-resolution images. Therefore, more robust feature extraction methods are required to allow effective retrieval results.

In low resolution aerial images, we cannot apprehend the appearance of certain objects in detail. Therefore, we need to use regions with a collection of tiny and complicated structures–such as man-made features including buildings, roads, parking lots, airports and bridges–in order to deal with them as texture motifs for image classification/segmentation. We can also regard the shapeless regions of natural resources such as forests, rivers and oceans as texture motifs [1,2].

This work presents a texture-based geographical image retrieval system using the MPEG-7 texture descriptors, which consist of the homogeneous texture descriptor (HTD), the texture browsing descriptor (TBD) and the edge histogram descriptor (EHD). These descriptors support effective nearest-neighbor and range similarity searches.

---

* Author for correspondence : +82-2-3408-3797.

The MPEG-7 texture description tools [3,4] have widely been applied to a variety of data sets for image retrieval. As an example, the HTD has already been used to annotate a large collection of aerial images of the Santa Barbara, California in USA [5], and its demonstration is available on line [6]. Also, there are many efforts to improve the performances of texture descriptors such as HTD [7,8], TBD [9] and EHD [10].

2 MPEG-7 Texture Descriptors

2.1 Homogeneous Texture Descriptor

HTD uses a mean value, a standard deviation and Gabor spectral filtering results of each gray scaled image, in order to represent geographical features for aerial image retrieval. Gabor spectral filtering [11] has been widely used for various classification and image segmentation tasks.

Gabor filters are useful for dealing with texture characterized by local frequency and orientation information. Gabor filters are obtained through a systematic mathematical approach. A Gabor function consists of a sinusoidal plane of particular frequency and orientation modulated by a two-dimensional Gaussian envelope. A two-dimensional Gabor filter is given by:

$$G(x, y) = \exp[\frac{1}{2}(\frac{x}{\sigma_x^2} + \frac{y}{\sigma_y^2})]\cos(\frac{2\pi x}{n_0} + \alpha) \tag{1}$$

By orienting the sinusoid at an angle α and changing the frequency n_0, many Gabor filtering sets can be obtained. An example (Fig. 1) of a set of eight Gabor filters is decided with different parameter values (n_0 = 2.82 and 5.66 pixels/cycle and orientations α = 0°, 45°, 90°, and 135°).

| n_0=2.82, α=0° | (2.82, 45°) | (2.82, 90°) | (2.82, 135°) |
| n_0=5.66, α=0° | (5.66, 45°) | (5.66, 90°) | (5.66, 135°) |
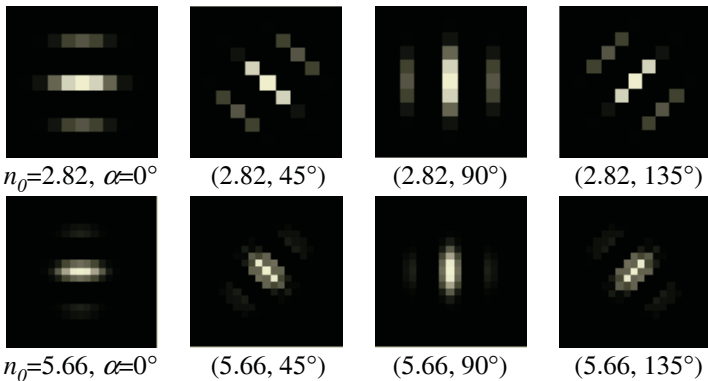
Fig. 1. A set of Gabor filters

2.2 Texture Browsing Descriptor

Directionality, coarseness, and regularity of patterns appearing on an aerial image are represented by texture. TBD is a useful descriptor to specify the perceptual charac-

terization of a texture since it considers a human characterization, in terms of directionality, coarseness, and regularity.

2.3 Edge Histogram Descriptor

The spatial distribution is represented by five types of edges in the EHD. These are four directional edges (one horizontal, one vertical, and two diagonal edges) and one non-directional edge for 16 sub-images into which the image space is divided without overlapping. Five types of edges are defined for each sub-image and we have a local edge histogram with total 16×5=80 histogram bins. Fig. 2 shows an example of 16 non-overlapping sub-images.
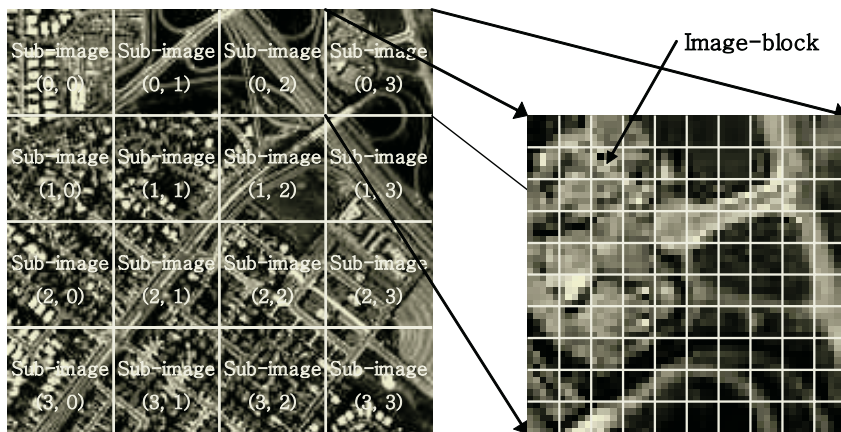
Fig. 2. Example of 16 non-overlapping sub-images

3 Experiments

Experimental data are provided by the UC Berkeley Library Web [12]. They are 184 aerial photographs of the San Francisco Bay area, California, taken in April, 1968 by the U.S. Geological Survey. The scale of the originals is 1:30,000. Each photograph image has the size of approximately 1300×1500 pixels with 256 grey-level (the size and resolution of each image are little different from each other). It is cut into about $100 (10 \times 10)$ non-overlapping sub-images (texture tiles) of size 120×120. A test bed for image retrieval has about $18,400 (184 \times 100)$ texture tiles. Fig. 3 presents an aerial photograph image and 100 non-overlapping sub-images.

For evaluation purposes, five categories are decided by human observers; 1) Sea/River, 2) Vegetation, 3) Downtown1, 4) Downtown2, and 5) Downtown3. Three of the most popular patterns in downtown regions are selected and correspond to downtown1, downtown2, and downtown3, respectively. Their examples are shown in Fig. 3. Three hundreds texture tiles are also selected for each category according to human visual experiences and indexed for retrieval performance evaluation.
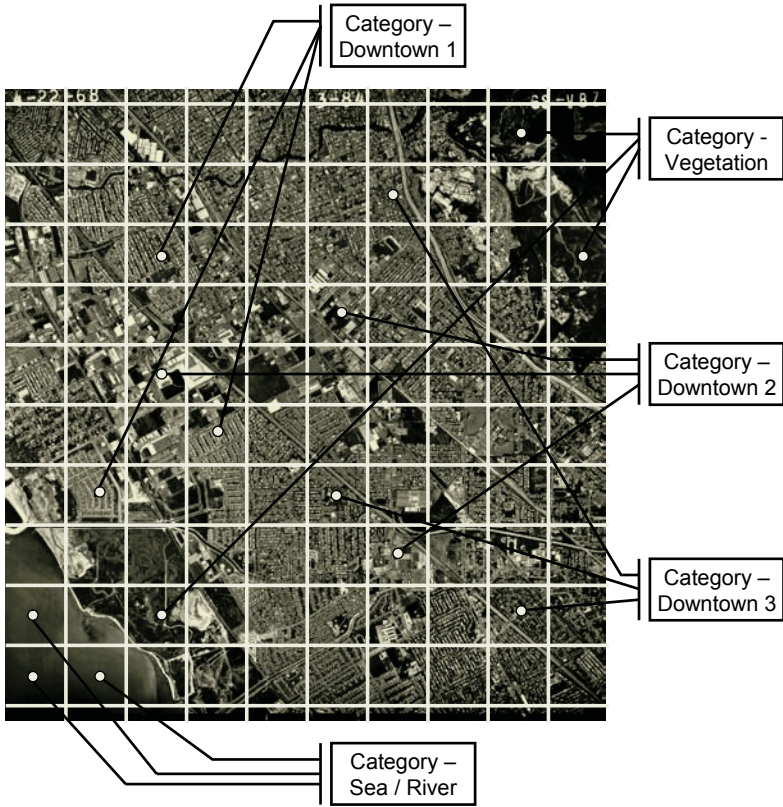
Fig. 3. An aerial photograph image and 100 non-overlapping sub-images
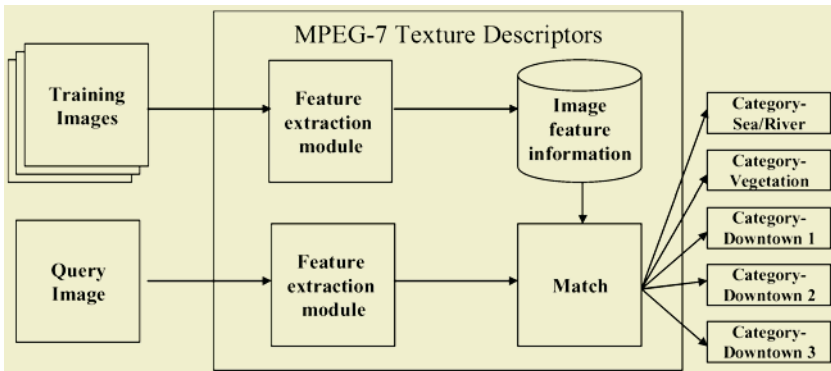
Fig. 4. Image retrieval process in a training phase and a matching phase

We evaluate the performance of the image retrieval method with different texture descriptors: 1) Homogeneous Texture, 2) Texture Browsing, and 3) Edge Histogram. Two hundreds texture tiles have been used to train the texture descriptors, and the rest tiles have been used to match with the training descriptors. Fig. 4 shows the image
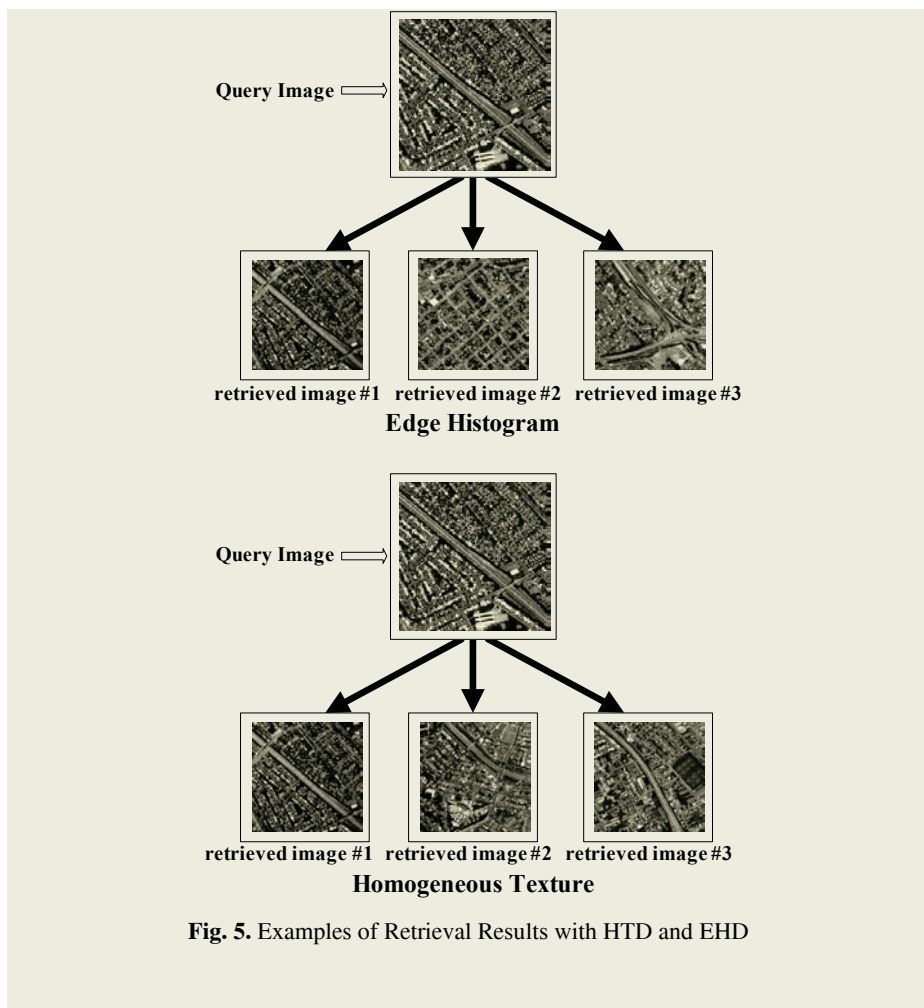
Fig. 5. Examples of Retrieval Results with HTD and EHD

retrieval process through two paths such as a training phase and a matching phase. Fig. 5 shows examples of Retrieval Results with HTD and EHD. Finally, Table 1 includes the comprehensive results for the performance of the aerial image retrieval methods.

4   Conclusion

This paper presented a texture-based geographical image retrieval system with MPEG-7 texture description tools. We found that HTD outperformed the other one when the description tools were applied to the historical aerial photograph image data.

Table 1. The performance (classification rates (%)) of the aerial image retrieval methods (A: Sea/River, B: Vegetation, C: Downtown1, D: Downtown2, E: Downtown3)

	Homogeneous Texture					Texture Browsing					Edge Histogram				
	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
A	95	5	0	0	0	94	6	0	0	0	94	6	0	0	0
B	6	94	0	0	0	6	94	0	0	0	7	93	0	0	0
C	0	0	71	18	11	0	0	69	19	12	0	0	66	21	13
D	0	0	19	69	12	0	0	20	66	14	0	0	22	60	18
E	0	0	13	12	75	0	0	14	14	72	0	0	18	17	65

References

1. A. Gasteratos, P. Zafeiridis and I. Andreadis, An Intelligent System for Aerial Image Retrieval and Classification, LNCS, Vol. 3025, pp. 63-71, 2004.
2. B. Zhu, M. Ransey and H. Chen, Creating a Large-Scale Content-Based Airphoto Image Digital Library, IEEE Transactions on Image Processing, Vol. 9, Issue. 1, pp. 163-167, 2000.
3. W. Peng, Y. M. Ro, S. W. Chee, Y. Choi, Texture Descriptors in MPEG-7, LNCS, Vol. 2124, pp. 21, 2001
4. Timo Ojala, Topi Mäenpää, Jaakko Viertola, Juha Kyllönen and Matti Pietikäinen, Empirical Evaluation of MPEG-7 Texture Descriptors with A Large-Scale Experiment, Proceedings 2nd International Workshop on Texture Analysis and Synthesis, Copenhagen, Denmark, pp. 99-102, 2002.
5. W. Y. Ma and B. S. Manjunath, A texture thesaurus for browsing large aerial photographs, Journal of American Society for Information Science, 49(7), pp. 633-648, 1998.
6. See http://vision.ece.ucsb.edu
7. Y. M. Ro and H. K. Kang, Hierarchical rotational invariant similarity measurement for MPEG-7 homogeneous texture descriptor, Electronic Letters, Vol. 36, No. 15, 2000.
8. B.-M. Straub, Investigation of the MPEG-7 Homogeneuos Texture Descriptor for the Automatic Extraction of Trees, Proceedings of the ISPRS Technical Commission III Symposium, pp. A-351, 2002.
9. L. L. Lee and L. H. Chen, An efficient computation method for the texture browsing descriptor of MPEG-7, Image and Vision Computing, Vol. 23, pp. 479-189, 2005.
10. S. W. Chee, D. K. Park and S. J. Park, Efficient Use of MPEG-7 Edge Histogram Descriptor, ETRI Journal, Vol. 24, No. 1, 2002.
11. L. Chen, G. Lu and D. Zhang, Effects of Different Gabor Filter Parameters on Image Retrieval by Texture, Proceedings of the 10th International Multimedia Modeling Conference, pp. 273-278, 2004.
12. http://sunsite.berkeley.edu/AerialPhotos/vbzj.html#index

Yet Another Induction Algorithm

Jiyuan An[1] and Yi-Ping Phoebe Chen[1,2]

[1] School of Information Technology, Faculty of Science and Technology
Deakin University, Melbourne, VIC 3125 Australia
[2]Australian Research Council Centre in Bioinformatics
{jiyuan,Phoebe}@deakin.edu.au

Abstract. Inducing general functions from specific training examples is a central problem in the machine learning. Using sets of If-then rules is the most expressive and readable manner. To find If-then rules, many induction algorithms such as ID3, AQ, CN2 and their variants, were proposed. Sequential covering is the kernel technique of them. To avoid testing all possible selectors, Entropy gain is used to select the best attribute in ID3. Constraint of the size of star was introduced in AQ and beam search was adopted in CN2. These methods speed up their induction algorithms but many good selectors are filtered out. In this work, we introduce a new induction algorithm that is based on enumeration of all possible selectors. Contrary to the previous works, we use pruning power to reduce irrelative selectors. But we can guarantee that no good selectors are filtered out. Comparing with other techniques, the experiment results demonstrate that the rules produced by our induction algorithm have high consistency and simplicity.

1 Introduction

Finding out a definition of a general category from the positive and negative training examples, is a classical topic [3][6][7][8][9] in the field of machine learning. Induction algorithm finds the subspaces in the feature space. These subspaces cover all positive training examples but do not include any negative examples are included. It is a NP-hard problem to find out the optimal subspaces. In order to make the search tractable, a common technique is to use heuristic search methods to find the "better" not "best" subspaces such as decision tree [11], GA [8], AQ [6] and CN2[2].

Training examples can be viewed as points in a high dimensional space whose axes are constructed by the features. Induction algorithm finds out border lines, planes, and hyperplanes to encompass positive training examples. But they do not cover negative examples. We can also use more than one cover to encompass all positive examples. This kind of covers is called positive covers. A positive cover is expressed by a conjunction of *selectors* such as $(X \geq 3 \cap y \geq 6)$. A selector in the rules can be a line, plane and hyperplane in feature space.

According to Occam's razor theorem, the simplest concept that is consistent with all training examples is the best. It is desirable that we can find fewer positive covers to cover more positive examples. At the same time, we should use fewer selectors to describe these positive covers. A complex of fewer selectors usually covers bigger

feature space. For example, the cover of (X ≥ 2) is bigger than that of (X ≥ 2 ∩ Y ≥ 2). Our algorithm starts from the simplest complexes, which consist of one selector (1-feature rules). Then we extend the complexes by adding another selector; the complexes become 2-feature rules. This process continues until no better positive cover exists.

Contrary to other induction algorithms, this algorithm does not make assumption for any hypotheses. As a result, it avoids the defect of other techniques as a good result depends on selection of right parameters such as (penalty factor, and etc.). Our algorithm is more robust than any of GA-base, AQ and CN2 algorithms.

Our algorithm enumerates all positive combinations of features. As the number of arbitrary feature combinations is very large, it is CPU costly. In this paper, we introduce pruning power to solve the problem.

We summarize contributions of this paper:

- A new induction algorithm is proposed. Comparing with AQ15 or GA-based algorithms, our technique is speedy, stable and accurate.
- In the process of extending k-feature rules to k+1-feature rules, we propose a feature pruning technique to reduce candidates. This technique makes our algorithm effectively.

2   Our Approach

Induction algorithm finds a set of positive covers which cover all positive examples but no negative examples. The answer is not unique; we can usually get many answers. It is because that we can use different set of covers to cover all positive examples. According to Occam's razor theorem, in order to find good rules to describe concepts, we have to minimize the number of positive covers. For example, if we can use one cover to cover all positive examples, we will not use two covers. However, finding this kind of positive covers is an NP-hard problem [4].
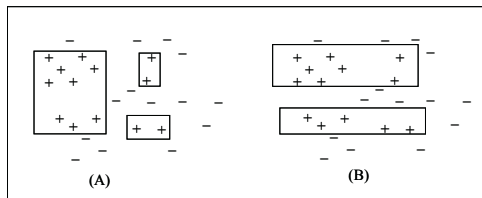
Fig. 1. Finding biggest positive cover.

Our algorithm finds positive covers one by one. First we find the biggest positive cover on original feature space. At that time, some positive examples are not covered by the positive cover. To find the next positive cover, the positive examples covered by the biggest positive cover are removed. In the remaining positive and negative examples, we find the biggest positive cover again. The process is repeated until all positive examples are covered by at least one positive cover. The number of positive covers should be small, although the result may be not the best answer as shown in Figure 1. Subfigure (A) uses three positive covers to cover all positive, because the

biggest cover is selected on the original domain. If we select the second biggest positive cover, only two positive covers are needed to cover all positive examples as shown in Figure 1 (B).

In this paper, we assume all features have discrete values (There are many effective discretization methods [4][8]). If we have n features, every feature has N discrete values. Then there are $C_n^1 \times C_N^1$ 1-feature combinations, and $C_n^i \times (C_N^1)^i$ i-feature combinations. For example, given 2 feature space, the discrete values N = 2 ("Yes" and "No"). Then, 1-feature combinations are f_0="Yes", f_0="No" f_1="Yes", f_1 = "No". 2-feature combinations are: (f_0, f_1) = {("Yes", "Yes"), ("Yes", "No"), ("No", "Yes"), ("No", "No")}. Our algorithm chooses the biggest positive cover from 1-feature, 2-feature, …, n-feature combinations. The number of all possible combinations is $\sum_{i=1}^{n} C_n^i \times (C_2^1)^i$.

```
Algorithm find rules
1. While positive data set ≠ Ø
2.    maxCover = 0
3.    for nc = 1 : n   {the number of combination changed from 1 to n}
4.       for each combination c₁,…,c_nc {check all possible combinations}
5.          if (maxCover < pureCover(c₁,…,c_nc ))
6.             maxcover = pureCover(c₁,…,c_nc )
7.             bestCover = c₁,…,c_nc
8.          endif
9.       endfor
10.   endfor
11.   delete all positive examples covered by bestcover
12.   if no positive cover is found
13.      return all bestCover
14.   endif
15. endwhile
```

Fig. 2. Algorithm for finding rules

Figure 2 shows the algorithm to find the set of biggest positive covers. The line 3-10 gives all possible combinations of features. Each combination is checked against every positive and negative example whether it is covered by the combination or not. In line 5, the number of positive and negative examples is calculated in terms of every combination. The function pureCover(c_1,…,c_{nc}) in line 5 is to calculate the number of positive examples in a positive cover. In line 11, the positive examples covered by selected combination are removed. A new biggest positive cover is found from the remaining positive and negative examples. The program is repeated until all positive examples are covered by at least one positive cover.

Figure 3 (A) gives 4 training examples consisting of 3 features (f_1, f_2, f_3). Each feature has two values 'Y' (yes) and 'N' (no). There are two positive examples and three negative examples. The combination of features has only 3, i.e. 1-feature, 2-features and 3-features. 1-feature rule likes f_1='Y' or f_2='N' shown Figure 3 (B).

In Figure 3 (B), the rule f_1='Y' covers 1 positive example and 2 negative example. The rule (f_1 = 'N') covers 1 positive examples but no negative examples are concluded, that is, pureCover(f_1='N')=1. So the rule is a positive cover in 1-feature rules. All 1-feature rules are extended to 2-feature rules. From Figure 3 (C), we find that

there are three rules which cover only positive examples and no negative examples are covered, i.e. pureCover(f_1='N' $\cap$ f_2='Y') = 1, pureCover(f_1='N' $\cap$ f_3='N') = 1, pureCover(f_2='Y' $\cap$ f_3='N') = 2. The biggest positive cover (f_2='Y' $\cap$ f_3='N') describes the concept of the 5 positive and negative examples.
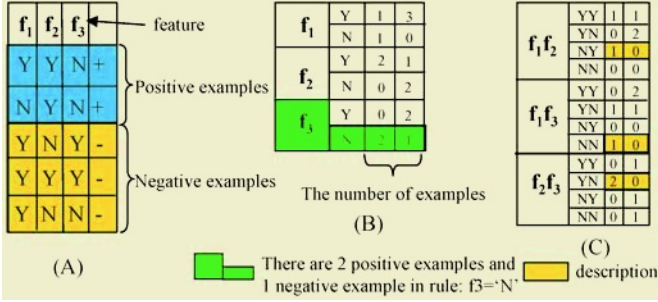
Fig. 3. An example of algorithm 1

However, this algorithm is only applicable in small feature spaces. Most practical learning tasks involve much larger feature spaces. The all possible rule set becomes very big when the number of feature increases. Researchers in spatial database call this "curse of dimensionality" [1]. In high feature dimensional space, it is very difficult to find a positive cover. In next section, we explain how to prune most irrelative combinations and get small set of candidates.

2.1 Pruning Power

A positive cover is represented with conjunction of selectors. For example, (f_1='Y' $\cap$ f_2 = 'N'). The description of concept can be viewed as a disjunction of some positive covers or rules. To make the algorithm tractable, we have to reduce the number of possible combinations. Fortunately, we find that most of combinations can be pruned without affecting the process of finding biggest position cover. The rule f_2='N' in Figure 3(B) does not cover any position examples. So it can be pruned. If a rule is denoted as R, cover(R) is a function which calculates the number of positive examples covered by rule R. Unlike the function pureCover(R), cover(R) does not consider weather R covers negative examples or not. For example, in Figure 3, pureCover(f_1='Y')=0, cover(f_1='Y') = 1.

LEMMA 1. If the number of positive examples covered by rule R is represented as cover(R), then we have cover(R) $\geq$ cover(r$\cap$R), where r denotes an arbitrary rule.

Proof: $\exists e \in r \cap R$, we have $e \in R$. So cover(R) $\geq$ cover(r$\cap$R). $\square$

In our approach, first we calculate all positive covers of 1-feature rules as illustrated in Figure 3(B). In all 1-feature rules, we can get a cover which covers the most number of positive examples but no negative examples. A rule is called best rule if it covers more number of positive examples but no negative examples than others. In our example, (f_1='N') is the best rule in 1-feature rules. pureCover(f_1='N') = 1. Ac-

cording to LEMMA 1, we can omit all combinations that cover less than one positive example. In the example, $\{f_1='Y', f_1='N', f_2='N', f_3='Y'\}$ and their combinations can be pruned. Only two 1-feature rules ($f_2='Y'$, $f_3='N'$) and their combinations become the candidates of concept description. We easily get pureCover($f_2='Y' \cap f_3='N'$)=2. It is the final answer. 11/12 combinations are pruned in the 2-feature rules. We will explain the pruning power in our algorithm by using Figure 4.

First, we calculate the covers of n 1-feature rules: cover(f_1), cover(f_2), ..., cover(f_n). Then we sort them into a queue in descending order. cover(r_{11}) $\geq$ cover(r_{12}) $\geq$ ... $\geq$ cover(r_{1n}) as shown in Figure 4(A). From the queue, we find out a rule that has the biggest positive cover. It is denoted as bestCandidate(1). In Figure 3(B), bestCandidate is rule $f_1='N'$. In Figure 4(A), rule r_{11} usually covers more number of positive examples than bestcandidate, i.e. cover(r11) > pureCover(bestcandidate). It is because that r_{11} usually covers negative examples. Second, we constitute 2-feature rules. Because the rule r_{11} covers the most number of positive examples in 1-feature rules, it is selected firstly to expand to 2-feature rules. By combining with each feature except itself, 2-feature rules are constituted. The bestcandidate is updated when a bigger pureCover(R) appears. Following r_{11}, the second candidate r_{12} in 1-feature rules is expanded to 2-feature rules. However, we must confirm weather r_{12} can be pruned or not. If r_{12} does not satisfy cover(r_{12}) > pureCover(bestCandidate), r_{12} is pruned. The process continues until cover(r_{1i}) $\leq$ pureCover(bestCandidate). If we can not find a rule that covers more positive examples than the rule bestcandidate, the program stops. The program usually stop before the rules are expanded into n-feature rules.
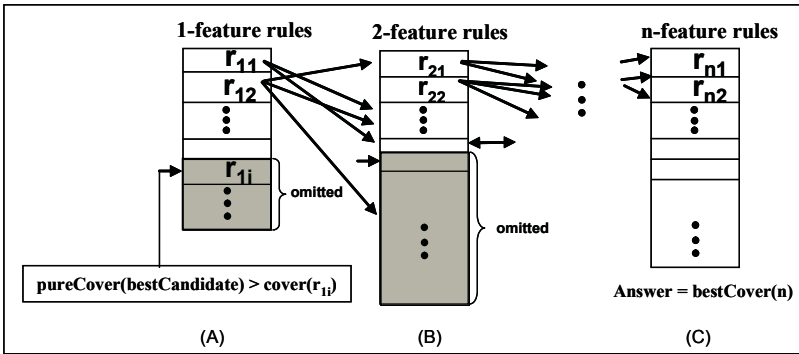
Fig. 4. Pruning power for finding best candidate

3 Experiment

We select voting dataset to explain our algorithm's effectiveness. The voting dataset has been widely used to evaluate concept learning algorithm. It recorded the votes of the U.S. House of Representatives Congressmen on 16 issues. The dataset has 435 training examples, including 267 Democrat and 168 Republican voting records. Each

example gives the vote cast by an individual on 16 difference issues. Each vote has 3 values ("Yes", "No" and "Abstain", simplified 'y', 'n' and '?').

3.1 Comparison to AQ15 and C4.5

The source code of decision tree (C4.5) was downloaded from [12]. AQ15 is a sophisticated symbolic inductive learning system. It can be run with its default settings [9]. Figure 5 shows the result of C4.5 decision tree. On the right side shows, the number of democratic examples is shown and their rules are shown on the left hand of Figure 5. Most rules consist of more than 4 features. The average length of rules is 4.47 features. In the 17 rules, we can also find that 9 rules cover only 1 or 2 positive examples. This means most rules do not reflect common character of democrats.

1.	$f3 = n$ & $f2 = y$	(219)
2.	$f3 = n$ & $f2 = ?$	(3)
3.	$f3 = n$ & $f2 = n$ & $f11 = y$	(10)
4.	$f3 = n$ & $f2 = n$ & $f11 = n$ & $f10 = y$	(9)
5.	$f3 = n$ & $f2 = n$ & $f11 = n$ & $f10 = n$ & $f5 = y$	(3)
6.	$f3 = n$ & $f2 = n$ & $f11 = n$ & $f10 = n$ & $f5 = n$ & $f13 = n$	(1)
7.	$f3 = y$ & $f10 = n$ & $f14 = y$ & $f2 = y$ & $f5 = u$ & $f0 = n$	(1)
8.	$f3 = y$ & $f10 = n$ & $f14 = y$ & $f9 = n$ & $f6 = n$ & $f11 = n$	(2)
9.	$f3 = y$ & $f10 = y$ & $f8 = n$ & $f2 = n$ & $f12 = n$ & $f1 = y$	(1)
10.	$f3 = y$ & $f10 = y$ & $f8 = n$ & $f2 = n$ & $f12 = y$ & $f15 = n$ & $f0 = n$	(1)
11.	$f3 = y$ & $f10 = y$ & $f8 = n$ & $f2 = n$ & $f12 = y$ & $f15 = ?$ & $f0 = y$	(1)
12.	$f3 = y$ & $f10 = y$ & $f8 = n$ & $f2 = y$ & $f1 = y$	(4)
13.	$f3 = y$ & $f10 = y$ & $f8 = y$ & $f12 = y$	(3)
14.	$f3 = y$ & $f10 = y$ & $f8 = y$ & $f12 = n$ & $f2 = n$	(1)
15.	$f3 = ?$ & $f11 = n$	(6)
16.	$f3 = ?$ & $f11 = ?$ & $f8 = n$	(1)
17.	$f3 = ?$ & $f11 = ?$ & $f8 = y$	(1)

Fig. 5. The conjunctive descriptions by Democratic Decision Tree

By using AQ15 method, 22 rules are produced to describe the democrats as shown in Figure 6. The length of rules are shorter than the rules produced by C4.5, and many rules cover more than 1 or 2 examples. The average length of rules is 3.4 features. However, the number of rules is up to 22, this means AQ15 fail to find good feature combinations to describe democrats. It has bad simplicity.

1.	$f3 = n$ & $f2 = y$	(219)	12.	$f2 = y$ & $f10 = ?$ & $f8 = n$ & $f6 = n$	(2)
2.	$f3 = ?$ & $f2 = y$	(5)	13.	$f2 = y$ & $f10 = y$ & $f8 = y$ & $f6 = ?$	(1)
3.	$f3 = n$ & $f11 = y$ & $f5 = y$	(25)	14.	$f2 = y$ & $f10 = y$ & $f8 = y$ & $f6 = n$	(9)
4.	$f3 = n$ & $f11 = n$ & $f5 = y$	(73)	15.	$f2 = y$ & $f10 = y$ & $f8 = n$ & $f6 = ?$	(1)
5.	$f14 = y$ & $f13 = y$ & $f1 = ?$	(2)	16.	$f2 = y$ & $f10 = y$ & $f8 = n$ & $f6 = n$	(18)
6.	$f14 = y$ & $f13 = y$ & $f1 = y$	(15)	17.	$f2 = y$ & $f15 = ?$ & $f12 = y$	(15)
7.	$f14 = y$ & $f13 = n$ & $f1 = ?$	(18)	18.	$f4 = n$ & $f14 = y$ & $f2 = n$	(11)
8.	$f14 = y$ & $f13 = n$ & $f1 = y$	(48)	19.	$f12 = n$ & $f1 = y$ & $f2 = n$	(7)
9.	$f2 = ?$ & $f10 = ?$ & $f8 = y$ & $f6 = ?$	(1)	20.	$f11 = n$ & $f10 = ?$ & $f15 = ?$ & $f2 = ?$	(1)
10.	$f2 = ?$ & $f10 = y$ & $f8 = y$ & $f6 = n$	(1)	21.	$f11 = n$ & $f10 = y$ & $f15 = ?$ & $f2 = n$	(3)
11.	$f2 = ?$ & $f10 = y$ & $f8 = n$ & $f6 = n$	(2)	22.	$f10 = y$ & $f1 = n$ & $f0 = n$ & $f15 = n$	(2)

Fig. 6. The conjunctive descriptions of democratic by AQ15

Figure 7 shows the democrat description rules produced by proposed algorithm. We use only 15 rules to describe democrat. The average length of rules is 2.53 features. Furthermore, many rules cover more than 20 positive examples. This means that these rules represent common characters of positive examples. 15 rules cover, on an average 44.8 positive examples. We find very good feature combinations to describe the characters of democrats. The rules produced by proposed algorithm satisfy completeness, consistency and simplicity. Our approach has the least and shortest rules of three methods, but has largest coverage for positive examples.

1.	f2 = y & f3 = n	(219)	9.	f1 = y & f3 = ?	(5)
2.	f3 = n & f10 = y	(113)	10.	f1 = n & f13 = n	(74)
3.	f5 = y & f8 = y & f11 = n	(50)	11.	f4 = ? & f8 = y	(7)
4.	f2 = y & f6 = n & f10 = y	(29)	12.	f2 = y & f6 = n & f9 = n	(22)
5.	f3 = n & f12 = y	(58)	13.	f1 = y & f6 = n & f14 = y	(17)
6.	f0 = y & f2 = ?	(2)	14.	f0 = y & f1 = y & f15 = ?	(26)
7.	f1 = y & f9 = n & f12 = n	(43)	15.	f0 = n & f1 = n & f10 = y & f15 = n	(2)
8.	f0 = y & f6 = ?	(5)			

Fig. 7. The rules to descript democratic by proposed algorithm

4   Conclusion

Searching a complex from complete feature space has so far been considered to be impractical. A new induction algorithm that incorporates novel pruning method is introduced to solve this problem. This new induction algorithm enumerates all possible complexes and is demonstrated with testing examples. Our algorithm can be considered to be the first one in the literature, which finds out complexes from all possible selectors for searching a positive cover. It guarantees that each rule is the best one in the training examples. We plan to analyse the pruning power in extremely high dimensional data such as gene microarray.

Acknowledgments

References

1. Beyer, K. S. Goldstein J. Ramakrishnan R. and Shaft U.: "When Is 'Nearest Neighbor' Meaningful". In proceeding s of ICDT'99 (1999) 217-235.
2. Clark, P. and Boswell, R. Rule induction with CN2: some recent improvements. In Machine Learning - EWSL-91. (1991) 151-163.
3. Divina, F. and Marc, E.: "Evolutionary Concept Learning", GECCO (2002): 343-350.
4. Fayyad, U. M. and Irani, K. B.. Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning. IJCAI93, pp 1022-1027, 1993.

5. Jong, K. and Spears, W.: "Using Genetic Algorithms for Concept Learning," Machine Learning, vol. 13, pp. 161-188, 1993.
6. Michalski, R. S.: "On the quasi-minimal solution of the general covering problem," in Proceedings of ISIP vol. A3, 1969, pp. 125–128.
7. Michalski. R. S. Carbonell, J. G. and Mitchell, T. M.: "Machine learning an artificial intelligence approach", Morgan Kaufmann Publishers, INC., 1983.
8. Mitchell, T. "Machine learning", McGraw Hill, 1997.
9. Michalski, R. S., Mozetic, I., Hong, J, Lavrac, N., "The AQ15 inductive learning system: An overview and experiments," Technical Report UIUCDCS-R-86-1260, University of Illinois, Urbana-Champaign. IL, 1986.
10. Potter, M. A. "The Coevolution of Antibodies for Concept Learning," presented at 5th International Conference of Parallel Problem Solving from Nature, pp. 530-539, 1998.
11. Quinlan J. R. Induction of Decision Trees. Machine Learning 1(1): pp. 81-106 1986
12. Winston, P. "C4.5 Decision Tree", http://www2.cs.uregina.ca/~hamilton/   2002

An Implementation of Learning Classifier Systems for Rule-Based Machine Learning

An-Pin Chen and Mu-Yen Chen

Institute of Information Management, National Chiao Tung University
Hsinchu, 300, Taiwan, ROC
{apc,mychen}@iim.nctu.edu.tw

Abstract. Machine learning methods such as fuzzy logic, neural networks and decision tree induction have been applied to learn rules, however they can get trapped into a local optimal. Based on the principle of natural evolution and global searching, a genetic algorithm is promising for obtaining better results. This article adopts the learning classifier systems (LCS) technique to provide a hybrid knowledge integration strategy, which makes for continuous and instant learning while integrating multiple rule sets into a centralized knowledge base. This paper makes three important contributions: (1) it provides a knowledge encoding methodology to represent various rule sets that are derived from different sources, and that are encoded as a fixed-length bit string; (2) it proposes a knowledge integration methodology to apply genetic operations and credit assignment to generate optimal rule sets; (3) it uses three criteria (accuracy, coverage, and fitness) to apply the knowledge extraction process, which is very effective in selecting an optimal set of rules from a large population. The experiments prove that the rule sets derived by the proposed approach is more accurate than the Fuzzy ID3 algorithm.

1 Introduction

Developing an expert system requires construction of a complete, correct, consistent, and concise knowledge base. The knowledge base construction always involves interaction and dialogue between domain experts and knowledge engineers. Therefore, acquiring and integrating multiple knowledge inputs from many experts or by various knowledge-acquisition techniques plays an important role in building an effective knowledge-based systems [1][2].

Generally speaking, knowledge integration can be thought of as a multi-objective optimization problem [10]. Due to the huge searching space, the optimization problem is often very difficult to solve. A genetic algorithm (GA) is usually used to discover a desirable but not necessarily an optimal set of rules. The application of a GA in search of an optimal rule set for machine learning is known as Genetic Based Machine Learning (GBML). A well-known GBML architecture is the so-called Learning Classifier Systems (LCS) developed by Holland [3][4]. More recent GBML architectures are the Extended Classifier System (XCS) developed by Wilson [8], the Anticipatory Classifier System by Stolzmann [7], and EpiCS by Holme [5].

Our research objective therefore was to employ the XCS technique, which can effectively integrate multiple rule sets into one centralized knowledge base. The rest of

this paper is organized as follows. Section 2 describes the system architecture of the XCS-based model. We then present the knowledge encoding methodology in Section 3. Section 4 explains the knowledge integration methodology for reinforcement. Experimental results are reported in Section 5. Finally, we make our conclusion and discuss future work in Section 6.

2   System Architecture

The system architecture is as shown in Figure 1. It is an implementation version of the general framework of the Wilson's XCS classifier system. The algorithm of the XCS-based knowledge integration (XCS-KI) is shown in Figure 2. At first, the system initializes the classifier set which is originally empty and will be covered by new classifiers automatically. Here we assume that all knowledge sources are represented by rules, since almost all knowledge derived by knowledge acquisition (KA) tools, or induced by machine learning (ML) methods may easily be translated into or represented by rules. After the next new run of the system, the initialization stage will load the learned classifier rule sets that are ready to be run.

In the knowledge encoding phase, the system detects the state of the environment, and encodes each rule among a rule set into the condition message of a bit-string structure. Then, the system generates the match set from the classifier set which contains the classifiers matched to the condition message. Next, the system produces a prediction array according to the prediction accuracy of matched classifiers, and generates an action set. After that, the system determines the winner action with the highest accuracy, and then executes this winner action within that environment.

In the knowledge integration phase, the system will obtain the rewards from the environment after executing an action, and then goes through the process of credit/rewards allocation to the classifiers in the action set of the previous step. In the meantime, the system instantly triggers the GA to implement the evolutionary module. Finally, it will report on the accomplishment, and stores the learning classifier set for reapplication in the next activity requirement.
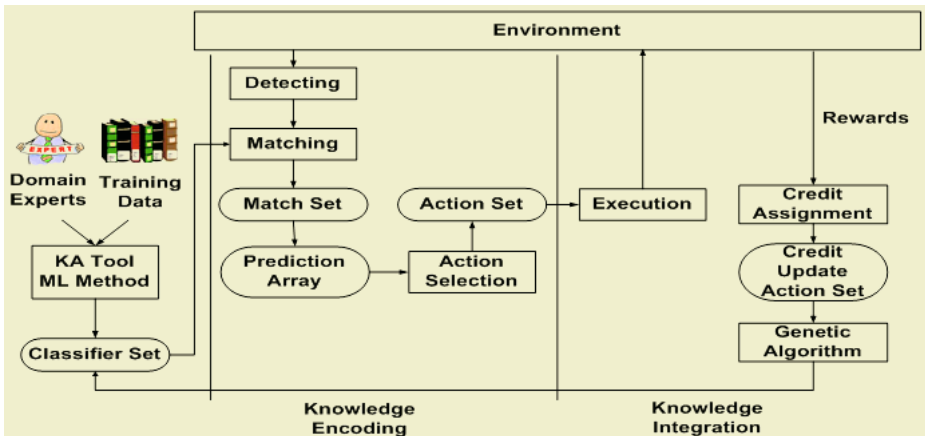
Fig. 1. System Architecture of XCS Classifier System

```
1:   procedure XCS-KI
2:      Initialize classifier set
3:      while (termination condition of XCS is false)
4:         Get environment state
5:         Generate match set
6:         Generate prediction array
7:         Select action
8:         Generate action set
9:         do winner action
10:           Get rewards
11:           Update attribute-values of relevant classifiers
12:           trigger Genetic Algorithm
13:              Selection
14:              Crossover
15:              Mutation
16:           end trigger
17:        end do
18:     end while
19:     Report the execution and learning performance
20:     Store the learned classifier set
21:  end procedure
```

Fig. 2. Algorithm of XCS-based Knowledge Integration (XCS-KI)

3   Knowledge Encoding Phase

We use a pure binary string to do the genetic coding. A classification rule can be coded since one chromosome consists of several segments. Each segment corresponds to either an attribute in the condition part (the IF part) of the rule or to a class in the conclusion part (the THEN part) of the rule. Each segment consists of a string of genes that takes a binary value of 0 or 1. Each gene corresponds to one linguistic term of the attribute or class. To improve the clarity of the coding, we use a semicolon to separate segments and a colon to separate the IF part and the THEN part.

In this paper we use a well known example for deciding what sport to play according to "Saturday Morning Problem" to demonstrate the process of encoding representation [6][9]. Three sports {*Swimming, Volleyball, Weight_lifting*} are to be deciding of four attributes {*Outlook, Temperature, Humidity, Wind*}. The attribute *Outlook* has three possible values {*Sunny, Cloudy, Rain*}, attribute *Temperature* has three possible values {*Hot, Mild, Cool*}, attribute *Humidity* has two possible values {*Humid, Normal*}, and attribute *Wind* has two possible values {*Windy, Not_windy*}. Also, assume that a rule set *RSq* from a knowledge source has the following three rules:

r_{q1} : IF (*Outlook* is *Sunny*) and (*Temperature* is *Hot*) THEN *Swimming*;

r_{q2} : IF (*Outlook* is *Cloudy*) and (*Wind* is *Not_windy*) THEN *Volleyball*;

r_{q3} : IF (*Outlook* is *Rain*) and (*Temperature* is *Cool*) THEN *Weight_lifting*;

The intermediate representation of these rules would then be:

r'_{q1} : IF (*Outlook* is *Sunny*) and (*Temperature* is *Hot*) and (<u>*Humidity* is *Humid* or</u>
<u>*Normal*</u>) and (<u>*Wind* is *Windy* or *Not_windy*</u>) THEN *Swimming*;

r'_{q2} : IF (*Outlook* is *Cloudy*) and (<u>*Temperature* is *Hot* or *Mild* or *Cool*</u>) and (<u>*Humid-*</u>
<u>*ity* is *Humid* or *Normal*</u>) and (*Wind* is *Not_windy*) THEN *Volleyball*;

r'_{q3} : IF (*Outlook* is *Rain*) and (*Temperature* is *Cool*) and (<u>*Humidity* is *Humid* or</u>
<u>*Normal*</u>) and (<u>*Wind* is *Windy* or *Not_windy*</u>) THEN *Weight_lifting*.

The tests with underlines are dummy tests. Also, r'_{qi} is logically equivalent to r_{qi},
for i = 1, 2, 3. After translation, the intermediate representation of each rule is composed of four attribute tests and one class pattern.

After translation, each intermediate rule in a rule set is ready for being encoded as
a bit string. Each attribute test is then encoded into a fixed-length binary string whose
length is equal to the number of possible test values. Each bit thus represents a possible value. For example, the set of legal values for attribute *Outlook* is {*Sunny*,
Cloudy, *Rain*}, and three bits are used to represent this attribute. Thus, the bit string
110 would represent the test for *Outlook* being "*Sunny*" or "*Cloudy*". With this coding schema, we can also use the all-one string 111 (or simply denoted as #) to represent the wildcard of "don't care". Continuing the above example, assume each intermediate rule in RS_q is to be encoded as a bit string. We may rewrite rule r'_{q1} as (100;
100; ##; ##: 100). As a result, each intermediate rule in RS_q is encoded into a chromosome as shown in Fig. 3. It should be mentioned that our coding method is suitable to
represent multi-value logic with OR relations among terms within each attribute, and
the AND relations among attributes.

	Outlook	Temperature	Humidity	Wind	Sports
r'_{q1}	100	100	##	##	100
r'_{q2}	010	###	##	01	010
r'_{q3}	001	001	##	##	001

Fig. 3. Bit-String Representation of RS_q

4 Knowledge Integration Phase

In the knowledge-integration phase, genetic operations and credit assignment are
applied at the rule level. In our approach, the initial set of bit strings for rules comes
from multiple knowledge sources. Each individual within the population is a rule, and
is of fixed length. Good rules are then selected for genetic operations to produce better offspring rules. The genetic process runs generation after generation until certain
criteria have been met. After evolution, all the rules in a population are then combined to form a resulting rule set.

4.1 The Strength of a Rule

The strength of the fitness of a rule can be measured jointly by its coverage, accuracy, and relative contribution among all the rules in the population [10]. Let U be the set of test objects. The coverage of a rule derived rule (r_i) is defined as follows:

$$Cg(r_i) = \frac{\left|\Gamma_{r_i}^U\right|}{n} \tag{1}$$

Where $\Gamma_{r_i}^U$ is the set of test objects in U correctly predicted by r_i. The n is the number of objects in U. The coverage $Cg(r_i)$ is the relative size of this condition set in the entire object space. Obviously, the larger the coverage, the more general the rule is.

The accuracy of a rule r_i is evaluated using test objects as follows:

$$Ac(r_i) = \frac{\left|\Gamma_{r_i}^U\right|}{\left|\Gamma_{r_i}^U\right| + \left|U - \Gamma_{r_i}^U\right|} \tag{2}$$

$(U - \Gamma_{r_i}^U)$ is the set of test objects in U wrongly predicted by r_i. The accuracy of the rule is the measure indicating the degree to which the condition set is the subset of the conclusion set, or the truth that the condition implies the conclusion. Obviously, the higher the accuracy, the better the rule is.

Since an object may be classified by many rules, we want to measure the contribution of each rule in the classification of each object. If an object is classified correctly by only one rule, then this rule has full contribution or credit which equals 1. If an object is classified correctly by n rules, then these rules should share the contribution, and thus each of them has only $1/n$ contribution. The contribution of a rule is the sum of its contribution to correctly classify each object. The contribution of a rule r_i is defined as follows:

$$Cb(r_i) = \sum_{u \in U} \frac{\Psi(r_i, u)}{\sum_{r_i' \in R} \Psi(r_i', u)} \tag{3}$$

Where $\Psi(r, u) = \begin{cases} 1 & \text{if } u \text{ is correctly classified by a rule } r \\ 0 & \text{otherwise} \end{cases}$

The R is the population of all the rules. The u is a test object in U. The contribution measure captures the uniqueness of the rule. A rule with high contribution is generally less overlapping with other rules. Finally, we integrate all the quality measures into a single fitness function. The fitness of a rule r_i is defined by Equation (4).

$$Fit(r_i) = \left(\left(Ac(r_i) - \frac{1}{L} \right) + \frac{Cg(r_i)}{L} \right) Cb(r_i) \tag{4}$$

The L is the number of all possible classes. The Ac(r_i) is subtracted by 1/L represents the accuracy of random guessing among L evenly distributed classes. The reason for making this subtraction is that a useful rule should have more accuracy than random guessing. We use 1/L as the weight for the coverage. When there are more classes, the coverage should have less weight because in this situation accuracy will be more difficult to achieve than coverage. Finally, the sum of the net accuracy and the weighted coverage is multiplied by the Cb(r_i) which represents the rule's competitive contribution in the population.

Table 1. A Set of Test Objects

Case	Outlook			Temperature			Humidity		Windy		Sports		
	Sunny	Cloudy	Rain	Hot	Mild	Cool	Hu-mid	Nor-mal	Windy	Not_Windy	Swim-ming	Vol-leyball	W-Lifting
1	1	0	0	1	0	0	1	0	0	1	1	0	0
2	1	0	0	1	0	0	0	1	0	1	0	1	0
3	0	1	0	1	0	0	0	1	0	1	1	0	0
4	0	1	0	0	1	0	0	1	0	1	0	1	0
5	0	0	1	1	0	0	0	1	1	0	0	0	1
6	0	1	0	0	0	1	1	0	0	1	0	0	1
7	0	0	1	0	0	1	0	1	0	1	0	0	1
8	0	1	0	0	0	1	0	1	0	1	0	1	0
9	1	0	0	1	0	0	1	0	1	0	1	0	0
10	1	0	0	0	0	1	0	1	1	0	0	0	1
11	1	0	0	1	0	0	1	0	0	1	1	0	0
12	0	1	0	0	1	0	0	1	0	1	0	1	0
13	1	0	0	0	1	0	0	1	1	0	0	0	1
14	0	1	0	0	1	0	0	1	1	0	0	0	1
15	0	0	1	0	0	1	1	0	1	0	0	0	1
16	1	0	0	0	1	0	0	1	0	1	0	1	0

5   Experimental Results

Two sample classification problems are used to illustrate the performance of the XCS-KI algorithm. The first is the Saturday Morning Problem, and the second is the Multiplexer Problem. They each have different structures and a different level of difficulties.

5.1   Saturday Morning Problem

We use the "Saturday Morning Problem" to generate sample data based on existing rules and then test if the hidden rules can be discovered from the sample data by a learning algorithm. The training data is shown in Table 1.

To apply our algorithm, the population size is fixed to 200. The initial population consists of 16 rules converted from 16 cases, and 184 rules randomly generated. In each generation, 20% of the population is selected for reproduction. The crossover probability is 0.8, and the mutation probability is 0.04. For selecting the replacement candidate, the subpopulation size is 20 and initial fitness is 0.01.

The number of rules extracted from the population decreases through generations. The final rule set extracted from the 28th generation consists of three rules. The three rules with * are the rules extracted from the population. They are listed in Table 2 along with 12 other rules that have high fitness values. Table 2 also shows the accuracy and coverage of these rules. They all have approximately the highest coverage. Finally, among them the rule with the highest fitness value, 0.22, is selected. The final rule set is denoted as set X which consists of rules X1-X3, and is listed in Table 3(B). For comparison, another set of rules is generated by using the Fuzzy Decision Tree Induction method [9] from the same set of sample data. It is denoted as rule set T which consists of rules T1-T6 and is listed in Table 3(A).

From Table 3, it is worth noting that the swimming rule X1 is equivalent to the combination of swimming rules T1 and T2. The weight-lifting rule X2 is a little more specific than the combination of rules T3, T4 and T5. The volleyball rule X3 is somehow more specific than T6 but is not fully covered by T6. As a result, the rules generated by the XCS algorithm seem to be both more accurate and more general.

Table 2. The Rules in the 28th generation for the Saturday Morning Problem

Rule	Outlook			Temperature			Humidity		Windy		Sports			Accuracy	Coverage	Fitness
	Sunny	Cloudy	Rain	Hot	Mild	Cool	Humid	Normal	Windy	Not_Windy	Swimming	Volleyball	W-Lifting			
1*	0	0	1	0	1	1	1	1	1	0	0	0	1	1	0.335	0.22
2	1	1	0	0	1	0	0	1	1	1	0	1	0	0.92	0.182	0.21
3*	1	1	0	0	1	0	0	1	0	1	0	1	0	1	0.312	0.2
4	1	0	0	1	0	0	1	0	1	1	0	0	1	0.74	0.284	0.2
5	0	0	1	0	1	1	1	1	1	1	1	0	0	0.96	0.142	0.19
6*	1	1	0	1	0	0	1	1	1	1	1	0	0	1	0.324	0.18
7	0	0	1	1	1	1	1	0	0	1	0	1	0	0.72	0.241	0.17
8	1	0	1	1	1	0	1	1	1	1	1	0	0	0.86	0.163	0.16
9	0	1	0	0	0	1	1	1	1	1	0	0	1	0.88	0.237	0.15
10	1	1	1	0	1	1	1	0	1	0	1	0	0	0.66	0.135	0.14
11	0	1	0	1	0	0	1	0	1	1	0	0	1	0.36	0.442	0.14
12	1	0	0	1	0	0	0	1	0	1	0	1	0	0.77	0.215	0.13
13	0	1	1	1	1	0	1	0	1	0	0	1	0	0.83	0.276	0.12
14	1	1	1	1	1	0	1	0	0	1	0	0	1	0.42	0.135	0.11
15	0	1	0	1	1	1	0	1	1	1	1	0	0	0.55	0.112	0.1

5.2   Multiplexer Problem

A simple experiment to demonstrate LCS is the 6-bit multiplexer. A Boolean 6-bit multiplexer is composed out of a 6 bit condition part (input) and a 1 bit action part (output). Figure 4 illustrates the 6-bit multiplexer. The first 2 bits of the input determine the binary address (00,01,10,11) which will select 1 of the 4 data lines. The output is determined by the value on the current selected dateline. Figure 5 shows the resulting classifier format for the 6-bit multiplexer. Where msb is the most significant bit and lsb is the least.

Table 3. Two sets of rules generated by fuzzy ID3 and XCS algorithm

---

A. Rule set T generated by the fuzzy ID3 method

Rule T1: IF (Temperature is Hot) and (Outlook is Sunny) THEN Swimming
Rule T2: IF (Temperature is Hot) and (Outlook is Cloudy) THEN Swimming
Rule T3: IF (Temperature is Mild) and (Wind is Windy) THEN W-lifting
Rule T4: IF (Temperature is Cool) THEN W-lifting
Rule T5: IF (Outlook is Rain) THEN W-lifting
Rule T6: IF (Temperature is Mild) and (Wind is Not_Windy) THEN Volleyball

---

B. Rule set X generated by the XCS algorithm

Rule X1: IF (Outlook is Sunny or Cloudy) and (Temperature is Hot) THEN Swimming
Rule X2: IF (Outlook is Rain) and (Temperature is Mild or Cool) and (Wind is Windy)
        THEN W-lifting
Rule X3: IF (Outlook is Sunny or Cloudy) and (Temperature is Mild) and (Humidity is
        Normal) and (Wind is Not_Windy) THEN Volleyball

Fig. 4. A 6-bit multiplexer

A1	A2	D0	D1	D2	D3	v
Bit 1	Bit 2	Bit 3	Bit 4	Bit 5	Bit 6	Bit 7
Address	Address	Data	Data	Data	Data	Data
msb	lsb	00	01	10	11	

Fig. 5. The classifier format for a 6-bit multiplexer

   With our multi-value logic coding method, each input and output has two values "On" and "Off" and thus can be represented by a string of two bits. We use *10* to represent "On", *01* to represent "Off", and # to represent a wildcard "don't care". As an example, (0;1;0;0;1;1) has an output of (0) because the input at data line 01 is 0. The format for this example is (0;1;0;0;1;1:0). It is clear that the input values of the data lines 00, 10 and 11 in this example are not needed to generate the output. Fur-

ther, the above case then can be represented as (01;10;01;01;10;10:01). The whole behavior of the 6-bit multiplexer can be summarized in the following eight rules:

Rule 1:   (01;01;01;##;##;##:01)
Rule 2:   (01;01;10;##;##;##:10)
Rule 3:   (01;10;##;01;##;##:01)
Rule 4:   (01;10;##;10;##;##:10)
Rule 5:   (10;01;##;##;01;##:01)
Rule 6:   (10;01;##;##;10;##:10)
Rule 7:   (10;10;##;##;##;01:01)
Rule 8:   (10;10;##;##;##;10:10)

   Since there are 6 binary inputs and 1 binary output, only a total of 64 different cases can be used for training. However, with a possible wildcard, we could have $3^7 = 2187$ different rules. The total number of different rule sets then is 2^{2187}. To determine the optimal set of rules in this huge search space is not an easy task. We used the XCS-KI algorithm to learn the rules. To apply our algorithm, the population size is fixed to 200. In each generation, 20% of the population is selected for reproduction. The crossover probability is 0.8, and the mutation probability is 0.04. For selecting the replacement candidate, the subpopulation size is 20 and the initial fitness is 0.01. The algorithm converges quickly. The optimal set of eight best rules is extracted from the 48th generation. For comparison, another set of eight best rules is generated by using the Fuzzy Decision Tree Induction method from the 68th generation. The number of rules extracted in each generation is illustrated in Figure 6. During the evolution process it gradually decreases, indicating the improvement of the rule set.
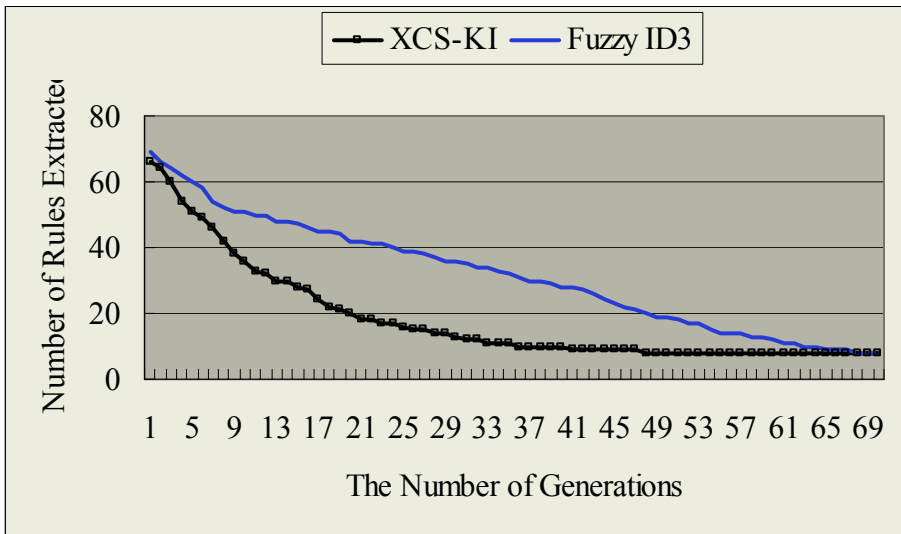
Fig. 6. The evolution process

6 Conclusion

In this paper, we have shown how the knowledge coding and knowledge integration methodology can be effectively represented and addressed by the proposed XCS-KI algorithm. The main contributions are: (1) it needs no human experts' intervention in the knowledge integration process; (2) it is capable of generating classification rules that can be applied when the number of rule sets to be integrated increases; (3) it uses three criteria (accuracy, coverage, and fitness) to apply the knowledge extraction process which is very effective in selecting an optimal set of rules from a large population. The experiments prove that the rule sets derived by the proposed approach are more accurate than those obtained by the Fuzzy ID3 algorithm.

References

1. Baral, C., Kraus, S., and Minker, J.: Combining Multiple Knowledge Bases. IEEE Transactions on Knowledge and Data Engineering, Vol. 3, No. 2 (1991) 208−220
2. Boose, J. H., and Bardshaw, J.M.: Expertise Transfer and Complex Problems: Using AQUINAS as a Knowledge-Acquisition Workbench for Knowledge-based Systems. International Journal of Man−Machine Studies, Vol. 26 (1987) 3−28
3. Holland, J.H.: Adaptation in Natural and Artificial Systems. University Press of Michigan, Ann Arbor (1975)
4. Holland, J.H., Reitman, J.S.: Cognitive Systems Based on Adaptive Algorithms. In: Waterman, D.A., Hayes-Roth, F. (Eds.), Pattern directed interference systems. Academic Press, New York (1978) 313-329
5. Holmes, J.H.: Evolution-assisted Discovery of Sentinel Features in Epidemiologic Surveillance, Ph.D. thesis, Drexel University, Philadelphia, PA (1996)
6. Quinlan, J.: Induction of Decision Tree. Machine learning, Vol. 1 (1986) 81−106
7. Stolzmann, W.: An Introduction to Anticipatory Classifier Systems. Lecture Notes in Artificial Intelligence, Vol. 1813, Springer, Berlin (2000) 175−194
8. Wilson S.W.: Rule Strength Based on Accuracy. Evolutionary Computation, Vol. 3, No. 2 (1996) 143-175
9. Yuan, Y., and Shaw, M. J.: Induction of Fuzzy Decision Trees. Fuzzy Sets and Systems, Vol. 69 (1995) 125−139
10. Yuan, Y., and Zhuang, H.: A Genetic Algorithm for Generating Fuzzy Classification Rules. Fuzzy Sets and Systems, Vol. 84 (1996) 1−19

Learning-by-Doing
Through Metaphorical Simulation⋆

Pedro Pablo Gómez-Martín,
Marco Antonio Gómez-Martín, and Pedro A. González-Calero

Dep. Sistemas Informáticos y Programación
Universidad Complutense de Madrid, Spain
{pedrop,marcoa,pedro}@sip.ucm.es

Abstract. After the doubtful success of content-based e-learning systems, simulations are gaining momentum within the e-learning community. Along this line we are working on JV^2M, a simulation-based learning environment to teach the Java Virtual Machine (JVM) and the compilation of object-oriented languages. This paper describes both the metaphorical simulation of the JVM and the knowledge our system possesses and details an execution example that reflects how all the information is used on it.

1 Introduction

Learning-by-doing is an experience-based style of learning where the student is aware what she needs to learn during the resolution of a problem. This way of learning is inherited from the days of the apprentices of manual works as carpenter or plumber. Computer virtual environments allows us to build simulators of machines and situations where a student acts as she would do in the real world, developing proficiency in a controlled and safe environment. For example [8] presents a simulation scenario to instruct tactical action officers, and in [7] the student boards into a virtual ship to learn how a high-pressure air compressor works. It can be also possible to create virtual environments that have not counterpart in the real world. For example, [3] presents a metaphorical computer where the user has to fetch instructions from the main memory to the CPU and execute them.

Similar to this idea, we propose the use of the learning-by-doing approach to teach the compilation of object-oriented languages, in particular teaching Java compilation. The resulting system includes a metaphorical simulation of the Java Virtual Machine (JVM) [6], where the student explores compilation scenarios. Using ideas of pedagogical agents (to mention just a few see [3] [4] [5] and [7]) our learning environment is complemented with a pedagogical agent, a human-like figure that assists the user in the learning process. Interaction is similar to a 3D-graphical adventure game that tries to guarantee student motivation while she is solving the exercises in the virtual environment. The system is called JV^2M,

---
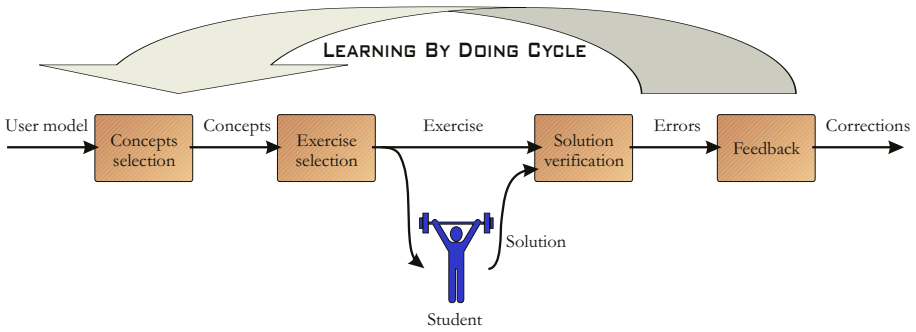
Fig. 1. Learning-by-doing cycle

meaning twice virtual because of the representation of the Java Virtual Machine in a virtual environment.

The rest of the paper runs as follows: next section describes our vision of the learning-by-doing approach of teaching. Section 3 describes the JV^{2}M main features. Section 4 introduces the system knowledge and how the application and the pedagogical agent use it on each execution, while Section 5 explains one example of exercise execution. Finally, in Section 6 conclusions and future work are presented.

2   Learning by Doing Cycle

We claim that trying to learn skills, abilities or creativity in a comfortable passive position is absolutely useless and ineffective. The learning-by-doing approach attacks this problem by giving students exercises that are considered a valuable learning experience. Each problem is selected for a purpose, putting into practise some concepts of the taught domain. In order to be effective, problems selection must be adequate, following a pedagogical direction. A balance is needed between the exercise difficulty and the student knowledge in order to neither overestimate nor bore her.

Our utilisation of the learning-by-doing model unfolds in a cycle, shown in Figure 1. It consists of five steps:

1. *Selection of concepts to practise:* concepts of the taught domain are chosen. As said before, if we want a successful learning, this selection cannot be arbitrary but depending on the current student knowledge. This selection opens the door for pedagogical decisions.
2. *Exercise selection:* using the previous selected concepts, the student is provided with an exercise to put them into practise.
3. *Exercise resolution:* the pupil solves (or tries to solve) the problem she faces.
4. *Solution verification:* the student answer is tested (compared with the system solution) in order to evaluate its correctness.

5. *Feedback:* using the conclusions reached in the previous step, advices are provided to the student. The tutor can give explanations about the mistakes made or even supply theoretical knowledge when needed.

This cycle is valid for both traditional teaching system, driven by human tutors, and for computer stand-alone application where the teacher is an interactive program, such as the Intelligent Tutoring Systems (ITS) [8].

When the student solves the exercise using a computer, the three last stages usually blend. As the system often provides contextualised help, it needs to test continuously the partial solution (stage 4) and to provide feedback before the student entirely finishes her answer (stage 5).

3   JV^{2}M: General Description

JV^{2}M is a system to teach the compilation process of high-level programming languages, or *how source code is translated into object code.* The skills we want to teach are, for example, code generation of the expressions and control structures. But to become proficient in these tasks, the student needs to know the object language that the compilation process has to generate. In our system, this is the language interpreted by the Java Virtual Machine [6] (JVM), that should be thought as the *assembly language* of this machine and it is known as Java *byte-codes.* Instead of forcing the user to know that language in advance, our system teaches it, becoming an educational program focused in two different aspects: *high-level concepts* related to the compilation process and *low-level concepts* related to the JVM structure and instructions.

JV^{2}M does not aspire to provide a theoretical and formal description of the compilers field. It tries to produce an intuitive learning of the process, based on exercises and conforming to the learning-by-doing approach described in the previous section. Exercises consist on the source code of a Java program that must be compiled. The way the next exercise is selected (first and second phases in the learning-by-doing cycle) is out of the scope of this paper.

Let's assume the next exercise has been chosen. A typical web-based learning application would ask the user to introduce, in some way, the translated object code. We are using a different approach: the student *executes* the compiled code instead of just writing it down. This compiled code is executed in a 3D virtual environment that symbolises a metaphorical JVM. The system implements a minimal Java Virtual Machine, only with its structures, to allow the virtual environment to present its state at every moment. Different objects and characters appear, each of them representing a specific structure of the real JVM. The user is symbolised by an avatar that inhabits in the virtual world (see Figure 6).

The student has an inventory where objects are kept. These objects represent, for example, operands, constants, references to classes, and other intermediate elements needed to execute the JVM instructions in the metaphorical world.

The learning environment is enhanced with a pedagogical agent called Javy (**J**av**a** taught **V**irtuall**y**) who supports the *feedback* phase. This avatar can provide the user with contextualised help about the exercise being solved. When

the user asks for help, the agent provides different information depending on the current exercise and state of the resolution process. Additionally, JAVY is able to finish by his own the problem, when the user requests it.

When the user wants to ask something to JAVY, she starts a conversation in a similar way to the used in some interactive entertainment software as *Grim Fandango* from LucasArts. Concretely, the user approaches to JAVY, and *the agent* provides a set of valid sentences to be said. The user chooses one of them, and JAVY answers it, providing a second set of related questions and starting a new conversation cycle. The next section will focus in the way this is accomplished.

4   System Knowledge

In order to be effective, every learning system needs to know whether the user has performed the task correctly or not once the resolution phase is finished. When the system possesses a pedagogical agent, it will need to monitor the solving state *throughout all the resolution process* (not just at the end) in order to be able to help the student when she gets stuck or asks for advice. In other words, the resolution (third phase in Figure 1) runs simultaneously to the solution verification and feedback stages (forth and fifth phases).

As said before, our system expects to teach the Java compilation process. It possesses a set of exercises that become the first source of knowledge. Exercises are indexed in order to be selected in the two first phases of the learning-by-doing cycle depending on the current student. Internally, each exercise is composed of pieces of Java code and their compiled versions. Exercises are designed by human tutors who enrich them with annotations about the way the translation process has taken place in them. Exercises support the teaching of the previously called "high-level" concepts, in other words, the compilation process.

Students must *execute* the JVM instructions of the compiled code. In order to know if they are acting correctly, the system has information about how JVM instructions are performed, enhanced with suitable explanations. This knowledge is also built by human tutors. It will be available at any moment, and it is shared by all the exercises. This knowledge sustains the "low-level" domain, the JVM execution model.

The joint point between both levels is the third source of knowledge, the conceptual hierarchy that stores the domain concepts students have to learn. It contains general explanations of all the concepts (not fit to a specific problem or JVM instruction). The next subsections will describe these knowledge parts, starting with the latter, the *conceptual hierarchy*.

4.1   Conceptual Hierarchy

The conceptual hierarchy is a net of related concepts. Each concept, has a *name*, a *textual description* that details it, and some *connections* with other related concepts.

The textual description is used by JAVY to explain the concept when the user asks for help. After one of such utterances (where JAVY just "reads" the
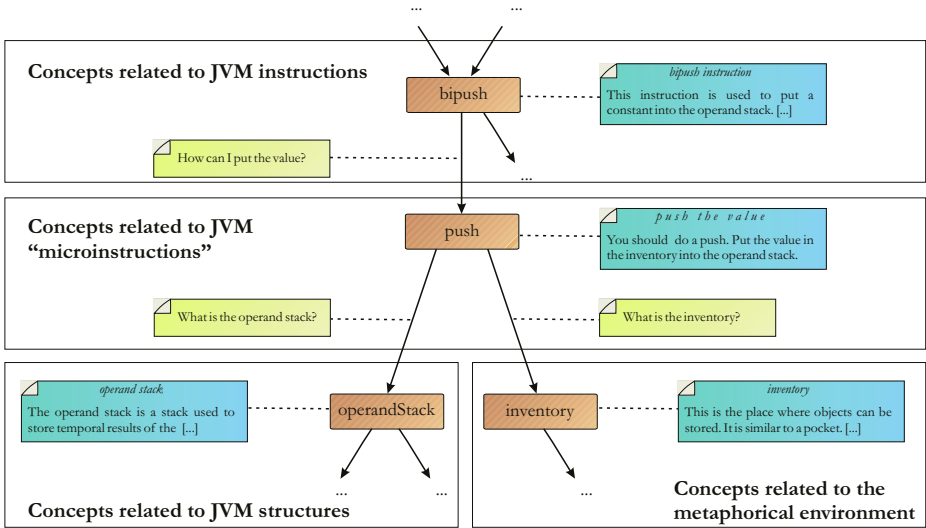
Fig. 2. Detailed conceptual hierarchy fragment

description), the agent must offer a set of questions the user can ask, related to the last explanation. We get the connected concepts using the *links* in the conceptual hierarchy. In order for JAVY to propose a question for each new related concepts, those links are enriched with a sentence used to guide the conversation. In that way, the next questions the user can ask are obtained directly from the links of that concept in the conceptual hierarchy.

Figure 2 exemplifies this idea. For instance, when JAVY is asked about the `push` concept, he provides the attached explanation: "You should do a push. Put the value in the inventory into the operand stack". Then, he iterates through the child nodes and recover two possible questions to provide the user: "What is the operand stack?" and "What is the inventory?". The conversation cycle is repeated depending on the student selection.

4.2   Execution Graphs

The system has to store in some way how each JVM instruction is executed. We have analysed all the instructions in order to identify the different steps (*microinstructions*) needed to perform all of them. We consider a microinstructions as an *atomic manipulation of a JVM structure*. The previously described conceptual hierarchy possesses a concept for each *microinstruction*, containing a general description about its meaning.

The system also stores an *execution graph* for each instruction. Graph nodes represent execution states for each instruction, and edges let advance the execution performing a microinstruction. Edges also store an explanation that exposes why the attached microinstruction is valid. Therefore, conceptual hier-
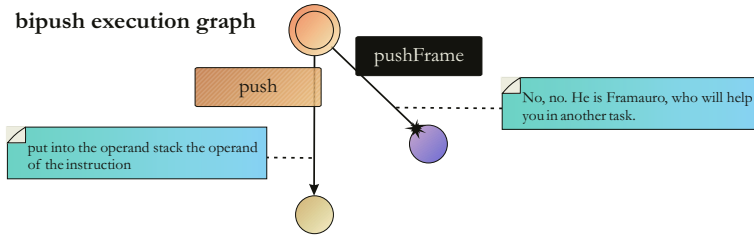
Fig. 3. Simple execution graph with an invalid path

archy maintains a general description of every microinstruction, and execution graphs contextualise them in the instructions where they are used.

Figure 3 shows the execution graph of a quite simple JVM instruction (`bipush`). Just the `push` microinstruction must be performed. As the figure shows, the edge that triggers its execution has been enriched with the explanation: "put into the operand stack the operand of the instruction". If the user makes a mistake doing a different step, Javy uses this text and mixes it with a *template* to assemble his advice: "If I were you, I would try to *put into the operand stack the operand of the instruction*".

If the student insists on going wrong, Javy uses the *name* of the concept for the valid microinstruction in the conceptual hierarchy (in this case "push the value", as Figure 2 shows) and another template. The agent would say: "If you want to *put into the operand stack the operand of the instruction*, you should try to *push the value*").

The execution graphs can explicitly keep invalid paths, with specialised explanations about common mistakes, so Javy can provide specific advices. Figure 3 shows that `pushFrame` microinstruction is invalid, and, if performed, the agent would say "No, no. He is Framauro, who will help you in another task".

4.3   Exercises

As said previously, exercises are composed of Java source code and its compiled version. They are developed by human tutors, assisted by an authoring tool. Source and object code is arranged in a tree of blocks, usually quite similar to the compiler syntactic tree. Tutors are supposed to attach an explanation to each block, referring to why it has been compiled in that way.

The tree of code blocks is also enriched with sentences in the edges that join the nodes. Javy will use them to guide the conversation. Figure 4 shows an exercise fragment, where each block in the tree has been labeled with a number for clarity. When executing the first JVM instruction (`bipush 5`) the user is allowed to ask the agent why the first operand in the "5*3" expression has been translated into it. Javy will answer "In order to calculate 5*3, you must stack the first operand, in other words, the number 5". As the link between the block `1` and its parent is labeled with "Why?", following the previous explanation the user would be able to ask the reason of the action to go up in the tree.
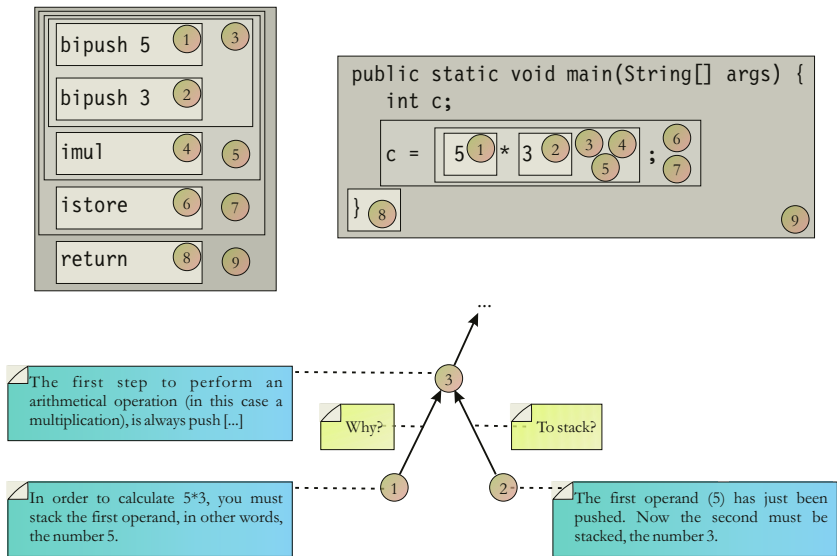
Fig. 4. Detailed exercise fragment

4.4   Putting All the Pieces Together

The three previous pieces of knowledge are not isolated, but interrelated by additional links, becoming a big source of information used by Javy to talk to the user. For example, code blocks in the exercises are usually linked with nodes in the conceptual hierarchy. Usually small blocks referring to one JVM instruction will be related with the concept of that instruction, and bigger blocks will be connected with higher concepts concerning to compilation issues.

All these additional links are also marked with questions that are provided to the user when the current conversation state is located in the node which is the source of the edge. Figure 5 sketchs this general structure.

5   Detailed Example

In order to settle the previous section, a short example will be now detailed[1]. The user is faced with the execution of the exercise presented in Figure 4. Using the keyboard, the student can consult the Java source code. At the very beginning, the first JVM instruction (`bipush 5`) must be executed. This instruction is related to the number `5` of the source code, that is shown highlighted. Figure 6 shows this situation.

The current instruction, "`bipush 5`", has a parameter that has been automatically put into the user inventory as a virtual object. Let's assume the student selects it and tries to give it to Framauro, an auxiliar character used

---

[1] Available in `http://gaia.sip.ucm.es/grupo/projects/javy/screenshots.html`
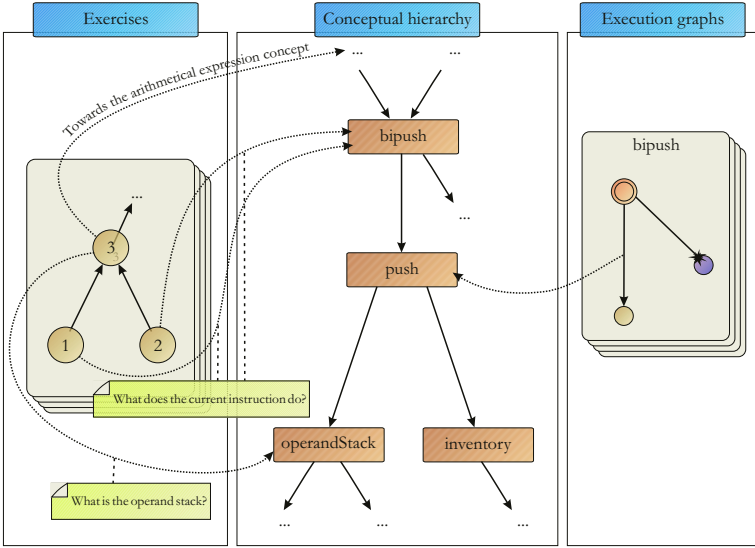
Fig. 5. General vision of the JAVY knowledge

to manipulate the JVM structure known as *frame stack*. Concretely, this action in the virtual environment corresponds to the `pushFrame` microinstruction. Using the current execution graph (shown in Figure 3), JAVY categorises this step as invalid, and provides the explanation adapted to this specific error from the graph: "No, no. He is Framauro, who will help you in another task".

Eventually, the user feels lost and asks JAVY. The agent must provide some questions the user can ask. In *the first conversation step*, he always provides the same alternatives:

- What am I supposed to do?
- Why do I have to do this instruction?
- Don't worry. I will manage.

The last sentence is available at any point of the conversation. If the user selects it, the dialogue will finish. The other two sentences are more interesting. Although they are the same at the beginning of all the conversations, they link with different pieces of information depending on the context.

The first sentence is related to the next microinstruction to be executed, established by the execution graph. When selected, JAVY uses the attached description and the template described in Section 4.2 to advise the user: "In order to *put into the operand stack the operand of the instruction* you should try to *push the value*".

After that explanation, the agent provides a different valid question, assembled using a template and the *name* of the microinstruction in the conceptual hierarchy: "*push the value*? Could you tell me more?". This sentence links with the concept in the hierarchy referring to the current microinstruction.

Fig. 6. System screenshoot

When the student chooses it, JAVY consults his knowledge to provide the attached explanation (that, as shown in Figure 2, is: "You should do a push. Put the value in the inventory into the operand stack"). In that moment, the user will be able to go deeper into the conceptual hierarchy as described in Section 4.1.

The previous description should make clear that the first question available to the the user at the very beginning of the conversation will let her get information about the *concepts* referring to the JVM instructions and structures. The specific information provided will depend on the current instruction and microinstruction.

On the other hand, the second sentence ("Why do I have to do this instruction?") always links to the exercise information, specifically with the deeper block containing the current JVM instruction. Let's assume the user has correctly executed the first JVM instruction and she is now facing the second one, `bipush 3`. If the user insists on making mistakes, JAVY will provided tips built as described in Section 4.2.

When JAVY is asked, he provides the same three possible questions. Now, we will focus on the second one. As just said, it links with the deeper code block in the exercise information, in this case the chunk labeled with number 2 in Figure 4. When selected, JAVY provides the attached explanation: "The first operand (5) has just been pushed. Now the second must be stacked, the number 3.".

Now the user will have *two* questions available. The first one links with the *parent code block* shown in Figure 4 (labeled with number 3). This connection will let the user to browse the code block tree, getting information about the compilation process, in other words, the *high-level concepts* taught by the system.

The second available question links with the *conceptual hierarchy* using the connection shown in Figure 5. The provided sentence is "What does the current instruction do?", that ends in the `bipush` concept.

6   Conclusions and Future Work

In this paper we have presented JV^{2}M and Javy, a pedagogical agent that inhabits in a learning environment for teaching Java compilation and the Java Virtual Machine internals. The text describes in some detail the knowledge stored by the system and its use has been explain by means of an example.

However, the success of the application relies on the availability of a huge exercise base, in order to recover the more suitable exercise to be presented to the student. Our current research focus on using Case-Based Reasoning [1] and its adaptation phase to reduce the amount of scenarios the human tutor has to create and insert into the system.

Finally, though explanations provided by Javy are contextualized enough, its generation is quite hard-coded. We are considering to use the current system as a test bed for using Conversational Case-Based Reasoning [2] in order to guide the conversation with Javy.

References

1. A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1):39–59, 3 1994.
2. D. W. Aha, L. Breslow, and H. Muñoz-Avila. Conversational case-based reasoning. *Applied Intelligence*, 14(1):9–32, 2001.
3. W. Bares, L. Zettlemoyer, and J. C. Lester. Habitable 3D learning environments for situated learning. In *Proceedings of the Fourth International Conference on Intelligent Tutoring Systems*, pages 76–85, San Antonio, TX, August 1998.
4. J. C. Lester, S. A. Converse, S. E. Kahler, S. T. Barlow, B. A. Stone, and R. Bhogal. The persona effect: affective impact of animated pedagogical agents. In *Proceedings Human Factors in Computing Systems (CHI'97)*, pages 359–366, Atlanta, March 1997.
5. J. C. Lester, J. L. Voerman, S. Towns, and C. B. Callaway. Cosmo: A life-like animated pedagogical agent with deictic believability. In *Working Notes of the IJCAI '97 Workshop on Animated Interface Agents: Making Them Intelligent*, pages 61–69, Nagoya, Japan, August 1997.
6. T. Lindholm and F. Yellin. *The Java Virtual Machine Specification. 2nd Edition.* Addison-Wesley, Oxford, 1999.
7. J. Rickel and W. L. Johnson. Animated agents for procedural training in virtual reality: perception, cognition and motor control. *Applied Artificial Intelligence*, 13(4):343–382, 1999.
8. R. H. Stottler. Tactical Action Officer Intelligent Tutoring System (TAO ITS). In *Proceedings of the Industry/Interservice, Training, Simulation & Education Conference (I/ITSEC 2000)*, November 2000.

Emergence of Immune Memory and Tolerance in an Asymmetric Idiotype Network

Kouji Harada

Tokuyama College of Technology
3538, Takajo, Kume, Shunan, Yamaguchi 745-8585, Japan
k-harada@tokuyama.ac.jp

Abstract. This study proposes an idiotype network model system
adopted an "asymmetric" idiotype - anti-idiotypic interaction, and shows
realizing immune tolerance as well as immune memory on a transient dy-
namics. To date, the immune tolerance was a problematic phenomenon
for traditional idiotype network models adopted "symmetric" idiotypic
interaction because they failed to reproduce it. This paper reports the
proposed model succeeds in the display of the immune tolerance by con-
sidering an asymmetric anti-idiotypic interaction. Actually, a computa-
tional experiment clarifies that its establishment associates with high
anti-idiotypic idiotype's population level of when antigens are re-dosed.
This result indicates what the anti-idiotypic idiotype functions effec-
tively plays a decisive role for the establishment of the immune tolerance.
Lastly, this study closes at the suggestion that the asymmetric modeling
is more immunologically relevant than the symmetric one which most
theoreticians supported so far.

1 Introduction

For researchers studying mechanisms of immune networks, whether a style of
an activation-suppression interaction between lymphocytes is "symmetric" or
"asymmetric", is a controversial problem. N.K.Jerne who proposed a ground-
breaking immune network theory called idiotype network theory accepted the
idea of the asymmetric interaction[1], and he made some binding sites on a
lymphocyte receptor discriminated between "being recognized sites" and "rec-
ognizing sites" . However, G.A.Hoffman raised a penetrating question about
this asymmetrical idea from the requisition that three major dynamical fixed
points representing for a virgin, an immune and a tolerant state must be stable:
"Hoffmann's stability criteria". He proved these fixed points except the virgin
state's fixed point become unstable in so-called a "prey-predator" type of the
asymmetric idiotype network model, and considered the asymmetrical model
unacceptable from the stability criteria [2]. In response to the consideration, he
proposed a plus-minus "symmetrical" idiotype network model. Thereafter many
network models based on the symmetrical idea, as represented by the bell-shaped
network models [3] were proposed and examined enthusiastically. Much of the in-
vestigations clarified symmetrical network models also had some problems from

a view of the idiotypic regulation. One major problem is that, although the symmetrical network model system possessed a suppressed state, the system does not have any procedure to reach the suppressed state from the virgin state[4]. That is, the suppressed state is a "unreachable state" by any antigenic perturbations. Then, in response to the insufficiency in the symmetrical network modeling approach, my previous study reconsidered an auto-catalytic idiotype network model adopted an asymmetrical idiotypic interaction, then clarified the model can retain immune memory against an antigen[5]. The most remarkable point with the development of the immune memory is that it is realized on a "transient" dynamics, not on a stable fixed point. The result means we do not have to stick to Hoffmann's stable criteria and to abort the idea of the asymmetrical idiotypic interaction. On the basis of the discussion, this study also adopts the asymmetric idea, in addition to that, considers newly an anti-idiotypic idiotype. This paper reports that by the introduction of the new immunological object, the proposed model can achieve the immune tolerance as well as the immune memory on a transient dynamics, however, for the success of the immune tolerance, there is the restriction that an antigen must be re-dosed just when the anti-idiotypic idiotype's population reaches the maximum value.

2   Description of the Model

Each lymphocyte has its own recognition molecular and its specific three dimensional shape is called *idiotype*[1]. The idiotype is characterized by a "recognizing" receptor site and a "being-recognized" ligand site, called *paratope* and *idiotope* respectively. An antigen has its own proper ligand site called *antigenic idiotope*. The strength of interaction between a paratope and an (antigenic) idiotope is determined according to degree of *specificity* between their respective three-dimensional shapes.

This paper presents an idiotype network dynamical model composed of an antigen, an antigen-specific idiotype and an anti-idiotypic idiotype, which facilitates activation and suppression to the idiotype through the idiotypic interaction. $X_1(t), X_2(t)$ and $A(t)$ represent population size of the idiotype, the anti-idiotypic idiotype and the antigen at the time, t, respectively. Following the idiotype network hypothesis[1], a growth dynamics of population of the idiotype, $X_1(t)$ and its anti-idiotypic idioype, $X_2(t)$ is described by,

$$\dot{X}_1(t) = (b - b')X_1(t)X_2(t) - dX_1(t) + s + b''X_1(t)A(t), \tag{1}$$

$$\dot{X}_2(t) = (b' - b)X_1(t)X_2(t) - dX_2(t) + s. \tag{2}$$

The antigen population, $A(t)$ follows a dynamics,

$$\dot{A}(t) = -b''X_1(t)A(t). \tag{3}$$

Figure 1 illustrates mutual asymmetric idiotypic interactions of among the three species, the antigen, the antigen specific idiotype and the anti-idiotypic
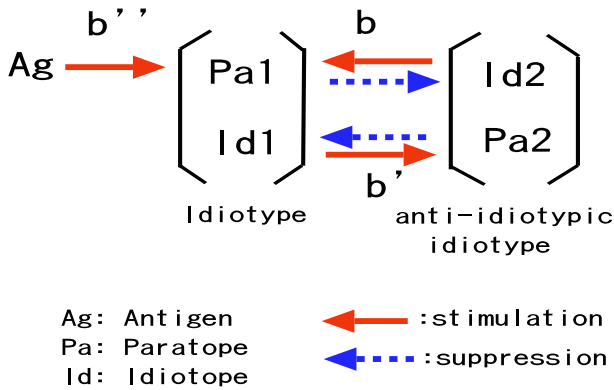
Fig. 1. Asymmetric idiotype - anti idiotypic idiotype network.

idiotype. The parameters b, b' and b'' mean the strength of the interaction of the idiotype's paratope - its anti idiotype's idiotope, the idiotype's idiotope - its anti-idiotype's paratope, and the antigenic idiotope - the antigen-specific paratope, respectively. The parameters d and s are dumping and source rate of an idiotype, respectively.

This study approximately solves the differential equations using the Runge-Kutta method of order 4 with a time-step size, $h = 10^{-2}$, and the system parameters were selected: $b = 0.1$, $b' = 15.0$, $b'' = 0.1$, $s = 10^{-6}$ and $d = 0.1$.

3   Long-Term Immune Memory

This section presents the proposed idiotype network model can retain immune memory by observing the enhancement of the second immune response against the primary one for two times of antigen doses. Figure 2 (a) shows an immune response when the same amount of the antigen with 1.1 units are dosed two times at the times, 200 and 1700, then the dynamic behavior indicates the second immune response is more boosted than the first one. Namely, this model system can acquire immune memory against the dosed antigen.

Next, I explain a mechanism of the development of the immune memory from the nullcline analysis of the model equations. A nullcline of a variable is given as a characteristic line which satisfies the condition of the time-derivative of the variable being zero, thus X_1-nullcline $X_1(A)$ as a function of the variable A, is given as,

$$X_1(A) = \frac{\gamma - \sqrt{\gamma^2 + 4(b' - b)(b''A - d)sd}}{2(b''A - d)(b - b')}, \tag{4}$$

where

$$\gamma = 2(b' - b)s - (b''A - d)d. \tag{5}$$

Also, a nullcline of the variable A as a function of the variable X_1, is given as,

$$A(X_1) = 0. \tag{6}$$

Figure 2 (b) shows $X_1(A)$ and $A(X_1)$ together with the orbit representing for the immune memory in Figure 2 (a), and an intersection point of the two nullclines means a unique stable fixed point, which is called "virgin state" because this state does not experience an antigenic infection. This virgin state is given as an initial state of the system in simulations.
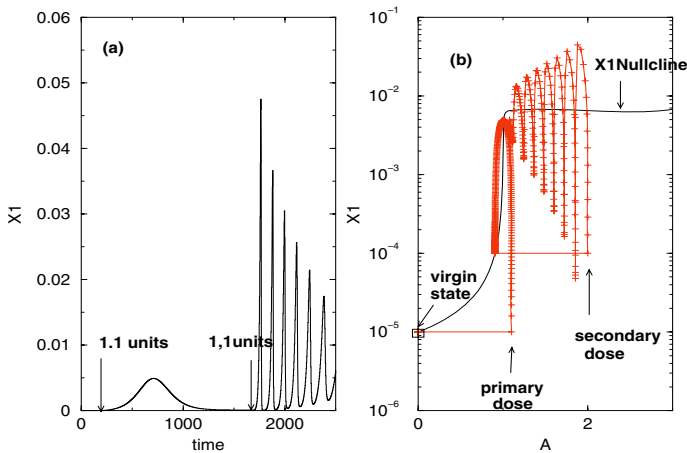
Fig. 2. Immune memory: In Figure (a), when 1.1 units of antigen is dosed twice, at the times, 200 and 1700, an accelerated and boosted recall immune response because of re-exposure to the primary antigen is observed. Figure (b) plots the thinned-out trajectory (crosses) of the immune memory response in Figure (a) together with the nullclines Eqs (4) and (6) on the $A - X_1$ plane.

A main reason for the establishment of the immune memory is that, after the primary dose of the antigen, the system's state takes much time to go back to the initial state, therefore can remain at a transient state, $(A, X_1) = (1.0, 10.0^{-4})$ on the X_1 nullcline for a long interval. Because the relaxation rapidly becomes slow on the nullcline. That results in boosting an immune response at the second dose of the antigen, even though the amount of the first and second antigen dose is the same, 1.1. The previous study presented that this type of immune memory appeared in a single self-catalytic idiotype model. In this study, on the basis of past achievements, it has demonstrated that the same type of the transient immune memory emerges even in a new model system including the anti-idiotypic idiotype. This result indicates a possibility that this type of immune memory succeeds even in asymmetrical and more high-dimensional idiotype networks.

4 Immune Tolerance and Timing of Antigen Dose

Immune tolerance is the immunological state of antigen-specific idiotype becoming unresponsive to the re-exposure of the same antigen as the primary one. It is also well-known as "negative" immune memory in immunology [6]. This section discusses the proposed network model can realize the immune tolerance against an antigen.
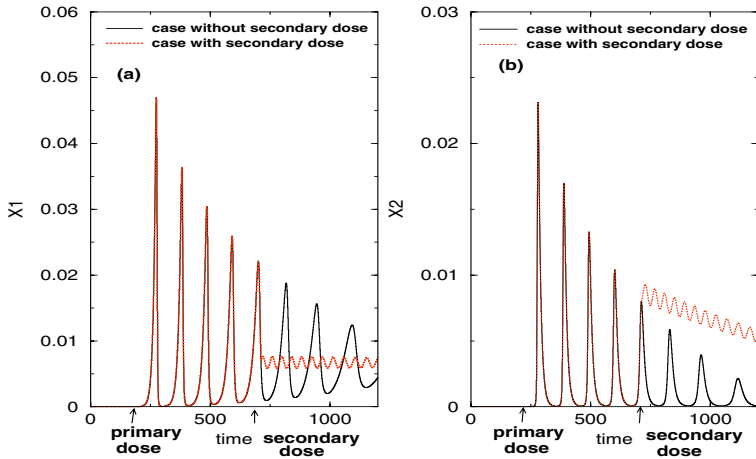
Fig. 3. Immune tolerance : a dotted curve in Figure (a) represents time evolution of the antigen-specific idiotype population X_1, against two consecutive doses of the same antigen with 2.1 units, at the time, 200 and 710. The population, X_1 become suppressed by re-dosing the antigen comparing with a case without considering secondary antigen dose (solid line). Figure (b) shows time evolution of the anti-idiotypic idiotype population, X_2 in the simulation of Figure (a). The population, X2 become sustained at a higher level by re-dosing the antigen at the time, 710.

In Figure 3 (a), a dotted and solid curves respectively represent time evolution of the antigen-specific idiotype population, X1 in the case of an antigen with 2.1 units being dosed two times at 200 and 710, and for comparison, in the absence of the secondary antigen dose at the time, 710. Comparing the solid line with the dotted one, we can observe the secondary antigen dose with 2.1 units suppresses the oscillatory immune response which the primary antigen induced. In other words, the system can become immune tolerance against the antigen. Then, we would like to focus on the behavior of the anti-idiotypic idiotype population, X2 in Figure 3 (b), in order to find out a reason why the immune tolerance becomes possible. Actually when the immune suppression against the antigen occurs, we can observe the anti-idiotypic idiotype population booming and sustained in a high level. That is, the highly sustained anti-idiotypic idiotype population enables the immune tolerance against the antigen.
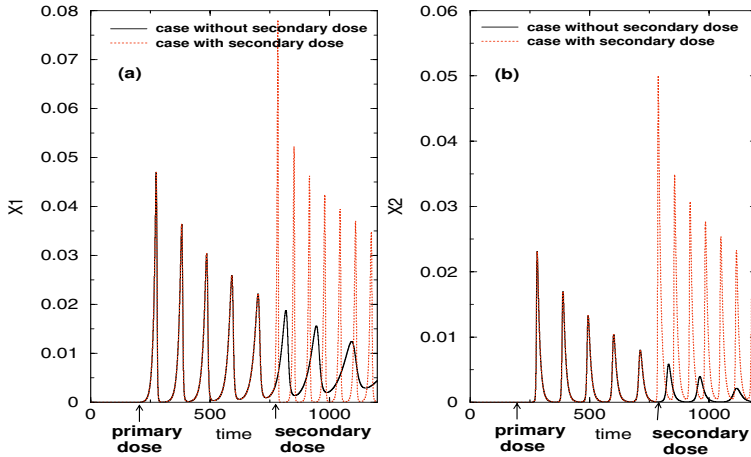
Fig. 4. Immune-response pattern depends on timing of an antigen dose: Each simulation of Figure (a) and (b) is the same as one of Figure 3 (a) and (b) except difference of timing of secondary antigen dose. Figure (a) shows that the antigen-specific idiotype population, X_1 is not suppressed as we have seen in Figure 3 (a), but rather boosted if the second antigen is dosed when the anti-idiotypic idiotype population is at a minimal level as we can confirm in Figure (b).

However, the immune tolerance does not always occur. Figure 3 and 4 suggest we need to adjust the secondary antigen-dose timing to succeed in the immune tolerance against the antigen. In fact, in the situation that the immune tolerance has succeeded, the secondary antigen-dose timing has been adjusted when the anti-idiotypic idiotype population, X2 reaches to a maximum value (see Figure 3 (a),(b)). If you upset the timing by design, and dose the antigen at the other timing when the anti-idiotypic idiotypic population, X2 is in a bust, then you must observe that the idiotype X1 is not suppressed but rather excited (see Figure 4 (a),(b)). To sum up, this investigation suggests that the immune tolerance against the antigen emerges dynamically depending on a level of the anti-idiotypic idiotype population of when the antigen is re-dosed.

5   Conclusions

This study has extended my previous "asymmetric" idiotype network model [5] composed of an auto-catalytic idiotype by considering the anti-idiotypic idiotype, and has demonstrated the updated asymmetric model can display not only the immune memory but also the immune tolerance on the "transient" dynamics, which my previous model did not. The result has an important implication that we do not have to stick to the idea of the immunological states, such as immune memory and immune tolerance being realized as stable dynamical fixed points. But hitherto idiotype network models adopted "asymmetric" idiotypic interaction has been unaccepted as immunologically relevant models. Because

Hoffmann has proved theoretically that some fixed points the asymmetric models have are inevitable to be unstable [2]. However, if immunological states as represented by the immune memory and the immune tolerance are realized as not a stable fixed point, but rather a long-term transient dynamics, Hoffmann's stability criteria for acceptable models will become unreasonable. So I believe this study must re-stimulate the activity of theoretical researches in terms of asymmetric idiotype network models.

References

1. Jerne, N.K.: Towards a network theory of the immune system, Ann. Immunol. **125C** (1974) 373–389.
2. Hoffmann, G.W.: Regulation of Immune Response Dynamics, DeLisi, C. and Hiernaux, J.R.J (Eds), **1** (1982) 137–162
3. De Boer, R.J.: Theoretical Immunology Parts II, Addison-Wesley Pub.Company Inc, New York (1988) 265–289.
4. De Boer, R.J and Hogeweg, P.: Bull. Math. Biol, **51, 2** (1989) 223–246.
5. Harada, K., Ikegami, T and Shiratori, N.: Proc. of Seventh International Conference on Knowledge-based Intelligent Information Engineering Systems, Springer-Verlag, Berlin (2003) 519–523
6. Elgert, K.D.: Immunology - Understanding the Immune System -. Wiley-Liss, Inc, New York (1996)

Mutual Repairing System
Using Immunity-Based Diagnostic Mobile Agent

Yuji Watanabe[1], Shigeyuki Sato[2], and Yoshiteru Ishida[2]

[1] Graduate School of Natural Sciences, Nagoya City University
Yamanohata, Mizuho-cho, Mizuho-ku, Nagoya, Aichi 467-8501 Japan
yuji@nsc.nagoya-cu.ac.jp
[2] Dept. of Knowledge-based Information Eng., Toyohashi University of Technology
Tempaku, Toyohashi, Aichi 441-8580, Japan
{sato,ishida}@sys.tutkie.tut.ac.jp

Abstract. In our previous research, we proposed a new approach for mutual repairing system using the immunity-based diagnostic mobile agents. However, we have not yet analyzed the performance by more detailed simulations and/or mathematical models. In this paper, the approach using the immunity-based diagnostic mobile agent is compared with other approaches using majority vote model or host-to-host communication. Some results show that the immunity-based diagnostic mobile agents can repair more abnormal hosts than the other methods. We also address a mathematical model to explain a phase transition.

1 Introduction

Fault tolerance has been discussed from a lot of standpoints, such as fault detection, fault avoidance and fault recovery. In recent years, the fault recovery has attracted much attention again, for example *self-repairing computers* [1] and *reliable cellular automata* [2].

We have studied autonomous distributed fault tolerant systems inspired by the biological immune system. The *immunity-based diagnosis model* based on the concept of the *idiotypic network hypothesis* [3] has been proposed by one of the authors in [4]. Similarly to immune cells circulating through the body, *mobile agents* with the immunity-based diagnosis model can monitor the state of host computer, moving around computer network [5]. In our previous research [6], we proposed a new approach for *mutual repair* using the immunity-based diagnostic mobile agents. The repair by agents can be regarded as the neutralization of the antigens toxicity by antibodies. In the approach, host computers and mobile agents mutually test based on the immunity-based diagnosis model, and then some hosts and agents try to replace their own data with data received from others. The effectiveness of the approach was partially confirmed by some simulations. However, we have not yet analyzed the performance by more detailed simulations and/or mathematical models.

In this paper, the approach using the immunity-based diagnostic mobile agent is compared with other approaches using majority vote model or host-to-host

communication. The majority vote is substituted for the immunity-based diagnosis, and host-to-host communication for mobile agent. Some results show that the immunity-based diagnostic mobile agents can repair more abnormal hosts than the other methods. We also address a mathematical model for the majority vote on host-to-host communication to explain a phase transition on the number of abnormal hosts.

2   Simulated Environment [6]

To make it easy to analyze the performance of diagnosis and repair, we employ a simulated computer network environment. The network consists of N host computers which are connected each other randomly. The number of connection per host is $N_c(2 \leq N_c < N)$. The network topology is fixed in simulation execution.

Each host computer has the following four components: *diagnostic module*, *repairing module*, *data*, and *buffer for repair*. The diagnostic and repairing modules are detailed in next section. For a simple data structure, we use a binary string, where each bit is assigned either TRUE or FALSE. The length of string is defined by L. We suppose that all bits of normal host are TRUE, while the bits of abnormal host are partially corrupted to FALSE at the rate e. The average number of FALSE bits in the abnormal host is equal to eL. The abnormal host can discover the FALSE bits, not by itself, but only by comparisons with other normal units. Note that the target of diagnosis and repair is only the data, not including diagnostic and repairing modules themselves. At the start of each simulation, $N_f(0)(\leq N)$ host computers are already abnormal.

3   Diagnosis and Repair Method

3.1   Two Approaches

In our previous research[6], we proposed an approach for mutual diagnosis and repair using mobile agents as illustrated in Fig. 1(a). At the beginning of the simulation, each host produces several mobile agents having a part of the host's data, diagnostic and repairing module, and buffer. Normal hosts therefore create fault-free mobile agents, while abnormal hosts create faulty mobile agents having some FALSE bits. Note that not only abnormal hosts but also faulty mobile agents are targets for repair. During the simulation, mobile agents migrate from host to host and diagnose each host. There are also mutual tests among agents acting on the same host. The total number of mobile agents and the length of agent's data string are represented by M and $L_a(\leq L)$, respectively.

An alternative approach of mobile agent is that host computer directly diagnoses and repairs the adjacent hosts as shown in Fig. 1(b). The approach of host-to-host communication does not employ mobile agents. The difference between the two approaches is whether the object which can test host computer is fixed or variable. In host-to-host approach, each host is tested by the same adjacent hosts repeatedly for the fixed network topology. In mobile agent approach, each host is checked by various agents.

(a) mobile agent   (b) host-to-host communication
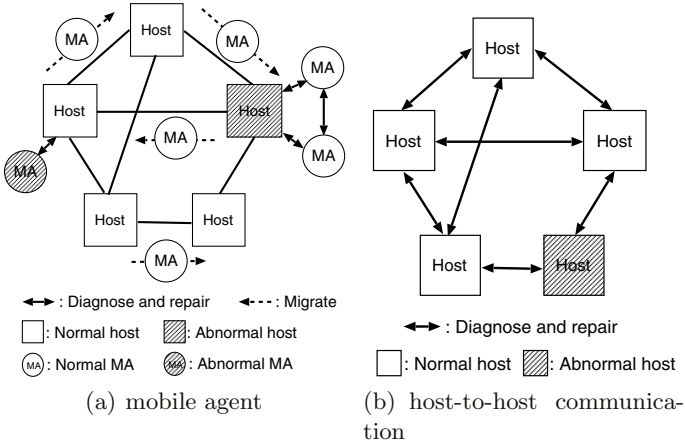
Fig. 1. Two approaches for mutual diagnosis and repair.

3.2 Diagnosis and Repair Process

We detail the diagnosis and repair process using the immunity-based diagnosis model proposed in [4]. First of all, a state variable R_i indicating the *credibility of unit* is assigned to each unit. Unit represents host computer or mobile agent. For one step, the following process can be executed:

1. Unit i sends a randomly-selected part of its data to unit j. The length of transferred data string is fixed by $L_t(\leq L)$. In Fig. 2, $L_t = 1$ and unit i sends the third bit to unit j.
2. The diagnostic module in unit j decides the test outcome T_{ji} by comparing the received data with its own data:

$$T_{ji} = \begin{cases} 1 & \text{if all corresponding bits are equal} \\ -1 & \text{if one or more mismatches exist} \\ 0 & \text{if unit } j \text{ does not test unit } i \end{cases} . \tag{1}$$

   In Fig. 2, $T_{ji} = 1$ because the third bit is the same.
3. Unit j returns both the test outcome T_{ji} and its own credibility R_j.
4. The diagnostic module in unit i receives some data from unit j and then outputs T_{ij} and R_i. If $T_{ij} = -1$ and $R_i < 0.5$, the repairing module in unit i stores the data received from unit j in the buffer for repair. In Fig. 2, the second bit from unit j is stored in the buffer.
5. After unit i is also tested by the other units, the diagnostic module updates the credibility R_i based on the immunity-based diagnosis model as follows:

$$\frac{dr_i(t)}{dt} = \sum_j T_{ji} R_j + \sum_j T_{ij} R_j - \frac{1}{2} \sum_{j \in \{k: T_{ik} \neq 0\}} (T_{ij} + 1), \tag{2}$$

$$R_i(t) = \frac{1}{1 + \exp(-r_i(t))}, \tag{3}$$
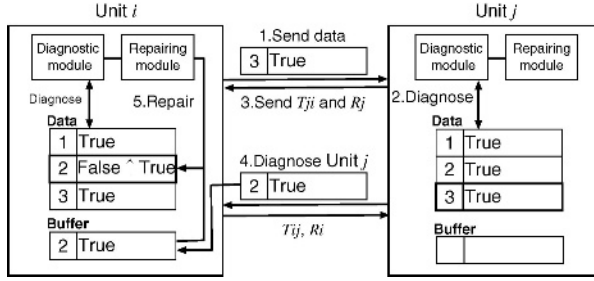
Fig. 2. Diagnosis and repair process.

where the credibility $R_i \in [0,1]$ is a normalization of $r_i \in (-\infty, \infty)$ using a sigmoid function. In addition, if the buffer has some data, the repairing module rewrites the data. In Fig. 2, unit i successfully recovers TRUE at the second bit.

A simpler substitution for the immunity-based diagnosis model is majority vote model. The following parts of the majority model are mainly different from the above-mentioned process in terms of not using the credibility R_i:

4'. Unit i receives some data from unit j and then outputs T_{ij}. If $T_{ij} = -1$, the received data is stored in the buffer for repair.
5'. After unit i is also tested by the other units, if the buffer has some data and $\sum_j T_{ij} < 0$, the repairing module rewrites the data.

4   Simulation Results

In each simulation, the parameters listed in Table 1 are fixed or variable. For the same communication costs at every step by the two approaches (mobile agent and host-to-host communication), the length of agent's data L_a is equal to the length of transferred data L_t. Additionally, the number of mobile agents per host is equivalent to the number of connection per host N_c on host-to-host communication. The reason is that each host is tested by M/N mobile agents, while it is checked by N_c adjacent hosts in the host-to-host approach.

Figure 3 illustrates the number of abnormal hosts after 100 steps $N_f(100)$ vs. the initial number of abnormal hosts $N_f(0)$ for the following methods:

- the immunity-based diagnosis using mobile agents (ID MA)
- the majority vote using mobile agents (MV MA)
- the immunity-based diagnosis on host-to-host communication (ID H2H)
- the majority vote on host-to-host communication (MV H2H)

Each result takes an average over 100 runs. If the repairing module is not installed, the number of abnormal hosts never changes. We make sure whether $N_f(100)$ is below $N_f(0)$.

Table 1. List of parameters.

Parameter	Description	Value
$R_i(0)$	Initial value of credibility	1.0
$r_i(0)$	Initial value of intermediate variable	5.0
N	Number of hosts	50
$N_f(0)$	Initial number of abnormal hosts	variable
e	Data corruption rate	1
L	Length of host's data string	variable
L_t	Length of transferred data	1
L_a	= Length of mobile agent's data	
N_c	Number of connection per host	25
M/N	= Number of mobile agents per host	

(a) L=1   (b) L=8
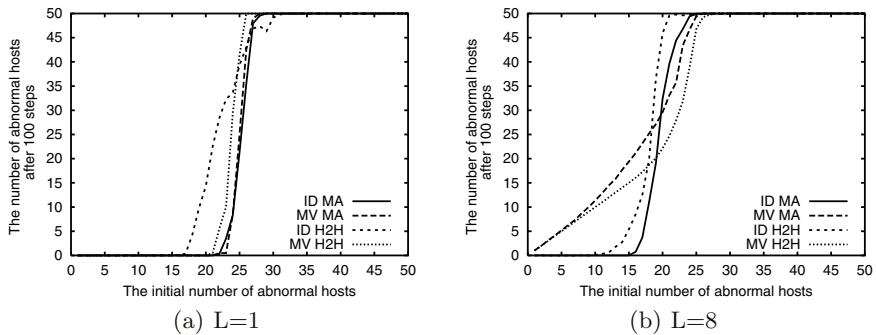
Fig. 3. The number of abnormal hosts after 100 steps $N_f(100)$ vs. the initial number of abnormal hosts $N_f(0)$ for four methods, also changing the length of host's data string L.

When the length of host's data string L is 1 in Fig. 3 (a), although ID H2H is slightly different, the other methods can repair all abnormal hosts until $N_f(0)$ becomes 20. In other words, a phase transition can be observed. A mathematical model for MV H2H when $L = 1$ will be discussed in the next section to explain the phase transition.

In Fig. 3 (b), both MV MA and MV H2H using the majority vote are of no effect for $L = 8$. Unlike the majority vote, the immunity-based diagnosis model can repair some abnormal hosts even for $L = 8$. In addition, the immunity-based diagnosis approach using mobile agent (ID MA) is better than the host-to-host approach (ID H2H).

5   Mathematical Model

We attempt to construct a mathematical model for the simple method, namely, the mutual repair using the majority vote on host-to-host communication. If $L = 1$, abnormal host has only one FALSE bit and always sends FALSE to

connected hosts. However, when $L > 1$, abnormal host sometimes sends TRUE because the bits of abnormal host are partially corrupted to FALSE.

Assuming that $L = 1$, if normal host i with a TRUE bit tests adjacent abnormal host j with a FALSE bit, then $T_{ij} = -1$ according to Eq. (1), and vice versa $(T_{ji} = -1)$. When all neighboring hosts of normal host i are abnormal, $\sum_j T_{ij} = -N_c (< 0)$. The probability of such case is represented by $p(t)^{N_c}$ where $p(t) = N_f(t)/N$, the probability of being abnormal host at step t.

In the majority vote, normal host i replace TRUE with FALSE by mistake when $\sum_j T_{ij} < 0$, that is, more than half of the neighboring hosts are abnormal. The probability that $\sum_j T_{ij} < 0$ is calculated by the sum of the binomial distribution under the conditions that $N_c = 2n + 1$ and $N \gg 1$:

$$P(t) = p^{N_c} + \binom{N_c}{1} p^{N_c - 1}(1 - p) + \cdots + \binom{N_c}{\frac{N_c-1}{2}} p^{\frac{N_c+1}{2}}(1 - p)^{\frac{N_c-1}{2}}, \quad (4)$$

where $p \equiv p(t) = N_f(t)/N$. In case of even-numbered N_c, the last term of the right-hand side of the above equation is slightly different. Strictly speaking, we should substitute the hypergeometric distribution for the binomial distribution when N is small.

If some normal hosts mistakenly change TRUE to FALSE, the number of abnormal hosts $N_f(t)$ increases. On the other hand, some abnormal hosts successfully replace FALSE with TRUE when more than half of the neighboring hosts are normal. Note that $T_{ij} = -1$ when abnormal host i tests adjacent normal host j. The probability that more than half of the neighboring hosts are normal is $1 - P(t)$. In sum, the number of abnormal hosts at step $t + 1$ is given as:

$$N_f(t+1) = N_f(t) + P(t)(N - N_f(t)) - (1 - P(t))N_f(t) = P(t)N, \quad (5)$$

where $N - N_f(t)$ is the number of normal hosts at step t. For example, when $N_c = 3$, the above equation is expanded by Eq. (4) as follows:

$$N_f(t+1) = \frac{N_f^2(t)}{N^2}(3N - 2N_f(t)). \quad (6)$$

Figure 4 plots the number of abnormal hosts at step t for $N = 50$. From the graph, if $N_f(0) < 25$, $N_f(1) < N_f(0)$ and finally $\lim_{t \to \infty} N_f(t) = 0$. Similarly, $\lim_{t \to \infty} N_f(t) = N$ when $N_f(0) > 25$. There is a fixed point for $N_f(t) = 25$. The phase transition agrees with the graph when $L = 1$ as observed in Fig. 3 (a). Figure 4 (b) also suggests that the curve becomes steep as N_c increases. Mathematical models for $L > 1$ and the other methods are under study.

6   Conclusions and Further Work

In this paper, the immunity-based diagnostic mobile agents can repair more abnormal hosts than other methods using majority vote model or host-to-host communication. The phase transition by the mathematical model agrees with the simulation result. In further work, we will go on analyzing the performance of mutual repair by both simulations and mathematical models.

(a) $N_c=3$ (b) $N_c=5$
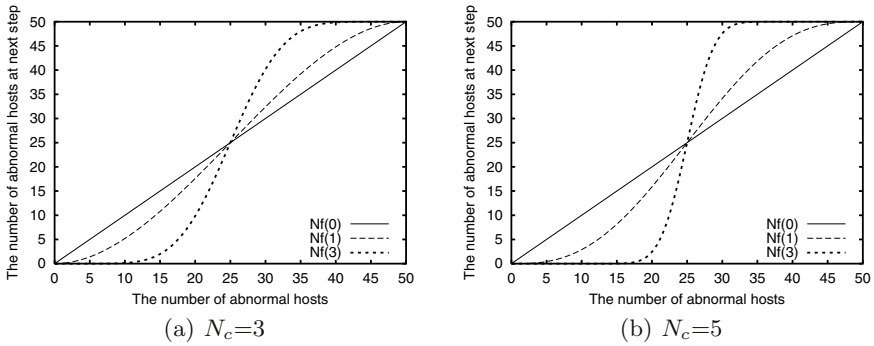
Fig. 4. The number of abnormal hosts at step t by the mathematical model for $N = 50$, also changing the number of connection per host N_c.

Acknowledgements

References

1. A. Fox and D. Patterson: Self-repairing computers. Scientific American **288-6** (2003) 54–61
2. P. Gacs: Reliable Cellular Automata with Self-Organization. Journal of Statistical Physics **103** (2001) 45–267
3. N. Jerne: The immune system. Scientific American **229-1** (1973) 52–60
4. Y. Ishida: Fully distributed diagnosis by PDP learning algorithm: towards immune network PDP model. Proc. of IJCNN (1990) 777–782
5. Y. Watanabe and Y. Ishida: Immunity-based approaches for self-monitoring in distributed intrusion detection system. KES 2003 (LNAI 2774) (2003) 503–510
6. Y. Watanabe, S. Sato and Y. Ishida: An approach for self-repair in distributed system using immunity-based diagnostic mobile agents. KES 2004 (LNAI 3214) (2004) 504–510

A Network Self-repair by Spatial Strategies in Spatial Prisoner's Dilemma

Yoshiteru Ishida and Toshikatsu Mori

Department of Knowledge-Based Information Engineering
Toyohashi University of Technology
Tempaku, Toyohashi, 441-8580 Japan

Abstract. We deal with a problem of cleaning up a contaminated network by mutual copying. This problem involves not only an aspect of "the double-edged sword" where copying could further spread contamination but an aspect of mutual cooperation where resource consuming copying could be left for others. The framework of "prisoner's dilemma" has been applied, aiming at emergence of appropriate copying strategies in an adaptive manner to the network environment.

1 Introduction

Many artificial networks such as the Internet have been revealed to be a *scale-free* network [1]. This finding has been making an impact on many researches such as on the security of the network [2]. On the other hand, self-repair systems including those inspired from the immune system have been studied extensively [3]. Adaptive nature of the immune system to the environment is of interest, which is realized by cooperative working of inhomogeneous agents with distinct strategies. Further, a new idea of recovery-oriented computing has been recently proposed [4]. We have studied the problem of cleaning up the network by mutually copying, and proposed a self-repair model [5] that can be equated with probabilistic cellular automata [6]. The model considers only synchronous and non-selective copying. This paper, however, focuses on an adaptive copying using a Prisoner's Dilemma approach [7-9].

In case of self-repair systems with autonomous distribution, abnormal units may adversely affect the system when they try to repair other normal units. Additionally, because repairing uses some resources, frequent repairs reduce the performance of the system. Although the frequency of repairs has to be decided giving consideration to the system environment, it cannot be easily determined because the environment within a system changes with time.

This paper proposes a model where the units select a strategy depending on the system environment. In this model, a unit is assigned some resources and a strategy; and it repairs according to this strategy. We aim at raising the reliability of the system without reducing the performance, in a fully distributed framework.

When the self-repair is done in an autonomous distributed manner, each unit does not voluntarily repair other units to save their own resources, thus leaving many faulty units not repaired. This situation is similar to the dilemma that would occur in

the Prisoner's Dilemma. Thus, we use an approach of a spatial version of Prisoner's Dilemma [10-14] for emergence of cooperative collectives and for controlling copying to save resources.

2 Spatial Strategies in SPD

We have studied a spatial version of Prisoner's Dilemma [15], and proposed a generalized TFT (Tit-for-Tat) such as *k1C, k2D* and their combination *k1C-k2D*, where *k1* is a parameter indicating generosity and *k2* contrariness. Dynamics of these spatial strategies in a two-dimensional lattice has been also studied in a noisy environment.

The PD is a game played just once by two players with two actions (cooperation, C, or defect, D). Each player receives a payoff (*R, T, S, P*) where $T>R>P>S$ and $2R>T+S$.

In IPD, each player (and hence the strategy) is evaluated. In SPD, each site in a two-dimensional lattice corresponds to a player. Each player plays PD with the neighbors, and changes its action by the total score it received.

Our model generalized SPD by introducing spatial strategy. Each player placed at each lattice of the two-dimensional lattice. Each player has an action and a strategy, and receives a score. Spatial strategy determines the next action dependent upon the spatial pattern of actions in the neighbors. Each player plays PD with the neighbors, and changes its strategy to the strategy that earns the highest total score among the neighbors. Table 1 is the Payoff matrix of PD. In our simulations, *R, S, T*, and *P* are respectively set to *1, 0, b (1 < b < 2)* and 0 in simulations below following the Nowak-May's simulations [12].

Table 1. The Payoff Matrix of the Prisoner's Dilemma Game. *R, S, T, P* are payoff to the player 1. (*1 < b < 2*).

		Other	
		C	D
Player	C	R (1)	S (0)
	D	T (b)	P (0)

Our SPD is done with spatial strategies: the next action will be determined based on the pattern of neighbors' actions. Score is calculated by summing up all the scores received from PD with 8 neighbor players. After *r* (strategy update cycle) steps of interactions with neighbors, the strategy will be chosen from the strategy with the highest score among the neighbors. Thus, the strategy will be updated at every *r* steps. In an evolutionary framework, strategy will be also changed by a mutation rate where mutation is operated on the string of the strategy code below.

To specify a spatial strategy, actions of the eight neighbors (the *Moore* neighbor) and the player itself must be specified (Fig. 1), hence 2^9 rules are required. For simplicity, we restrict ourselves on a `` totalistic spatial strategy'' that depend on the number of D (defect) action of the neighbor, not on their positions.

To represent a strategy, a bit sequence is used [10] whose *l*-th element is C (D) if the action C (D) is taken when the number of D of the neighbor players is *l* (*l*=0,1, …,8). For example, *All-C* is [CCCCCCCCC], *All-D* is [DDDDDDDDD]. The number of strategies is 2^9 =512. As a typical strategy, we define *kD* that takes D if *l>k* and C otherwise. For example, *2D* is [CCDDDDDDD]. This *kD* can be regarded as a spatial version of TFT where *k* seems to indicate the generosity (how many D actions in the neighbor are tolerated.).

The strategy *kC* can be similarly defined: it takes C if if *l>k* and D otherwise. For example, *4C* is [DDDDCCCCC] as shown in Fig. 1. The number *k* in the strategy *kC* seems to indicate the contrariness that would cooperate even if *k* players in the neighbor defect. Further, *k1D-k2C* is a combination of *k1D* and *k2C*. For example, *2D-7C* is [CCDDDDDCC]. In our previous studies, the action error has an effect to favor more generous spatial strategies in the sense that more D actions should be forgiven in the neighbor's actions [15].
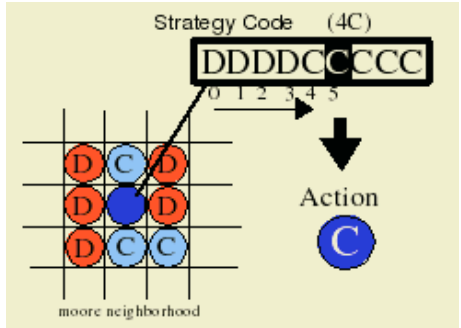
Fig. 1. A Strategy Code for Spatial Strategies.

3   A Model Incorporating SPD

The units in the system are assigned tasks that they should do. The units are assumed to use corresponding resources while repairing, making it difficult for them to perform much of the tasks. The units have to do the tasks assigned to them; but without doing repair, abnormal units increase and hence the performance in a system level decreases; hence a dilemma. Our model adopts the spatial version of the Prisoner's Dilemma (SPD) approach. The SPD model arranges players that interact with neighboring players in a two-dimensional lattice space. Each unit is assumed to use *spatial strategies* [15].

The unit becomes abnormal depending on the failure rate λ . Each unit has a Strategy Code and repairs according to that code. Repairing is done depending on the repairing rate γ and a repaired unit becomes normal. Also the unit uses the quantity

of repair resource R_γ for repairing. The unit is able to repair more than one unit, provided that the quantity of maximum resource `Rmax` is not exceeded. We call the resource that is not used for repairing the available resource `Ra`, and consider it as *score* of a unit. If an abnormal unit repairs another unit, the failure rate λ of the repaired unit is increased by damage rate δ . Through the strategy update cycle r, the strategy of unit is updated to that of the strategy that got the highest profit among the *Moore* neighborhood units. Each unit repairs(C) or does not repair (D). The strategy code of the unit is expressed as a string character of 9 bits consisting C's and D's.

Fig. 1 shows an example of a strategy code. The unit counts the number of D units in its *Moore* neighborhood in the previous step. In this case, five D units are present; then in the next step, the unit does the action (C) represented in the fifth bit.

When the unit copies its content (software that can be copied and be contaminated), the strategy of the unit is also copied. Thus, the strategy will be changed at copying in addition to every strategy update cycle (and at mutation in a evolutionary framework).

This paper restricts strategies to the *kC* strategy, which is composed of only 9 among the 512 existing strategies. In the *kC* strategy, the unit does the action of D if the number of D units is less than k; and it does the action of C if the number of D units is greater or equal to *k*.

4   Computer Simulations with a Square Lattice

The simulation is done in a two-dimensional lattice space with a unit existing in each lattice. Throughout simulations stated in this section, the parameters are: failure rate 0.001, repair rate 0.01, quantity of maximum resource 8, quantity of repair resource 2, damage rate 0.1 and strategy update cycle 200.

Fig. 2 plotted the number of failure units (average number of 10 independent trials) with damage rate varying from 0.00 to 0.30. In this simulation, all the units are set to be normal initially. It can be observed that there is a threshold between 0.12 and 0.13 over which the number of failure units explode. The threshold is also observed in a somewhat different but similar (and simpler) model for self-repair [5].

To examine the efficacy of the strategic repair, we compare it with two trivial strategies: the one where the units always repair; and the one where the units never repair. We used the same parameters for these three strategies with the damage rate fixed at 0.1. To observe the effect of repair, 50% of units out of the units comprising the 50×50 space are randomly selected and set to be abnormal in the initial configuration.

Fig. 3 shows the available resource (averaged over 40 independent trials) in each model. We can observe that both *the always repair* and *the strategic repair* become stable after some number of steps (An oscillation observed in the strategic repair is due to the change of numbers in strategies at the strategy update cycle.). However, more available resource remains in *the strategic repair*. Moreover, *the strategic repair* becomes stable twice as fast.
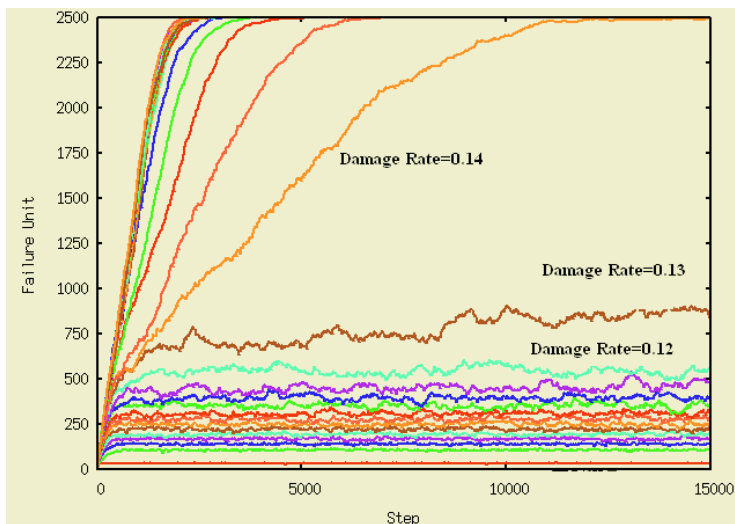
Fig. 2. A Threshold in Damage Rate.
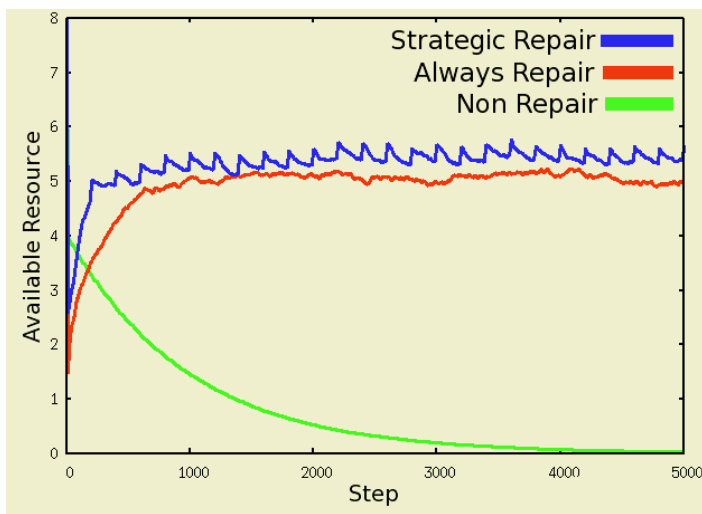
Fig. 3. Time Evolution of Available Resources.

Fig. 4 shows the number of units in each strategy of *the strategic repair*. We can see that when the number of abnormal units increases, units with easily repairing strategies like *0C* also increase. Otherwise, units using hardly repairing strategies like *8C* increase. Therefore, the units choose an appropriate strategy depending on the environment.
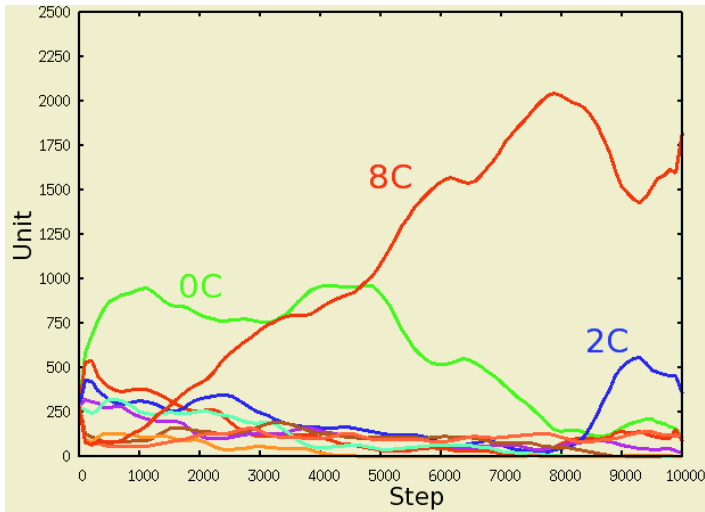
Fig. 4. Time Evolution of the Fraction of Units with Spatial Strategies.

5   Conclusion

We have studied the problem: when the network can be cleaned up by mutual copy-ing. We proposed a probabilistic cellular automaton model for synchronous copying for simplicity [5]. In this paper, a game theoretical approach of spatial Prisoner's Dilemma has been used. With the approach, mutual copying can be controlled so that units will mutually repair by copying but not too excessively to exhaust their re-sources. It is also shown that a strategy appropriate for the environment will emerge when the system environment changes.

Acknowledgements

References

1. Barabasi, A.-L.: Linked: The New Science of Networks , Perseus, (2002)
2. Dezso, Z. and Barabasi, A.-L.: Halting viruses in scale-free networks, Phys. Rev. E 65, 055103, (2002)
3. Ishida, Y.: Immunity-Based Systems: A Design Perspective, Springer, Berlin Heidelberg New York   (2004)
4. Brown, A. and Patterson, D.: Embracing Failure: A Case for Recovery-Oriented Comput-ing   (ROC), High Performance Transaction Systems Workshop (TTPS '01) (2001)

5. Ishida, Y.: A Critical Phenomenon in a Self-Repair Network by Mutual Copying, LNAI, (2005) this volume

6. Domany, E. and Kinzel, W.: Equivalence of cellular automata to Ising models and directed percolation, Phys. Rev. Lett. 53   (1984) pp. 311

7. Axelrod, R.: The Evolution of Cooperation, Basic Books, New York (1984)

8. Axelrod, R.: The Evolution of Strategies in the Iterated Prisoner's Dilemma, in Davis, L (ed.): Genetic Algorithms and Simulating Annealing, Pitman (1987) pp.31-41

9. Axelrod, R. and Dion, D.: The Further Evolution of Cooperation, Science, vol.242, pp.1385-1390

10. Matuo, K. and Adachi, N. : Metastable Antagonistic Equilibrium and Stable Cooperative Equilibrium in Distributed Prisoner's Dilemma Game, Proc. Int. Symp. Syst. Res., Infor. Cybern. (1989)

11. Boyd, R.: Mistakes Allow Evolutionary Stability in the Repeated Prisoner's Dilemma Game, J. theor. Biol., vol. 136 (1989) 47-56

12. Nowak, M.A. and May, R.M.: Evolutionary games and spatial chaos, Nature, Vol. 359 (1992) pp.826-829

13. Grim, P.: The greater generosity of the spatialized prisoner's dilemma, J. theor. Biol., Vol. 173, (1995) pp. 353-359

14. Nakamaru, M., Nogami, H. and Iwasa, Y.: Score-dependent fertility model for the evolution of cooperation in a lattice, J. theor. Biol., Vol. 194 (1998) pp. 101-124

15. Ishida, Y. and Mori, T.: Spatial Strategies on a Generalized Spatial Prisoner's Dilemma, J. of Artificial Life and Robotics, (2005) to appear

A Critical Phenomenon in a Self-repair Network by Mutual Copying

Yoshiteru Ishida

Department of Knowledge-Based Information Engineering
Toyohashi University of Technology
Tempaku, Toyohashi, 441-8580 Japan

Abstract. This paper reports a critical phenomenon in a self-repair network by mutual copying. Extensive studies have been done on critical phenomena in many fields such as in epidemic theory and in percolation theory with an effort of identification of critical points. However, from the viewpoints of cleaning up a network by mutual copying, critical phenomena have not much studied. A critical phenomenon has been observed in a self-repair network. Self-repairing by mutual copying is "the double-edged sword" that could cause outbreaks with inappropriate parameters, and careful investigations are needed.

1 Introduction

The internet has been pointed out to be a scale-free network [1], and it is further suggested that the network is tolerant to a random attacks but vulnerable to selective attacks [2]; hence the selective defense at the hub seems to be effective. Further, it is also pointed out that the critical point (over which the epidemic breaks out) in scale-free networks is 0, thus it is difficult to eradicate computer viruses and worms from the Internet [2].

Against the selective attacks and computer virus spread with 0 critical point, the selective or even adaptive defense [3] can be considered. For the physical systems such as mechanical systems, they are repaired by identifying the faulty components and replacing them with non-faulty ones. For information systems, however, they can be repaired by simply copying the clean system to the contaminated system. As the first step toward the adaptive defense of the information systems, we consider the self-repairing of the network by mutual copying.

2 Self-repair by Copying: The Double Edged Sword

In biological epidemic spreads [4-6], spontaneous recoveries could occur and even the immune system could give arise. However, in information systems such as computer networks, human must repair or computers must repair mutually when contaminated with computer viruses and worms. This paper deals with the latter case where the computers mutually repair. In this case, computers could contaminate other ones when the repairing computers themselves are contaminated. This is because com-

puters can repair software faults by copying, quite differently from mechanical systems. In hardware faults of mechanical systems, the repairing units simply cannot repair others when the repairing units are faulty. However, in software faults of information systems, the repairing units can repair others simply by copying their content. Further, the self-repair by copying could have spread contamination when the network has already highly contaminated or the infection rate is quite high. This paper concentrates on the naive problem of cleaning up the network by mutually copying: **When can a contaminated network be cleaned up by mutually copying?**

We have studied the immunity-based systems, and pointed out the possibility that they will be the double edged sword [7]. For example, the system based on the Jerne's idiotypic network framework has recognizing units that are also being recognized. Self-nonself recognition involving self-referential paradox could also lead to the situation of the double edged sword: recognition done by units credible enough is credible, but not credible otherwise.

The repair by copying in information systems is also the "double edged sword" and it is engineering concern to identify when it can really eradicate abnormal elements from the system. This paper considers CA (cellular automata), probabilistic CA specifically, to model the situation where computers in a LAN mutually repair by copying their content. We will give clear priorities on simplicity over reality of the system.

The studying the reliability of information systems by a simple model of a probabilistic CA is not new. The results by Peter Gacs [13] for his probabilistic CA amount to; "*there exists a simple finite-state machine M such that for any computer program π and any desired level of reliability p<1, the program π can be run with reliability p on a sufficiently large one dimensional array of copies of M, each one communicating only with its nearest neighbors, even if these machines all have the same small but positive error rate*" [14], when interpreted to the computer networks.

3   Models of Probabilistic Cellular Automata

In our model, the system consists of units capable of repairing other units connected. We call the connected units as neighbor units based on the terminology of CA. The repairing may be done by copying its content to the other units, since the application to a computer network by LAN in our mind.

Although mutual repairing and testing may be done in an asynchronous manner, our model considers synchronous interactions for simplicity and for comparison with existing probabilistic CA models [8-10]. Each unit tries to repair its adjacent units, however, since the repairing is done by copying it could make the adjacent units abnormal rather than normal.

In a mathematical formulation, the model consists of three elements (*U, T, R*) where *U* is a set of units, *T* is a topology connecting the units, and *R* is a set of rules of the interaction among units. In this paper, a set of units is a finite set with *N* units, and the topology is restricted to the one-dimensional array as shown in Fig. 1 (which could be n-dimensional array, complete graph, random graph, or even scale-free network) that could have *S* neighbors for each units with a boundary condition, i.e. the

structure of the array is a ring with unit 1 adjacent to the unit *N*. Also, we restrict the case where each unit has a binary state: normal (0) and abnormal (1).

Each unit tries to repair the adjacent units in a synchronous fashion with a probability **Pr**. As shown in Fig. 2, the repairing will be successful with the probability **Prn** when it is done by a normal unit, but with the probability **Pra** when by an abnormal unit (**Pra < Prn**). In this paper, we assume **Prn** =1. The repaired units will be normal when all the repairing is successful. Thus, when repairing is done by the two adjacent units, all these two repairing must be successful in order for the repaired unit to be normal.

Fig. 1. One-dimensional array with two states: normal (0) and abnormal (1).
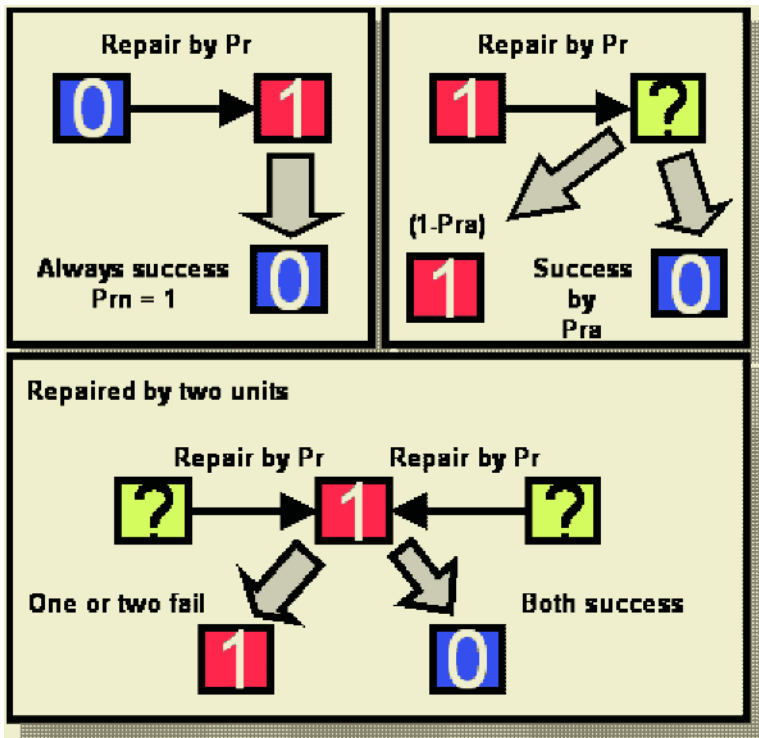
Fig. 2. Probabilistic repair by normal units (*above left*) and by abnormal units (*above right*). Repair by two units (*below*).

As a probabilistic cellular automaton, the transition rules are as follows (where the self-state is the center in the parenthesis and the two neighbor states are at the left and the right in the parenthesis. The self-state will be changed to the state indicated to the right of the arrow, with the probability indicated after the colon):

(000)→0: 1, **(010)→1:** $(1-\mathrm{Pr})^2$,

(001)→1: α, **(011)→1:** α+ $(1-\mathrm{Pr})^2$,

(101)→1: 2 $(1-\mathrm{Pr})$ α+β, **(111)→1:** 2 $(1-\mathrm{Pr})$ α+β+ $(1-\mathrm{Pr})^2$,

where α= Pr $(1-\mathrm{Pr}\,a)$, β= Pr^2 $(1-\mathrm{Pr}\,a^2)$.

In such models, it is of interest to determine how the repairing probability **Pr** should be set when the success probability by abnormal unit **Pra** is given. Also, when **Pr** is fixed to some value and **Pra** moves continuously to large value, does the number of abnormal units change abruptly at some critical points or does it just gradually increase? Further, when given some value of **Pra**, **Pr** should always be larger, which requires more cost.

4   Relation with Domany-Kinzel Model and Other PCA

Domany-Kinzel model [8] is a one-dimensional two state and totalistic probabilistic cellular automaton (PCA) with the interaction timing is specific. The interaction is done in an alternated synchronous fashion: the origin cell with state 1 is numbered as 0. The numbering proceeds {1,2,.... } to the right, and {-1,-2,.... } to the left. At N-th step the even numbered cells will act to the odd numbered cells and the odd numbered cells will act at the next step. The neighbor is two cells adjacent to oneself without self-interaction.

The interaction rule is as follows (Fig. 3):

(0*0)→0:1, **(0*1)→1:**p1, **(1*1)→1:**p2
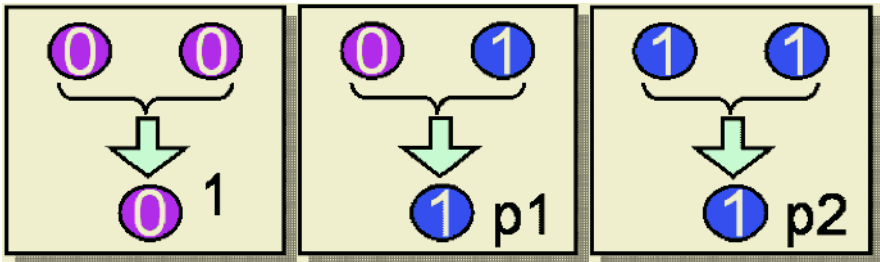
Fig. 3. Transition of the Domany-Kinzel model when both side is 0 (*left*) , one side is 1 (*middle*), and both side is 1 (*right*).

When the number of units is finite, **0** (all the units are normal) is only one steady state and all the initial states will be exponentially [11] attracted to **0** when *p1*<1, *p2*<1. When the number of units is infinite, the system has other steady states with 1 and 0 mixed with a certain ratio, other than all 0 states. Our PCA model can be equated with the DK model when Pr=1 (i.e. units always repair) with the parameters: *p1*=α (=$(1-\mathrm{Pr}\,a)$), *p2*=β (=$(1-\mathrm{Pr}\,a^2)$); i.e. the case of the directed bond percolation. For two-dimensional cases, some results of p-Voter model [12] are available.

5   Simulation Results

Computer simulations are conducted for one-dimensional CA with boundary condition: that is, a ring-shaped network. Initially, only one unit is abnormal and the unit is numbered as 0. The number of units is 500. Initially only one unit is abnormal (1). One run stops at 800 steps. Fig. 4 shows the number of normal units (averaged over 10 times).
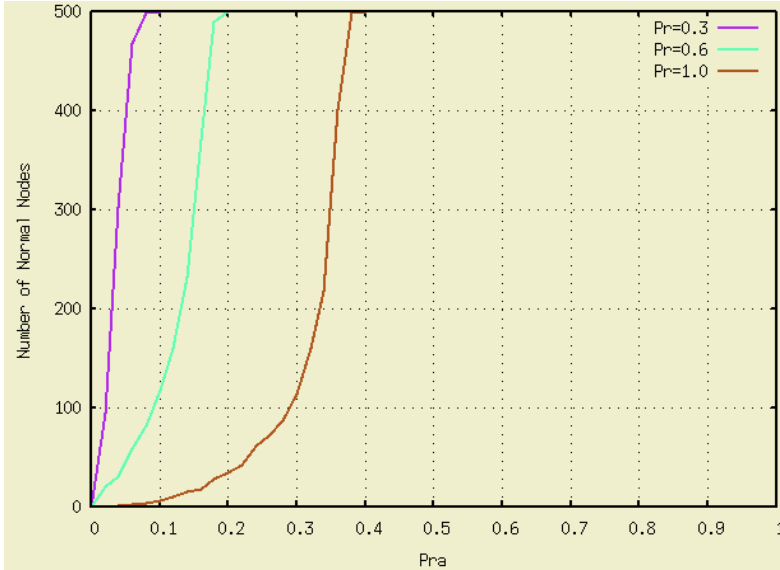
Fig. 4. The number of normal units after 800 steps plotted when the successful repair probability **Pra** varies. Three lines correspond to the **Pr**:,0.3,0.6,1.0 from the left to the right.

Under the approximation that the probability that the state of unit 0 is constant **p0** (mean field approximation and steady state), the following steady state probability of **p0** is obtained.

$$\frac{\Pr a(2 - 2\Pr + \Pr a \Pr)}{\Pr(1 - 2\Pr a + \Pr a^2)}$$

This steady state probability also matches qualitatively with the above simulation results.

When **Pra** increases, the number of normal units rapidly increases. The steeper the curves, the smaller the probability **Pr**. That is, when **Pra** is less than 0.4, **Pr** should be small and hence repair should be done infrequent.

Further, there is a critical point of **Pra** that must be exceeded to clean up the network. The critical point becomes larger, as the probability **Pr** becomes larger. In Fig. 5, the critical points are plotted. The left region is the *frozen* region where all the
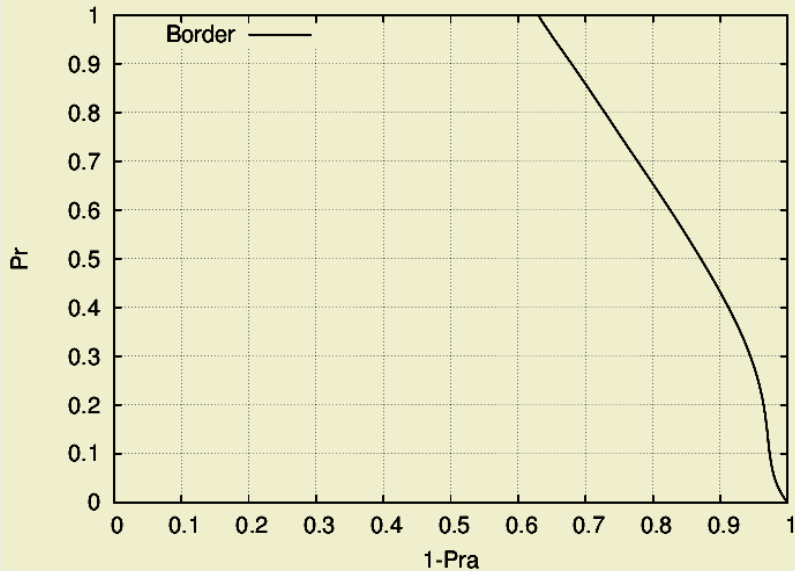
Fig. 5. Frozen Phase (*left region* where all the units are normal) and Active Phase (*right region* where some units remain to be abnormal).

units become normal, and the right region is the *active* region where some abnormal units remain in the network.

6 Conclusion

As a first step toward an adaptive defense of the network, we studied a simple problem of cleaning up the network by mutually copying. In a probabilistic framework, a model similar to the Domany-Kinzel model can be obtained. A critical phenomenon is observed: when copying from faulty units has a certain rate of infecting rather than cleaning, then mutual copying could spread contamination explosively; hence the rate should be carefully identified and controlled.

Acknowledgements

References

1. Barabasi, A.-L.: Linked: The New Science of Networks , Perseus, (2002)
2. Dezso, Z. and Barabasi, A.-L.: Halting viruses in scale-free networks, Phys. Rev. E 65, 055103, (2002)

3. Ishida, Y. and Mori, T.: A Network Self-Repair by Spatial Strategies in Spatial Prisoner's Dilemma (submitted)
4. Boccara, N. and Cheong, K.: Critical behaviour of a probabilistic automata network SIS model for the spread of an infectious disease in a population of moving individuals, J. of Physics A: Math. Gen. 26 (1993) pp. 3707--3717
5. Rhodes, C.J. and Anderson, R.M.: Dynamics in a Lattice Epidemic Model, Phys. Rev. Lett. A, 210 (1996) pp. 183--188
6. Rhodes, C.J. and Anderson, R.M.: Forest-fires as a Model for the Dynamics of Disease Epidemics, J. of Franklin Inst. 335B(2) (1996) pp. 199-211
7. Ishida, Y.: Immunity-Based Systems: A Design Perspective, Springer, Berlin Heidelberg New York   (2004)
8. Domany, E. and Kinzel, W.: Equivalence of cellular automata to Ising models and directed percolation Phys. Rev. Lett. 53   (1984) pp. 311
9. Bagnoli, F., Boccara N. and Palmerini, P.: Phase Transitions in a Probabilistic Cellular Automaton with Two Absorbing States, arXiv:cond-mat/9705171 v2 (1997)
10. Vichnicac, G.Y., Tamayo P. and Hartman, H.: Annealed and quenched inhomogeneous cellular automata, J. Statist. Phys. 45 (1986) pp. 875.
11. Kinzel, W.: Phase transition of cellular automata, Z. Phys. B 58 (1985) pp. 229
12. Eloranta, K.: Contour dynamics in cellular automata, Physica D 89 (1995) pp. 184
13. Gacs, P.:Reliable Cellular Automata with Self-Organization, J. Stat. Phys. 103(2001), pp. 45-267
14. Gray, L.F.: A Reader's Guide to Gac's "Positive Rates" Paper, J. Stat. Phys. 103(2001), pp. 1-44

A Worm Filter Based on the Number
of Unacknowledged Requests

Takeshi Okamoto

Department of Network Engineering, Kanagawa Institute of Technology
1030 Shimo-ogino, Atsugi, Kanagawa 243-0292, Japan
`take4@nw.kanagawa-it.ac.jp`

Abstract. We propose a new filter for preventing computer worms from spreading. The new worm filter limits the number of unacknowledged requests, rather than the rate of connections to new computers. Normal network traffic is analyzed to determine appropriate parameters for the worm filter. Performance evaluation showed that the worm filter stops not only high-speed worms in the wild, but also simulated slow-speed worms. Finally, the weaknesses of the worm filter is discussed.

1 Introduction

Anti-virus systems protect computers and networks from malicious programs, such as computer viruses and worms, by discriminating malicious programs from harmless programs, and removing only the former. Anti-virus systems can therefore be considered as the computers' immune system.

An innovative method, a "*virus throttle*," was proposed by Williamson in 2002 [1, 2]. The *virus throttle* slows and halts worm propagation without affecting normal network traffic. Williamson focused attention on the differences in network behavior between normal and infected computers. The infected computer attempts to connect with many different computers as fast as possible, while the normal computer tends to connect repeatedly with the same computers. Based on these differences, the *virus throttle* limits the rate of connections to new computers. However, the *virus throttle* passes over slow-speed worms, which propagate to a new computer every few minutes or hours, because it was designed for restricting high-speed worms.

This paper presents a new worm filter for preventing both slow- and high-speed worms from spreading. The worm filter is based on the differences in the number of unacknowledged requests rather than the rate of connections to new computers. In one-to-one communications, a reply is sent for each request. TCP receives an acknowledgment (a SYN-ACK packet) for a connect request (a SYN packet) in the 3-way handshake of TCP. Some popular UDP services, such as the DNS service, provide a reply for a request. Therefore, the normal computer generally receives a reply, although sometimes no reply is received due to human errors, network problems, etc. On the other hand, the infected computer receives almost no replies for worms' requests, because the destinations of their requests are generated almost randomly and active servers on the Internet are very sparse.

The worm filter counts the unacknowledged requests, monitoring both incoming and outgoing packets on a router or a bridge. The unacknowledged count is decreased at regular intervals, because a normal computer sometimes does not receive a reply. If the unacknowledged count exceeds a threshold value, the worm filter drops outgoing packets. If not, the worm filter allows outgoing packets to pass through. Thus, the worm filter is based on the information regarding the normal computer's traffic ("self") rather than the infected computer's traffic ("nonself"). This approach is analogous to that of the immune system in that the immune system defines the "self" and eliminates others as "nonself."

The algorithm of the worm filter is described in detail in Section 2, and Section 3 presents analysis of normal network traffic to determine appropriate parameters for the worm filter. Section 4 describes the implementation of the worm filter and performance evaluation against some worms. Section 5 discusses the weaknesses of the worm filter. Section 6 presents a summary of this paper.

2   Algorithm of the Worm Filter

The worm filter limits the number of unacknowledged requests. It monitors both incoming and outgoing packets on a router as shown in Fig. 1. It monitors SYN packets and SYN-ACK packets of the 3-way handshake of TCP, UDP packets, and ICMP packets. Only outgoing packets whose destination is a broadcast address or a multicast address (e.g., 224.0.0.0 – 239.255.255.255) are excluded, because there are no replies from theses addresses.

The filter has two parameters. One is the time interval for decreasing the number of unacknowledged requests, and the other is a threshold, which is the maximum number of unacknowledged requests allowed.

To count the number of unacknowledged requests, the worm filter has a request table per local computer. The index of the request table is the set of destination addresses, the destination port, and the protocol number of an outgoing packet (i.e., a request). The value of the request table is the time when the worm filter received the request. In the case of ICMP packets, the destination port is set to 0, because ICMP has no port.

The worm filter is divided into two processes for incoming and outgoing packets. The process of outgoing packets makes an index for a request, and the index is looked up in the request table. If an index is found, its value is updated to the time when the process received the request, and the request is allowed to pass through. If no index is found, the process checks the unacknowledged count. If the count is less than the threshold parameter, the index is newly added to the request table, and then the request is allowed. If the count is greater than or equal to the threshold, the worm filter checks the value of the least recently used index in the request table. If the difference between the current time and the value exceeds the time interval parameter, the process removes the least recently used index and its value. Then, the index is newly added and the request is allowed. If the difference does not exceed the threshold, the request is eventually dropped.
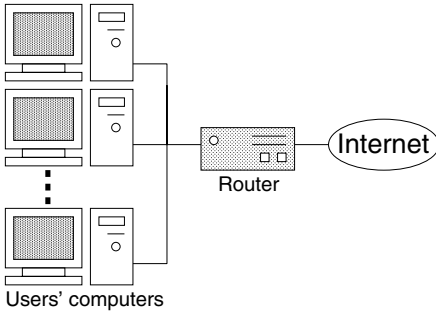
Fig. 1. Network configuration
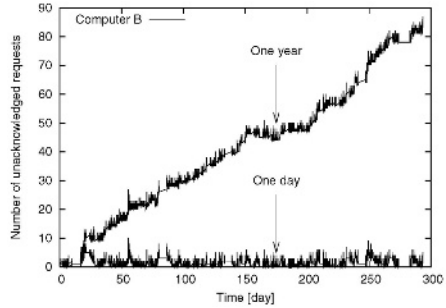
Fig. 2. Dynamics of the number of unacknowledged requests
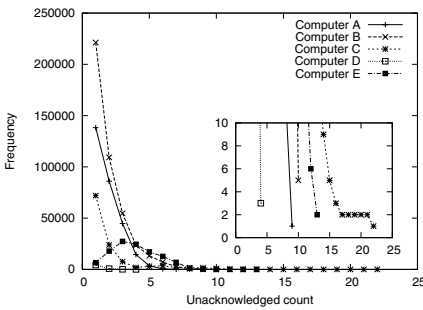
Fig. 3. Frequency of the unacknowledged count
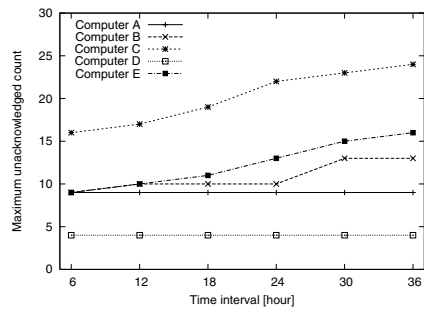
Fig. 4. Time interval vs. maximum unacknowledged count

The process of incoming packets makes an index for an incoming packet, and the index is looked up in the request table. If the index is found, the process removes the index and its value. Note that all incoming packets are allowed to pass through regardless of whether the index is found.

3 Analysis of Normal Behavior

First, we analyzed the TCP network traffic, as most popular network applications, such as web browsers, mail user agents, and instant messengers use TCP. The network traffic was captured from 5 computers, A, B, C, D and E, each with a different network but all running Microsoft Windows® XP Professional operating system. Their capture periods were 301, 299, 194, 43, and 98 days, respectively.

Figure 2 shows two dynamics of the unacknowledged count for computer B. The time interval parameter was set at one year and one day, respectively, and the threshold parameter was not set. In the case of one year, the unacknowledged count gradually increased over a period of 299 days and eventually reaches 84. This means that there were a few unacknowledged requests per day. With the
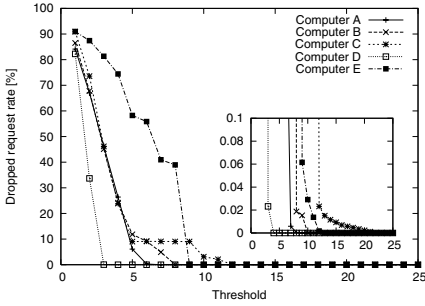
Fig. 5. Threshold vs. dropped request rate
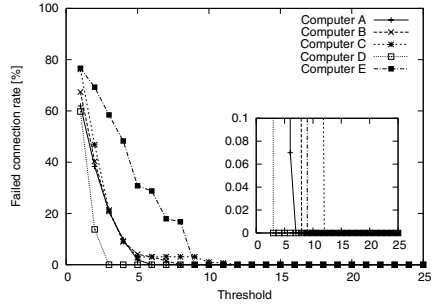
Fig. 6. Threshold vs. failed connection rate
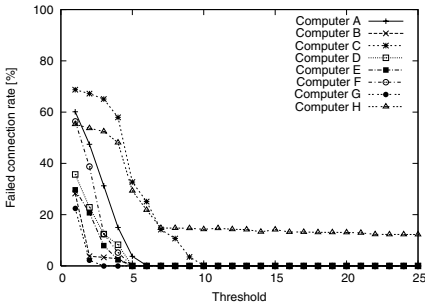
Fig. 7. Threshold vs. failed connection rate (Skype^{TM} packets included)
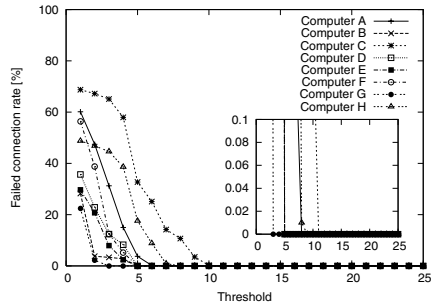
Fig. 8. Threshold vs. failed connection rate (Skype^{TM} packets excluded)

time interval set to one day, the unacknowledged count oscillated between 0 and 10. The unacknowledged count sometimes reached more than 8 instantaneously. Such a sharp momentary increase in number of requests has been observed in other computers. This happens when a user surfs a web page that has images or banners served from many different sites.

Figure 3 shows the frequency of the unacknowledged count. The small inset shows the scaling up of the frequency range from 0 to 10. The time interval parameter was set at one day, and the threshold parameter was not set. The maximum unacknowledged count of computer C was much greater than those of the other computers. The dynamics of the unacknowledged count of computer C indicated that, for only one second, the unacknowledged count increased from 5 to 22 and returned instantaneously to 5. This occurred when the user of computer C surfed more than about 17 web sites at the same time.

Next, the parameters of the worm filter were examined. Figure 4 shows the maximum number of unacknowledged counts as a function of the time interval parameter. The maximum unacknowledged counts for computers B, C, and E increased monotonously, while those of computers A and D remain unchanged. From here, the time interval parameter was set at 24 hours.

Figure 5 shows the rate of dropped requests as a function of the threshold parameter. The small inset shows the scaling up of the range of the rate from 0.0 to 0.1. When the threshold exceeded 21, no requests were dropped. However, this threshold is still not appropriate, because TCP retransmits a connect request (a SYN packet) before aborting the attempt. For Windows® XP Professional, the number of retransmissions is 2 by default. The retransmission timeout is doubled with each successive retransmission in a particular attempt. The initial timeout is 3 seconds. Therefore, if the unacknowledged count decreases to less than the threshold by 9 seconds, the retransmitted request passes through the worm filter before aborting the attempt. Even when the worm filter drops a request, most retransmitted requests would pass through when the sharp increase in the number of requests was only momentary. To evaluate the number of failed connections due to the worm filter, a simulator of the worm filter was implemented. The simulator showed that, even with the threshold set to 12, no connections failed as shown in Fig. 6.

Finally, the network traffic of TCP, UDP, and ICMP were analyzed. The network traffic was captured from 8 computers on a LAN over a period of about 4 weeks. Each computer transmitted more than 10,000 packets, and the total number of outgoing packets analyzed was 450,455.

Figure 7 shows the rate of failed connections as a function of the threshold parameter. It is interesting to note that only computer H maintained a high failed connection rate. For the network traffic of computer H, there were almost no replies for requests from one source port. The source port was used by Skype™, which is a P2P network application for IP phone service. As the P2P network is an ad-hoc network, the Skype™ client does not always receive replies from computers that have sent a reply. In addition, Skype™ attempts to exchange UDP packets with many computers to maintain the connectivity of the P2P network. Therefore, Skype™ caused too many unacknowledged requests. Fortunately, the source port used by Skype™ can be specified by the user. When the worm filter accepted outgoing packets from its source port, the failed connection rate was decreased markedly as shown in Fig. 8.

In conclusion, the appropriate parameters for the worm filter were a time interval of one day and a threshold of the number of unacknowledged requests of 12. Note that outgoing packets from the source port of Skype™ must be accepted if there are Skype™ users in a local network, although the accepted port is a hole in the worm filter.

4   Implementation and Evaluation

The worm filter is implemented as a `netfilter` module on the Linux kernel 2.4.22. The `netfilter` is a framework for packet filtering, network address translation, and other packet mangling. The filter module is built in the Linux kernel using `iptables` 1.2.7a. The use of `iptables` enables us to set up, maintain, and inspect the tables of IP packets filter rules in the Linux kernel.

The worm filter was tested on a network as shown in Fig. 1. The network was a virtual network created using VMware® Workstation 4.5, which enables

multiple operating systems to run concurrently on a single physical computer. The worm filter was installed in the router shown in Fig. 1, which runs Vine Linux 2.4R6. The parameters of the worm filter were a time interval of one day and a threshold of 12. One of the users' computers ran Microsoft Windows® 2000 server, and this computer was ready for infection with the worms, CodeRedIv2, CodeRedII, Slammer, and Blaster, captured in the wild.

To evaluate the time required for the worm filter to stop a worm, the Windows® 2000 server was intentionally infected with each worm. The worm filter stopped each worm in under 0.01 seconds. Only 12 requests were falsely allowed, which corresponded to the threshold parameter. Although the number of allowed requests was a little larger than that of the *virus throttle*, all the stopping times of the worm filter were faster than those of *virus throttle* [2]. In addition, slow-speed worms that attempt to propagate slowly (e.g., only one computer per hour) were simulated, and the worm filter stopped their propagation although it took much longer to stop them than high-speed worms.

To evaluate the usability of network applications, popular network applications, such as Internet Explorer, Outlook®, and MSN Messenger, were used for a period of about 2 weeks. There were no failed connections and no degradation was observed in network performance. When 24 web sites were opened at the same time, only 3 requests were dropped, but these requests were retransmitted after approximately 3 seconds by TCP retransmission. Currently, the worm filter is being tested on a real network of 12 computers. In addition, it is necessary to evaluate network scalability of the worm filter.

5   Weaknesses of the Worm Filter

The worm filter is based on the number of unacknowledged requests. If a worm propagates to only those computers that send a reply, it passes over the worm filter. Such worms do exist, although their numbers are small and they have not yet caused massive epidemics. They are classified into 3 types of worm according to the taxonomy of worms [3]: a *metaserver worm*, a *topological worm*, and a *passive worm*.

The *metaserver worm* obtains a list of target computers from a metaserver, which keeps a list of all the servers that are currently active. For example, Santy worm, which was found on December 21st, 2004, sends search keywords to Google™(`www.google.com`), and obtains a list of target computers from the search results. However, it is not difficult to stop their propagation, because they are dependent on one or more metaservers.

The *topological worm* obtains a list of target computers from the devices of the infected computer. For example, a worm that infects a web server obtains a list of target computers from the web pages stored in its infected web server. As these computers would usually be active, the infected computer would receive a reply.

The *passive worm* either waits for target computers to visit or follows user's requests into target computers. For example, whenever a user on the infected

computer surfs other sites, the worm attempts to infect the sites. The infected computer would receive replies from those sites.

Not only do these worms receive a reply, but also their network traffic is similar to that of an actual user. These behaviors make it difficult for the worm filter to stop their propagation. An alternative filter would be needed to prevent these worms from spreading.

6   Summary

A worm filter based on the number of unacknowledged requests was proposed and implemented on the Linux operating system. Normal network traffic was analyzed and an appropriate parameter set for the worm filter was shown. Performance evaluation showed that the worm filter stopped not only 4 random-scanning worms in the wild, but also simulated slow-speed worms. Finally, some worms that can escape the worm filter were discussed. To stop their propagation, an alternative worm filter would be needed.

Acknowledgments

References

1. Williamson, M. M.: Throttling viruses: restricting propagation to defeat malicious mobile code. In: ACSAC Security Conference 2002. (2002) 61 – 68
2. Twycross, J., Williamson, M. M.: Implementing and testing a virus throttle. In: the 12th USENIX Security Symposium. (2003) 285 – 294.
3. Weaver, N., Paxson, V., Staniford, S., Cunningham, R.: A taxonomy of computer worms. In: the 2003 ACM Workshop on Rapid Malcode, ACM Press. (2003) 11–18

Comparison of Wavenet and Neuralnet
for System Modeling

Seda Postalcıoğlu*, Kadir Erkan, and Emine Doğru Bolat*

Kocaeli University, Faculty of Technical Education Department of Electronics & Computer
Kocaeli, Turkey
{Postalcioglu,ebolat}@ieee.org, erkan@kou.edu.tr

Abstract. This paper presents nonlinear static and dynamic system modeling using wavenet and neuralnet. Wavenet combines wavelet theory and feed-forward neuralnet, so learning approach is similar to neuralnet. The selection of transfer function is crucial for the approximation property and the convergence of the network. The purelin and the tansig functions are used as the transfer functions for neuralnet and the first derivative of a gaussian function is used as the transfer function for wavenet. Wavenet and neuralnet parameters are optimized during learning phase. Selecting all initial values random, but for wavenet, it may be unsuitable for process modeling because wavelets have localization feature. For this reason heuristic procedure has been used for wavenet. In this study gradient methods have been applied for parameters updating with momentum. Error minimization is computed by quadratic cost function for wavenet and neuralnet. Nonlinear static and dynamic functions have been used for the simulations. Recently wavenet has been used as an alternative of the neuralnet because interpretation of the model with neuralnet is so hard. For wavenet learning approach, training algorithms require smaller number of iterations when compared with neuralnet. Consequently, according to the number of training iteration and TMSE, dynamic and static system modeling with wavenet is better as shown in results.

1   Introduction

This paper presents the nonlinear static and dynamic system modeling with wavenet and neuralnet. Neuralnet is an approach towards function approximation and time series forecasting, and it has obtained wide approval in the fields of approximation and forecast [1]. The feed-forward neuralnet have three layers. These are the input layer, the hidden layer and the output layer. The neuralnet performs a non-linear transformation of input data in order to approximate the output data. Wavenet combine wavelet theory and feed-forward neuralnet. Wavenet can be considered as an alternative to neural and radial basis function networks [9]. This property can be expressed as any function of $L^2(R)$ can be approximated to any given accuracy with a finite sum of wavelets. Radial basis functions have been applied in the area of neural networks where they are used as the transfer function. The wavelet theory has found many applications in numerical analysis and signal processing, the wavelet transform of which provides another novel approach towards the function approximation problem. The wavenet is an approach for system identification [4,5]. The goal of system

---

* Student Member, IEEE

identification is the process of generating workable models from observed data. Recently wavenet has been used as an alternative of the neuralnet because of interpretation of the model with neuralnet is so hard [2]. On the other hand training algorithms for wavenet require a smaller number of iterations when compared to neuralnet [3]. The selection of transfer function is crucial for the approximation property and the convergence of the network [6]. Transfer function calculates a layer's output from its net input. In this paper the purelin function is used for output layer and the tansig function is used for hidden layer, which are used as the transfer functions in neuralnet. The first derivative of a gaussian wavelet is used as the transfer function for wavenet. Wavelets show local characteristics [7,12]. Fig. 1 shows respectively the first derivative of a gaussian wavelet, purelin and tansig functions.

(a)   (b)   (c)
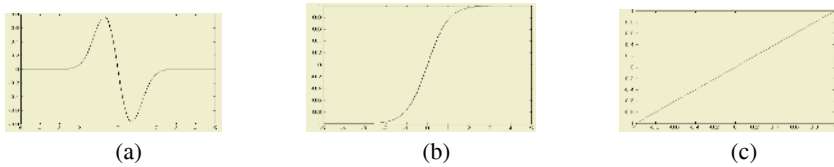
Fig. 1. a) first derivative of a gaussian wavelet b) tansig function c) purelin function

For neuralnet transfer function is constant and functions can not be changed or adjusted during the simulation. On the other hand for wavenet, the transfer function is adjusted by means of dilation and translation parameters. The parameters of wavenet are dilation (d), translation (m), bias (a_0), and weights (a,c). The parameters of neuralnet are weights and bias. The parameters are optimized by gradient method with momentum for neuralnet and wavenet.

2   Structure of Neuralnet

In order to train a neuralnet must calculate how the error changes as each weight is increased or decreased slightly. The back propagation algorithm is the most widely used method for determining the error derivative of the weights. Feed-forward neuralnet sends to data from input to output. The weights specify the strength of the influence [8]. The aim is to determine a set of weights which minimizes the error. The common way for learning is the least mean square convergence. The weight of each connection is changed, so that the neuralnet produces a better approximation of the desired output. Fig. 2 presents three layers feed-forward neuralnet. Radial basis function networks are also feed-forward, but have only one hidden layer. In this learning approach the tansig and the purelin functions are used in hidden and output layers respectively. In this study the neuralnet structure has one hidden layer and output layer.

The back-propagation algorithm computes each error derivative of the weights by first computing the error derivative, the rate at which the error changes as the activity level of a unit is changed. For output units, the error derivative is the difference between the actual and the desired output. The error derivative is computed for a hidden unit in the layer just before the output layer. All the weights between hidden unit and the output units to which it is connected, must be identified. Then multiply those

weights by the error derivatives of those output units and add the products. This sum equals the error derivative for the hidden unit. This is called the back propagation algorithm. In this study the back-propagation algorithm has been used which consists of nine steps for training neural network [13]:
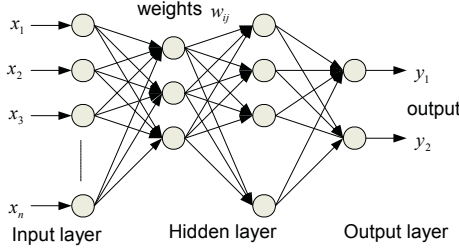
Fig. 2. Three layers feed-forward neuralnet

1. Initial values of weights are determined.
2. Input data x_n are sent from input layer to hidden layer.
3. Hidden layer sums of the input data with weights and this total is applied to the transfer function for providing output data. So output data is sent to the output layer as shown in equation (1)

$$z_{net_{pj}} = \sum_{i=0}^{n} w_{ji} x_{pi} \qquad z_{pj} = f(z_{net_{pj}}) \tag{1}$$

4. Input data with weights are computed using the equation (2) for the output layer.

$$y\_net_{pk} = \sum w_{kj} z_{pj} \quad j = 1...p \quad i = 1...n \quad k = 1...m \tag{2}$$

5. $y_{\_net}$ is applied to the transfer function for obtain output data.

$$y_{pk} = f(y\_net_{pk}) \tag{3}$$

6. Error and (w_{kj}) weights variation are computed using the equation (4) for output layer.

$$\delta_{pk} = (d_{pk} - y_{pk}) f^1(y\_net_{pk}) \qquad \Delta w_{kj} = \mu \delta_{pk} z_{pj} \tag{4}$$

7. Error and (w_{kj}) weights variation are computed using the equation (5) for hidden layer.

$$\delta_{pj} = f^1(z\_net_{pj}) \sum_{k} \delta_{pk} w_{kj} \qquad \Delta w_{ji} = \mu \delta_{pj} x_i \tag{5}$$

8. Weights and bias are updated for output layer and hidden layer using equation (6).

$$w_{kj}(t+1) = w_{kj}(t) + \Delta w_{kj} \qquad w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji} \tag{6}$$

9. Stopping conditions are controlled and the training is stopped.

3   Structure of Wavenet

Recently, due to the similarity between the wavelet transform and a one-hidden-layer neuralnet, the idea of combining both wavelets and neuralnet has been proposed. This

has resulted in the wavenet. The wavelet theory provides useful guidelines for the construction and initialization of the networks and consequently, the training times are significantly reduced. Furthermore wavenet has more freedom by introducing two new parameters for the dilation (d) and translation (m) of wavelet transform and therefore can provide better performance [1]. Wavelet functions can be classified by two categories. These are orthogonal wavelet and wavelet frames. Wavelet frames are used for application of function approximation and process modeling due to the orthogonal wavelets can not be expressed in closed form [10]. Wavelet frames are constructed by mother wavelet which is a prototype for generating the other window functions [14]. A wavelet $\Phi_j(x)$ is derived from its $\phi_z(x)$ mother wavelet. It's shown in equation (7).

$$\Phi_j(x) = \prod_{k=1}^{N_i} \phi(z_{jk}) \qquad z_{jk} = \frac{x_k - m_{jk}}{d_{jk}} \tag{7}$$

N_i, is the number of inputs. N_w is a layer of wavelets. The network output y is computed using equation (8).

$$y = \Psi(x) = \sum_{j=1}^{N_w} c_j \Phi_j(x) + a_0 + \sum_{k=1}^{N_i} a_k x_k \tag{8}$$

The wavenet structure is shown in Fig. 3.
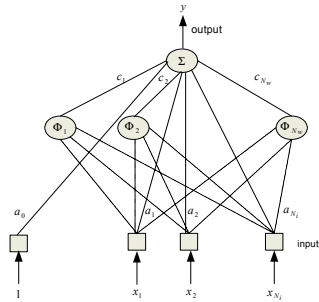
Fig. 3. Wavenet structure

Bias (a_0), weights (a,c), translation (m), dilation (d) are adjustable parameters of wavenet. For wavenet, the first derivative of a gaussian wavelet function is used as a transfer function. It's shown in equation (9).

$$\phi(x) = -xe^{-\frac{1}{2}x^2} \tag{9}$$

The main purpose is to update the parameters during the training phase. θ is the set of adjustable parameters as shown $\theta = \{m_{jk}, d_{jk}, c_j, a_k, a_0\}$ and $j=1...N_w$, $k=1...N_i$. The training is based on the minimization so that quadratic cost function has been used as shown equation (10).

$$\mathbf{j}(\theta) = \frac{1}{2} \sum (y_p - y)^2 \tag{10}$$

For equation (10), y is the network output as shown in equation (8) and y_p is the process output which information flowing out of the system. The minimization is per-

formed by iterative gradient based methods. The partial derivative of the cost function with θ is computed using equation (11).

$$\frac{\partial j}{\partial \theta} = -\sum e \frac{\partial y}{\partial \theta} \text{ and } \theta = \{m_{jk}, d_{jk}, c_j, a_k, a_0\} \tag{11}$$

For parameters updating, equation (12) has been used.

$$\theta_{t+1} = \theta_t + \mu(-\frac{\partial j}{\partial \theta}) \tag{12}$$

θ_t shows parameters and θ_{t+1} shows updated parameters. μ is the learning rate for equation (12). In this study momentum has been used which increases the learning rate with out an oscillation. In equation (13) α is the momentum. Selection of α is done interval (0,1).

$$\theta_{t+1} = \theta_t + \mu(-\frac{\partial j}{\partial \theta}) + \alpha[\theta_t - \theta_{t-1}] \tag{13}$$

4   Initializations of the Network Parameters

Selecting initial values of the network parameters are important because initial value affects the speed of the training and approximation to the global or local minimum. Weights are updated according to the derivative of the transfer function at selecting initial value stage. Functions and derivative of the functions can not be zero. For neuralnet, initial values of the parameters must not be big otherwise it goes to the saturation and learning is occurred slowly. Parameters of the initial values are determined by means of boundary input-output data and feature of transfer function for neuralnet. Initial value is more important for wavenet because wavelets have localization feature [9] and wavelets may be exit from the related domain because of initial value. On the other hand initial values may make some wavelets too local and make the components of the gradient of the cost function very small in areas of interest. So selecting initial values of dilation (d_{jk}) and translation (m_{jk}) randomly may not be suitable for process modeling. In this study the vector m of wavelet j at the center of parallelepiped defined by the N_i intervals $\{[\alpha_k, \beta_k]\}$ and the dilation parameters are initialized to the value as shown in Equation (14) [10].

$$m_{jk} = \frac{1}{2}(\alpha_k + \beta_k) \text{ and } d_{jk} = 0.2(\beta_k - \alpha_k) \tag{14}$$

Equation (14) guarantees that the wavelets extend initially over the whole input domain. The choice of the weights (a,c) is less critical for wavenet. They are initialized to interval (0,1).

5   Stopping Conditions for Training

Parameters of the wavenet and neuralnet are training during the learning phase for approximation the desired function. Gradient methods have been applied for adjustable parameters. When variation of gradient and variation of parameters reaches a lower bound or the number of iterations reaches a fixed maximum, training is stopped.

6   Simulation Results

6.1   Static System Simulation

The approximation of a function of a single variable, without noise, given for the static system modeling has been used, as shown in equation (15) [4]. This example was first proposed in [5], which is one of the seminal papers on wavenet [9].

$$f(x) = \begin{cases} -2.186x - 12.864 & x \in [-10,-2[, \\ 4.246x & x \in [-2,0[, \\ 10e^{(-0.005x-0.5)\sin(x(0.03x+0.7))} & x \in [0,10] \end{cases} \tag{15}$$

System shows different characteristic for different input domains. The domain of the x data is transformed into [-1,1]. The learning procedure is applied on this domain. The number of training sequence is 200 examples which is uniformly distributed in the interval of interest.

Fig. 4 (a) and (b) present the nonlinear static model output and the nonlinear static system output for neuralnet and wavenet. Wavenet result has been obtained with networks of 10 wavelets. Learning iteration is 3000. Learning rates are respectively μ=[1e-2, 1e-2, 1e-3, 1e-3,1e-3] for model parameters ($\theta = \{m_{jk}, d_{jk}, c_j, a_k, a_0\}$). Training Mean Square Error (TMSE) is calculated as 5.0595e-004. Neuralnet result has been obtained with 13676 learning iteration and learning rate is taken 0.1 for all parameters. TMSE is calculated 2.688e-003.

(a)   (b)
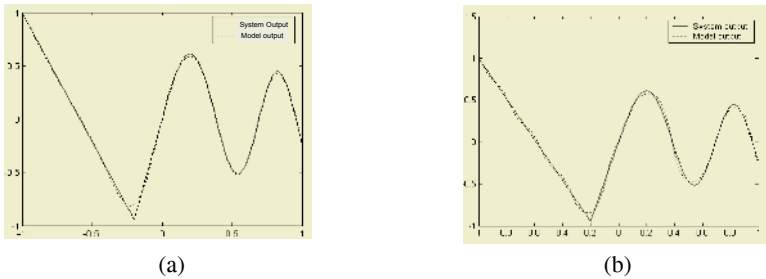
Fig. 4. (a) Wavenet result (b) Neuralnet result

As shown in results, TMSE and number of training iteration are better for wavenet.

6.2   Dynamic System Simulation

This study, propose the use of wavenet and neuralnet to the dynamic modeling for this reason second order nonlinear function is used. The input sequence is u(n) and the measured process output is $y_p(n)$. As in the static case, the aim is to approximate the function by a wavenet and neuralnet. Inputs are past outputs of the process (y_p) and the external input is (u). For the dynamic system modeling we use equation (16) [11].

$$y_p(n) = f(y_p(n-1), y_p(n-2), u(n-1)) = \frac{24 + y_p(n-1)}{30} y_p(n-1) - 0.8 \frac{u(n-1)^2}{1+u(n-1)^2} y_p(n-2) + 0.5u(n-1)$$

$$y(n) = \Psi(y_p(n-1), y_p(n-2), u(n-1), \theta) \tag{16}$$

The input sequence for training consists of pulses with random amplitude in the range [-5,5] and with random duration between 1 and 20 sampling periods.

Fig. 5 (a) and (b) present the nonlinear dynamic model output and the nonlinear dynamic system output for wavenet and neuralnet.

(a)   (b)
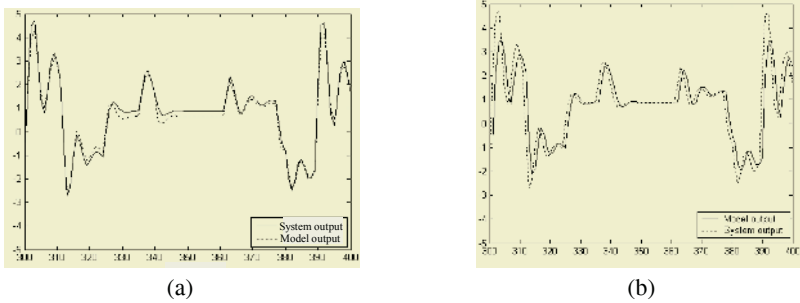
Fig. 5. (a) Wavenet result (b) Neuralnet result

Wavenet result has been obtained with networks of 5 wavelets. Training is computed for 4000 iteration. TMSE is 3.02e-002. Learning rate is respectively μ=[1e-2, 1e-2, 1e-3, 1e-3,1e-3] for model parameters (θ ={mjk,djk,cj,ak,a0}). Neuralnet result has been obtained with 26350 iteration, TMSE is computed as 4.5103e-001 and learning rate is taken 0,1 for all parameters.

As shown in results, TMSE and number of training iteration are better for wavenet.

7   Conclusions

This paper presents nonlinear static and dynamic system modeling using wavenet and neuralnet. For parameters updating gradient method has been used. Error minimization is computed using quadratic cost function. The first derivative of a gaussian function has been selected as the transfer function for wavenet because of universal property [10]. The tansig and the purelin functions have been used as a transfer function for neuralnet which are common used. Wavelets show local characteristic, so the initial values of translations and dilations requires more care than the initial value of the weights. If selecting initial values of the translation and the dilation properly, training is shorter than which is modeled by neuralnet. According to the number of training iteration and TMSE, dynamic and static system modeling with wavenet is better as shown in results.

References

1. He, Y., Chu, F., Zhong, B.,: 2002, "A Hierarchical Evolutionary Algorithm for Constructing and Training Wavelet Networks", pp. 357-366, Neural Comput & Applications 10.
2. Zhang, Q., Benveniste, A.: 1992, "Wavelet Networks". IEEE Transaction on Neural Networks, pp.889-898, vol.3, No. 6.
3. Galvao Roberto K. H., Becerra V. M.; 2002, "Linear-Wavelet Models for System Identification". IFAC15th Triennial World Congress, Barcelona, Spain.

4. Zhang, Q., Benveniste, A.: 1992, "Wavelet Networks". IEEE Transaction on Neural Networks, pp.889-898, vol.3, No. 6.
5. Zhang, Q.: 1997, "Using Wavelet Network in Non parametric Estimation", IEEE trans. Neural Networks, Vol. 8, No.2, 227-236.
6. Jiao, L., Pan, J., Fang, Y.,: 2001, "Multiwavelet Neural Network and Its Approximation Properties", pp. 1060-1066, IEEE Transactıons on Neural Networks, Vol. 12, No. 5.
7. Polycarpou, M., Mears, M. and Weaver, S.: 1997, "Adaptive Wavelet Control of Nonlinear Systems,'" Proceedings of the1997 IEEE Conference on Decision and Control pp. 3890-3895.
8. Liu, Z., He, Z., Qian, Q.: 2002, "Research on Feedforward Neural Network, Wavelet Transformation, Wavelet Network and Their Relations", IEEE, Proceedings of the 2nd International Workshop on Autonomous Decentralized System, pp.277-281.
9. Oussar Y, Dreyfus G.: 2000, "Initialization by Selection for Wavelet Network Training", Neurocomputing, 34, 131-143
10. Oussar Y., Rivals I., Personnaz L., and Dreyfus G.: 1998, "Training Wavelet Networks for Nonlinear Dynamic Input Output Modeling", Elsevier neurocomputing 20, 173-188
11. Oussar Y., Rivals I., Personnaz L., and Dreyfus G.: 1998, "Training Wavelet Networks for Nonlinear Dynamic Input Output Modeling", Elsevier neurocomputing 20, 173-188
12. Becerikli, Y., 2004,"On Three Intelligent Systems: Dynamic Neural, Fuzzy and Wavelet Networks For Training Trajectory", Neural Computing & Applications (NC&A), vol.13, n0.4, pp.339-351.
13. Yıldırım, M., 1998,"Bulanık Mantıklı Yapay Sinir Ağı ile Doğrusal Olmayan Sistem Modelleme", Fen Bilimleri Enst. Y. lisans tezi, Kocaeli Üniversitesi.
14. Polikar R.: 2001, "The wavelet tutorial"
   http://engineering.rowan.edu/~polikar/WAVELETS.

Neurone Editor: Modelling of Neuronal Growth with Synapse Formation for Use in 3D Neurone Networks

Johan Iskandar and John Zakis

Electrical and Computer Systems Engineering Department, Monash University
Clayton, Victoria 3168, Australia
{Johan.Iskandar,John.Zakis}@eng.monash.edu.au

Abstract. This paper describes Neurone Editor as an editing tool that allows different types of neuronal cells to be modelled by using Java and Java3D. It provides a means of generating realistic neuronal geometries and their expected growth patterns. Each neurone type, as described by the parameters and the variance attributed to each parameter, allows the generation of multiple instances of similar neurons. In this way, we can generate large neurone networks in a biological pattern with the aim of generating computational structures of that type and also as an aid to visualize these structures. With these aims in view, detailed and anatomically correct renditions are not provided but the aim is to replicate natural growth patterns and the resulting interconnections on the assumption that shape defines function.

1 Introduction

Various types of neurones in the nervous system can be distinguished by differences in cell body size and shape, and by different branching patterns of the dendrites and axons. Therefore many applications have been built to model neurones of different types for different purposes such as Morphology Editor [1], L-Neuron [2], [3] and NEURON [4]. We built Neurone Editor [5] as a visualization tool to model complex neurone structures and their growth characteristics. It is created by using Java and Java3D to complement NeuroneSim by adding neurone editing capability to NeuroneSim [6], [7]. Once a neuronal geometry is defined using the neurone editor, these features can then be used to generate a neurone network with controlled randomised growth patterns. The resultant networks and their interconnections can be saved in files for later analysis.

1.1 Neuronal Geometry and Neurone Networks

A neurone is made of three components, the cell body (soma), the dendrites, and the axon. The dendrites are normally elaborate tree-like structures whose purpose is to collect electrical impulses from the axons of neighbouring cells; the axons therefore serve to transmit electrical impulses to the dendrites of other cells [8]. The electrical impulses collected by the cell's dendrites are summated, and if a threshold is reached, then the cell generates an electrical impulse which is transmitted by its axon [9]. These connections are known as *axodendritic synapses.* It is these connections that describe a *neurone network.*

The branching patterns of dendritic arrays are probably established by random adhesive interactions between dendritic growth and the axons during development. There is an overproduction of dendritic segments in early development, and as the network matures and information is processed through the dendritic tree, the branches become pruned and segments redistributed, perhaps in response to functional demand. Neurones with similar functions have similar tree structure as shown in Fig. 1.
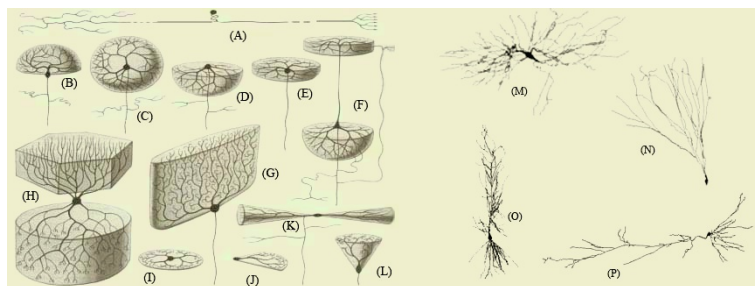
Fig. 1. Pattern variations of neuronal geometry [1], [10]. (A) Unipolar, (B) Bipolar, (C) Stellate, (D), (E) and (K) are modifications of Stellate, (F), (O) and (P) Pyramidal, (G) Purkinje, (H) Golgi, (I) and (J) Amacrine, (L) Glomerular, (M) Hilar, (N) Granule

Table 1 describes the Hillman parameters for various cells which are related to the 3D structure of a dendrite. For instance, high daughter-branch ratios means that many uneven bifurcations exist; this produces a relatively narrow tree growing from a single thick trunk, as in the Pyramidal cell in Fig. 1 (F), (O), and (P). In contrast, a daughter ratio closer to one leads to "bushy" trees, as the bifurcations are more even [11].

Table 1. Statistical distributions of Hillman's fundamental parameters for several morphological classes [11]. The values for the number of trees (N), branch section area (S), branch length (L), daughter diameter ratio (R), taper (DA), Rall's power (n), and bifurcation angle (a) are reported as normal (mean±standard deviation) or uniform (range) distribution

	Purkinje	Stellate	Granule	Motoneuron	Pyramidal (Apical)	(Basal)
N_{tree}	1 ± 0	4 ± 1	3.5 ± 1.5	10 ± 4	1 ± 0	6 ± 2
S_{stem}, µm² (S_{trunk} µm²)	109.3 ± 14	2.75 ± 1.35 (11.0 + 3.4)	1.51 ± 0.79 (5.48 + 0.94)	88.4 ± 56 (886 + 71)	12.8 ± 13 (76.8 ± 21)	
S_{stem}	1.12 + 0.60	0.51 + 0.27	0.49 + 0.28	N.D.	1.41 ± 0.64	
L_{trunk} µm (L_{stem} µm)	11.6 + 9.2 (+50%)	31.7 + 23 (+0%)	10.7 + 8.4 (−0%)	N.D.	70.8 ± 65 (+75%)	
R(d-d)	[1-6]	[1-2]	[1-2]	[1-2]	[2-6]	[1-2]
AA (Taper)	>10% (stem) 0% (else)	0%	0%	>10%	>10%	0%
Rall's v	2.36 ± 1.2	2.24 + 1.12	2.58 ± 1.8	1.69 + 0.48	1.99+0.79	2.28+0.89
α, deg	36 ± 29	39 ± 10	42 ± 13	17 ± 9	61 ± 21	

2 Using L-Systems for Modelling of Trees

In order to model a neurone; there have to be a set of parameters which will lead to the generation of a virtual neurone. Research done by Hillman [11], Burke [12], or Tamori [13] describes the extraction of these necessary parameters and construction of algorithms for neurone generation. However, a full description of the orientation of

each branch segment in space needs to be determined first in order to fully create a 3D structure. In none of the work by Hillman, Burke, or Tamori has shown how to do this.

Aristid Lindenmayer in 1968 proposed L-systems as a mathematical formalism for the description of plant development. Its core functionality is the flexibility to specify virtually all kinds of branching patterns and thus it can be used to model the morphology of a variety of organisms such as tree structures [14]. Since dendrite growth is very similar to the growth of plants and trees, L-systems can also be applied to the generation of neurone growth.

Imagine a turtle that moves through space and traces out a line segment as it moves. Its movements are specified by a series of productions. At each point in space, a turtle has three vectors associated with it – a vector that specifies the current direction in which the turtle is heading (the **H** vector, or Heading), a vector that points to the turtle's left (the **L** vector, or Left), and a vector that points up from the turtle's point of view (the **U** vector, or Up). A turtle's heading, left or up directions can be changed by applying **R** rotations. All three vectors are orthogonal to each other as shown in Fig. 2 (A).
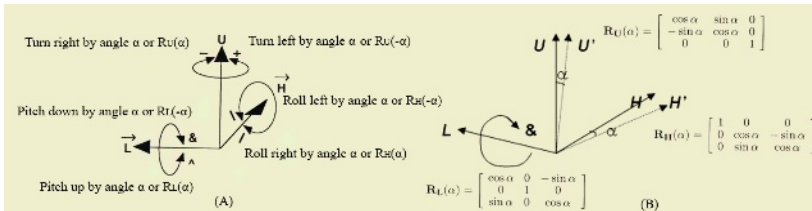
Fig. 2. (A) The vectors **H**, **U**, and **L** associated with a turtle as it moves through space. Note that **H** × **L** = **U.** (B) The **U** and **H** vectors are rotated anti-clockwise about the **L** vector by angle α, resulting in vectors **U'** and **H'** [14]

2.1 Turtle Interpretation of an L-System

The following is an example of a set of L-system productions:

N = 20
ω : A
p1: A → F[&(a)B][/(b)C]
p2: B → F[^(d)C][&(c)B]
p3: C → F[\(e)C][+(f)B]

ω is the start of a production which invokes the subsequent productions. This effectively states that wherever A appears, it is to be replaced by F[&(a)B][/(b)C]. Then B and C are replaced by F[^(d)C][&(c)B] and F[\(e)C][+(f)B] respectively, as specified by productions p2 and p3. The result is then F[&(a)F[^(d)C][&(c)B]][/(b)F[\(e)C][+(f)B]]. B and C are then once again replaced by the productions p2 and p3, N times. Note that left square bracket is an instruction to save the current state of the vectors **H**, **L**, and **U**. In contrast, the right square bracket is an instruction to restore the old vectors **H**, **L**, and **U** [14]. All symbols can be seen in Fig. 2 (A). The following production p1 is a summary of this translation from an L-system production into pseudo code.

```
A([H L U], current turtle position)
{
    //draw line of unit length in space, whose
    //orientation is determined by the vector H
    F(H, current turtle position)
    Save state of vectors H, L, U //push to stack
    Compute new state [H' L' U'] = [H L U]R_L(-a)
    B([H' L' U'], current turtle position)
    Restore state of vectors H, L, U // pop off stack
    Save state of vectors H, L, U // push to stack
    Compute new state [H' L' U'] = [H L U] R_H(b)
    C([H' L' U'], current turtle position)
    Restore state of vectors H, L, U // pop off stack
}
```

It can be seen that from the production by p1, that the functions B and C are called. This can be seen as a bifurcation or branching as both functions receive headings or directions used for drawing the line segments. Different line segments are drawn by function B() and C(), both segments originating from the same point. This is the formation of a branch.

3   Method and Design

The most important class of the Neurone Editor is named LBranch, which extends the Branch class and contains the L-system methods such as RotateH(), RotateL() and RotateU(), which perform the matrix rotations $\mathbf{R_H}(\alpha)$, $\mathbf{R_L}(\alpha)$, and $\mathbf{R_U}(\alpha)$ as shown in Fig. 2 (B). This class also has a rotate() method which can rotate an entire dendrite or axon tree (not just change the orientation of one segment). There is also a BoundsChecker class, which is used by the Axon and Dendrite classes to restrict their growth to within certain geometric bounds. Furthermore, RotatableCell extends the abstract class Cell, this class is virtually empty, except for overriding the method grow().

Additional parameters are added to Neurone Editor whilst preserving the parameters that are used in NeuroneSim. These parameters are the Tropism vector (**T**); a vector **V**, that aids in defining a branch plane; the amplitude (or magnitude) of the tropism; initial direction of growth of the dendrite from the cell body (vector **d**); and number of dendrites.

- Tropism is the bending effect of an entire tree structure in a particular direction. There are two ways in specifying tropism. Firstly, axial tropism which acts parallel to the starting direction of growth and secondly, the direction of the tropism vector relative to the growth axis, alpha and beta (in degrees).
- The amplitude of tropism corresponds to the coefficient *e,* is shown in Fig. 3 (B). The magnitude of this parameter represents the bending effect, if the amplitude of tropism is zero, there will be no bending effect.
- The vector **V** is a parameter that defines the plane in which a bifurcation occurs. A bifurcation can take place either in the **L-H** or **H-U** planes as in Fig. 3(A)

Dendrites initially start growing at right angles to the soma surface [10] which is the surface of a sphere. If vector **d** is the initial direction of growth, *r* is the radius of the cell body, and *p* is the centre of the cell body, then *p* + *r***d** will be a vector pointing

to the starting position of growth relative to the location of the cell body. All three vectors **T**, **V**, and **d** are specified by the user, using spherical coordinate. Fig. 4 shows the spherical coordinate system, which is used to allow independent variances in the growth parameters.

(A)   (B)
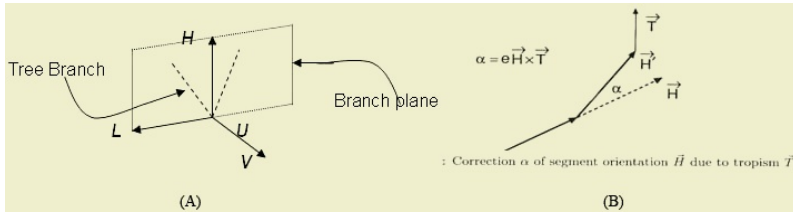
Fig. 3. (A) **V** ×**H** = **L** defines a plane where a tree branch can exist. Here, **V** is normal to the branch plane. If **V** is perpendicular to **H** then **V** = **U**. (B) The heading of the turtle is altered by vector **T**. This causes the tree structure to be bent in the direction of the tropism vector (**T**) [14]
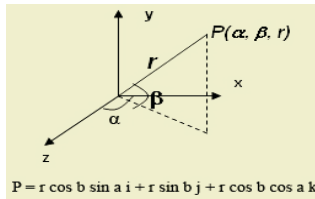
Fig. 4. Spherical coordinate system used to describe growth of cells. Vectors **i**, **j**, and **k** are parallel to axes x, y and z respectively

3.1   L-Systems Implementation in Neurone Editor

TrunkTreeDendrite and TrunkTreeAxon are both sub-classes of LBranch class that implements the following L-system specification:

```
ω: Trunk(l)
p1: Trunk(l)→F(l)[&(a)Branch(l)][$Trunk(l)]
p2: Branch(l)→F(l)[+(a/2)Branch(l)][-(a/2)Branch(l)]
```

where l is the segment length, a represents the bifurcation angle, and $ means that the **U** and **L** vectors are randomly positioned for the growth of each new segment of the trunk. This L-system produces a straight trunk with branches bifurcating from it as illustrated in Fig. 5 (A).

We can also have a stochastic L-system, where there is a probability that a particular production is used. For instance, we could have the probability k to use production `Trunk(l)→F(l)[&(a)Branch(l)][Trunk(l)]`, and the probability (1-k) to use production `Trunk(l)→F(l)Trunk(l)`.

The production   `Trunk(l)→F(l)Trunk(l)` corresponds to the situation where a bifurcation does not occur, and similarly, so does   the production `Branch(l)→F(l)Branch(l)`.

The RegularDendrite and RegularAxon classes which are derived from LBranch, implement the following L-system specification:

```
ω: Branch(l)
p1: Branch(l)→F[+(a/2)Branch(l)][-(a/2)Branch(l)]
```

This produces a simple branch pattern as shown in Fig. 5 (B).
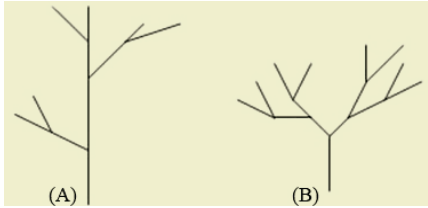
Fig. 5. (A) A tree rooted at a trunk and (B) a regular branch pattern [14]

3.2 Modelling of Neuronal Morphology

The graphical user interface of the Neurone Editor allows the input of parameters to create different neuronal geometries by using the data in Table 1. As an example, Fig. 6 shows models of Stellate, modification of Stellate, Pyramidal and Purkinje cell. The resulting models compare favourably with Fig. 1 (K), (C), (G) and (P).
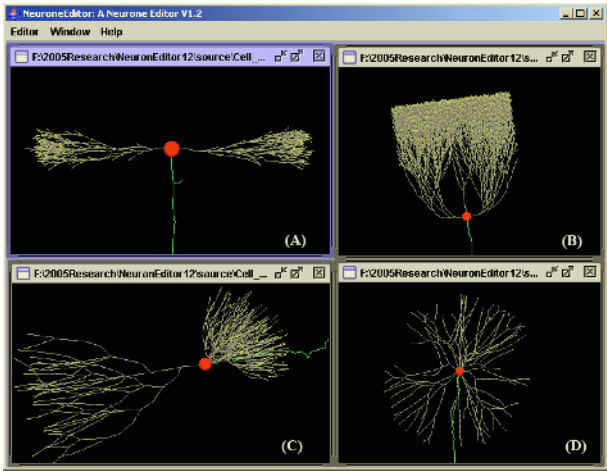
Fig. 6. The resulting (A) modification of Stellate and (D) Stellate, (B) Purkinje, and (D) Pyramidal cells

4 Conclusions and Future Directions

Neurone Editor is built to model different types of cells and their growth patterns. The resulting neurone objects can then be used in NeuroneSim to build large structures of neurones that may model different parts of the brain. The choice of Java 3D for this application was for reasons of portability and integration with OpenGL and Direct3D. This allows sophisticated graphics development at a high level of abstraction and the efficient utilisation of available graphics hardware [15].

Neurone Editor is used to generate dendritic structures that morphologically re-semble real neuronal dendrites by using constrained randomised growth. By careful selection of constrains we can grow neurones that resemble biological neurones of any desired type, but where each one is different to its neighbours just as in real life. By doing this we achieve three useful outcomes:

1. The neurone editor can be used to visualise the organic growth of a single neu-ronal cell in 3D, and it can be viewed from any angle and degree of magnification.
2. Generated cells can be saved in a database of neurones. Various types of neuronal cells then can be selected from the database for the purpose of modelling different parts of the brain. Neuronal cells in the database can be re-edited for different in-put parameters thereby creating a new type of cell.
3. By using NeuroneSim we can obtain a listing of neurone connections or synapses for use outside of this graphical modelling application to create large neurone networks.

These neurone networks can then be simulated using neurone network simulation tools. One reason for doing this is to find if large network structures, like those of the brain, determine function based on structure as determined by the types of neurones and their growth patterns [16]. By choosing VHDL as our neurone network construc-tion and simulation tool, thus we can build them on a silicon chip. Silicon chips con-taining a large number of neurones are not only possible but will be extremely fast compared with biological processes. They will also much faster than algorithmic computations of neural networks, because the silicon neurones all operate concur-rently.

References

1. J. Chad, "Duke / Southampton Archive of Neuronal Morphology," Available: http://www.cns.soton.ac.uk/~jchad/cellArchive/cellArchive.html
2. G. Ascoli and J. L. Krichmar, "L-neuron: A modeling Tool for the Efficient Generation and Parsimonious Description of Dendritic Morphology," *Neurocomputing*, vol. 32-33, pp. 1003-1011, 2000.
3. G. A. Ascoli, "Neuroanatomical algorithms for dendritic modelling," *Network: Computa-tion in Neural Systems*, vol. 13, pp. 13, 2002.
4. M. L. Hines and N. T. Carnevale, "The NEURON simulation environment," *Neural Com-putation*, vol. 9, pp. 1179-1209, 1997.
5. D. Heibert, "Java 3D Modelling of Neurons," in *ECSE*. Melbourne: Monash, 2001, pp. 70.
6. J. D. Zakis, A. Zubovic, and B. J. Lithgow, "Modelling of the Inferior Colliculus using Java 3D," presented at IEEE Engineering in Medicine and Biology Society Conference - Biomedicial Research, Monash University, Clayton Campus, Victoria, Australia, 2001.
7. J. Iskandar and J. Zakis, "Neurone Network Modelling of the Inferior Colliculus using Java3D," presented at The 2nd International Conference on Artificial Intelligence in Sci-ence and Technology, Hobart, Tasmania, 2004.
8. J. Nolte, *The Human Brain: An Introduction to Its Functional Anatomy*, 5th ed. St. Louis, Mo.; London: Mosby, 2002.
9. R. Rhoades and R. Pflanzer, *Human Physiology*, Third ed. Orlando: Saunders College Pub-lising, 1996.
10. H. Gray, P. L. Williams, and L. H. Bannister, *Gray's anatomy*, 38th ed. New York; Ediburgh: Churchill Livingstone, 1995.

11. D. E. Hillman, "Neuronal Shape Parameters and Substructures as a Basis of Neuronal Form," in *The Neurosciences: Fourth Study Program*, F. O. Schmitt, and Worden, F.G, Ed. Cambridge, Mass: The Mit Press, 1979, pp. 1185.
12. R. Burke, W. Marks, and B. Ulfhake, "A parsimonious description of motoneuron dendritic morphology using computer simulation," *J. Neuroscience*, vol. 12, pp. 2403-2416, 1992.
13. Y. Tamori, "Theory of dendritic morphology," *Physical Review E*, vol. 48, pp. 3124-3129, 1993.
14. P. Prusinkiewicz and A. Lindenmayer, *The algorithmic beauty of plants*. New York: Springer-Verlag, 1990.
15. D. Selman, *Java 3D Programming*. Greenwich, Conn.: Manning, 2002.
16. J. Iskandar, "Modelling of Neurone Networks in Java and VHDL," presented at Cybernetics and Intelligent Systems, Singapore, 2004.

A Hybrid Decision Tree – Artificial Neural Networks Ensemble Approach for Kidney Transplantation Outcomes Prediction

Fariba Shadabi[1], Robert J. Cox[2], Dharmendra Sharma[2], and Nikolai Petrovsky[1]

[1] Medical Informatics Centre, Division of Health, Design and Science
University of Canberra, ACT, 2601, Australia
[2] School of Information Sciences and Engineering
University of Canberra, ACT, 2601, Australia

Abstract. The learning strategy employed in neural networks offers a good performance even in the situations where a model is presented with incomplete and noisy data. However, neural networks are known as 'black boxes' as how the outputs are produced is not clear. In this study, a hybrid learning strategy, namely RDC-ANNE (Rules Driven by Consistency in Artificial Neural Networks Ensemble) is proposed. This paper looks at the use of RDC-ANNE in the graft outcome prediction domain as a prototypical medical application. At first, for a better generalization, a committee of binary neural networks is trained. Then, a partial C4.5 decision tree is built from a specifically selected dataset, generated based on the graft data used to test the trained neural networks ensemble. Finally the most appropriate leaf in every path is converted into an understandable rule. In this approach, for the rule generation process, we enforced the model to mainly consider the patterns that their class labels were consistently causing agreement across the neural network classifiers. Experimental results show that the RDC-ANNE method is able to extract partial rules from an ensemble model and reveal the important embedded information of a trained neural network ensemble.

1 Introduction

An ideal data mining tool for medicine decision making should be able to learn the underlying relationships presented in large clinical datasets, extract maximal predictive information, and clearly reveal new and useful patterns that predict particular outcomes or treatments.

Artificial Neural networks (ANN), decision trees and rule generating data mining techniques have been implemented successfully in diverse domains including medical diagnosis [1], [2]. The attractiveness of decision trees and rule generating models is largely due to their ability to clearly explain and justify their decisions as symbolic rules. However where a decision tree model is presented with incomplete and noisy data, it may not generate reasonable results. The learning strategy employed in ANN, on the other hand, offers a good performance even in situations where the problem domain is not well understood. ANN models provide insights into the relationships and patterns hidden in a set of data based on patterns of connectivity that bond the nodes of the network together. The ANN model uses these patterns to classify unseen inputs, the problem being that the way classifications or predictions are derived is not obvious. Several studies have investigated the design of hybrid learning models that

combine the power of ANN with rule induction methods [3], [4], [5]. These models can be useful for revealing previously unknown complex relationships that have been learned by ANN during the training process. A survey paper written by Andrews et al. [6] provides a good review of some of the existing hybrid knowledge based-ANN approaches.

It has been shown that clinical databases that contain considerable noise often significantly deteriorate the generalization ability of ordinary prediction models. This paper proposes a hybrid learning approach namely RDC-ANNE (Rules Driven by Consistency in Artificial Neural Networks Ensemble) and uses a trial data set from a kidney transplant database as a prototypical medical application. In this study, at first we investigate the possibility of building a model that could provide better generalization for complex and noisy clinical data by using a committee of binary neural networks. Then, we modify the model further in order to extract the input patterns (examples) that were included across the ANN series in the final results. Patterns that consistently generate agreement across the classifiers can be considered as examples with positive impacts or higher predictive sources. Finally, we aim for a better clarity by extracting rules from a trained ANN ensemble, using a learning-based rule-extraction technique.

2   Kidney Transplant Challenges

The financial cost of kidney disease in Australia is extremely large and with an ageing population, the number of people with end-stage kidney diseases is expected to increase dramatically in the next decade. The treatment for patients with end-stage kidney diseases is either dialysis or kidney transplantation. Currently, there are thousands of men, women and children on the kidney transplantation waiting list, who are anxiously waiting for a donor organ that could give them a second chance at life.

Given the critical shortage of kidneys for transplantation, a variety of techniques have been proposed to estimate the probability that a kidney graft will continue functioning for a certain period of time after surgery. Generally, studies have revealed that clinical data gathered from patients who have undergone graft transplant surgery are very complex and not easy to generate predictions from [7], [8]. As mentioned earlier, in terms of accuracy, ANN models have the ability to provide good solutions in complex situations where a large number of variables contributes to an outcome but their individual influence is not well understood. However unlike decision tree models, the classification rules are not directly accessible from an ANN model. This study is concerned with creating a hybrid decision tree - ANN ensemble approach for kidney transplantation outcomes prediction. Our method is similar to [3] and [5], but this approach provides improvement in terms of comprehensibility of extracted rules from the trained ANN ensemble, especially where many noisy examples are available.

3   Material and Methods

For the purpose of this study, the following methodology was employed:

1. Pre-process the data set. This includes: extracting the data from different tables, cleaning the data, transforming the nominal attributes into numeric attributes, and

choosing the appropriate parameters to be included in the dataset with the help of domain expert.

2. Split the dataset for training and testing (with balanced distribution of success and failure cases).
3. Perform classification using the ANN ensemble approach.
4. Generate new training sets by extracting the patterns (examples) that were consistently causing agreement across the ANN classifiers, in the testing phase.
5. From the new training sets generated in step 4, choose a reasonably big training set that has provided both a good level of   accuracy (based on its corresponding classification table) and a reasonable amount of model agreement.
6. Extract rules from the patterns in the chosen training set, using the WEKA PART decision list.
7. Analyze the results.

It should be noted that prior domain knowledge can also be applied to choose the significant rules that could be useful for predicting the outcomes of a medical event (e.g. failure or success of graft transplants) or to form new concepts. The pseudo-code of RDC-ANNE strategy is shown in Table 1.

3.1 Data Set

The dataset used in this project was obtained from a kidney transplant and dialysis database [9]. Data was sampled at the time point of 2 years after transplantation. In effect, we were trying to predict success or failure of the kidney graft in the 2 years after the transplant. After balancing the data, we obtained a final data set of 657 fail points and 657 success points. Some variables were removed because they are measured after the transplant has been made and they were therefore actually an indication of the outcome of the transplant (see also [10] and [11]). The 16 variables that were retained are AGE (Recipient age at transplant), MISA (Number of mismatches A), MISB(Number of mismatches B), MISDR (Number of mismatches DR), MISDQ (Number of mismatches DQ), REFHOSP (Referring hospital), REFSTAT (Referring state), DONHOSP (Donor hospital), DONSTAT(Donor state), TRANHOS (Transplant hospital), TRANSTA (Transplant state), DONSOUR (Donor source), DONAGE (Donor age), DONSEX (Donor sex), ISCHEMIA (Total ischemia to nearest hour),and KIDPRESI (Initial kidney preservation). For ANN training, we divided the data set into three equal sized sets, the training set, the overtraining prevention (tuning) set and the test set. We also pre-processed the data further by performing normalization. It should be noted that for classification using PART decision list in the data mining tool WEKA [12], the data was converted to ARFF format.

3.2 Artificial Neural Network Ensemble

ANN ensembles have been used successfully in a number of medical domains [1]. An ANN ensemble consists of several individually trained ANN classifiers that are jointly used to solve a problem. This form of regularization could offer a significant improvement in prediction accuracy (for a detailed discussion and description, see [11] and [13]).

Table 1. The RDC-ANNE approach

Given a training set S (with n cases and their desired classes), learning algorithm L, number of bootstrap samples T and a test set V, the following steps were used to generate rules from an ANN ensemble:

Step 1: Train and test an ANN ensemble.

1. Model Generation and Training
 For i = 1 to T:
     Generate a new training set (bootstrap sample) with n cases, using random drawing (with replacement) from S

     Apply the learning algorithm L to the bootstrap sample

     Keep the resulting neural network model M_i for future use

2. Prediction
Repeat the following procedure for every case in the test set V:
For i = 1 to T:

     Predict class of case using M_i

Return class that appears most frequently

Step 2: Obtain a list of training data as the possible candidates for rule induction.
For i= T/2 to T

     Generate training set N_i by extracting all cases and their desired classes in test set V

   that have caused agreement on one category (fail or success) across at least i models

     Modify the new training set N_i by replacing the desired classes of all cases with their

   corresponding class labels (i.e. the class assigned by the trained ensemble)

     Produce the classification table C_i by comparing the corresponding class labels output

    from the ensemble and the desired classes

     Keep the resulting training data N_i and C_i for future use

Step 3: Extract rules from the most appropriate training set.

First, choose a reasonably big training set from N_T, N_{T-1}, N_{T-2} … $N_{T/2}$ that has pro-

vided both a good level of   accuracy (based on its corresponding classification table C) and reasonable amount of model agreement.

Then, grow a C4.5 decision tree from samples and their new class labels in the chosen train-
ing set.

Finally, generate rules by converting the most appropriate leaf in every path into a rule.

   In this study, a series (n=500) of feed-forward, back-propagation networks were trained independently of the others, to differentiate between successful and unsuc-
cessful transplants. We implemented the majority voting strategy where each compo-
nent classifier votes for a category and the category with the majority of votes defines the ensemble category. All networks were implemented using the Delphi program-
ming language and consisted of a set of 16-input neurons and 2-output neuron (suc-
cess and failure). Sigmoid transfer functions were used for both the hidden layer and output layer. The training algorithm used a trial and error approach to determine the appropriate ANN topology (i.e. the best training constant, number of epochs and hidden nodes) automatically.

3.3   Bagging

Bagging (bootstrapping aggregates) was originally proposed by Breiman [13]. Bagging is a popular method for training component neural networks. This technique generates several training sets using random drawing (with replacement) from the original training set. Consequently, in every new training set there are data points which appear more than once while others do not appear at all.

3.4   PART Decision List

Decision trees and rule generating data mining techniques have the ability to model non-linear relationships with logical rules [14]. A study conducted by Murthy [2] provides a good multidisciplinary survey of recent real world applications of decision trees such as C4.5 [15]. This study shows that decision tree programs can perform well in many situations and generally they are not very difficult to construct. However to construct decision trees from real-world data, a fair amount of understanding of the problem under investigation and the methodology behind this technique is required.

For the purpose of this study the WEKA PART decision list [12] was used. The PART decision list is a rule induction strategy that builds a partial C4.5 decision tree in each iteration and converts the most appropriate leaf in every path into a rule [16].

4   Results

Evolution of an ANN ensemble's predictive ability was started by training a series (n=500) of Multilayer Perceptron (MLP) networks independently of the others, to differentiate between successful and unsuccessful transplants. In this experiment rather than reporting predictive accuracy alone to show best model choice, we modified the program in order to show the input patterns (examples) that were included across the ANN series in the final results. Patterns that consistently were in agreement across the classifiers can be considered as examples with positive impacts or higher predictive sources.

In this experiment, the balanced test set reached 70% accuracy rate with 87%-agreement among the networks (435 of 500 networks), based on 19% of data points (i.e. 84 cases). The results are shown in Fig. 1. This model was able to classify about 87% of successful transplants and 54% of unsuccessful cases. It should be mentioned that the accuracy rate also reached 76% with 89%-agreement among the networks (446 of 500 networks), using only 63 examples (14% of data points).

For rule generation using the PART decision list, at first, we applied a more conventional strategy and generated a new training data set by feeding the entire examples in the test set to a trained ensemble and replacing the true class labels of the original test instances with the class labels assigned to them by the ensemble. The rule set generated by PART decision list produced 26 rules. The rule set includes 2 rules with 7 clauses (conditions), 2 rules with 6 clauses, 1 rule with 5 clauses, 4 rules with 4 clauses, 4 rules with 3 clauses, 8 rules with 2 clauses, and 5 rules with 1 clause. The rules were valid for 94 % of cases in the data set. The rule set can be made substantially easier to understand by enforcing the model to mainly consider the examples whose class labels consistently caused agreement across the ANN classifi-

ers. In effect, this strategy tries to safely remove the branches of rules that were generated by the presence of noise in the data set. The following rule set was produced by applying RDC-ANNE approach based on the 84 examples whose class labels (outputs from the ensemble) were in agreement across 87% of classifiers:
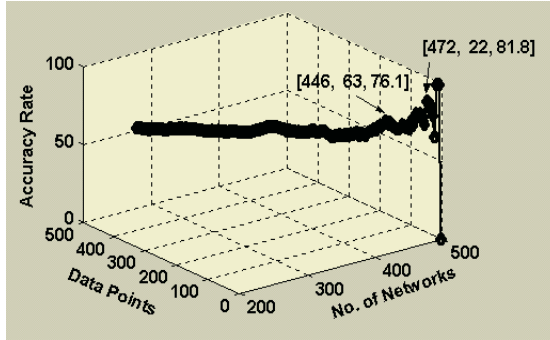
Fig. 1. The results for neural network ensemble with 500 bagging

1. donstate <= 4 AND donage <= 43: **Success** (38 cases) **2**. donhosp <= 109 AND misa <= 1 AND refstate > 3 AND misb > 0 AND refhosp > 94: **Failure** (11 cases) **3.** misa <= 1 AND misb <= 1: **Success** (16 cases /1 false positive) **4.** donsex > 0 (i.e. Female): **Failure** (9 cases) **5.** misa > 1: **Failure** (5 cases) **6.** age > 27: **Success** (3 cases /1 false positive) **7.** Else: **Failure** (2 cases).

As it can be seen, in this experiment, the rule set produced fewer rules (only 7 rules) with fewer clauses. These rules were valid for 97 % of cases (i.e. 82 cases) in the data set.

5   Conclusion and Discussion

In this paper, we implemented an ANN ensemble model, using a simple voting mechanism, with each classifier having one vote. After training the ANN ensemble, we fed each pattern or example in the test set to the ensemble model and recorded the votes. Then we searched for patterns or examples that caused at least 50% of the networks to agree on one category (failure or success). Obviously this data set included the entire patterns. After replacing the desired classes of these patterns with their class label (i.e. the class assigned by the trained ensemble), we produced our first candidate for training the decision tree. Then we searched for patterns that caused at least 51%-agreement among the networks. We repeated this procedure until the model reached 100%-agreement among the networks for one category (failure or success). It is clear that as the model approaches the 100%-agreement condition, although the accuracy may increase, fewer patterns can satisfy the condition and be included in their corresponding data set. Therefore, the patterns presented in the data set (driven from the test set) that were chosen for training the decision tree (e.g.87%-agreement data set) may not be adequately representative of all possible scenarios. This, sadly, means that the symbolic rules extracted from the decision tree may not capture all possible interactions among attributes. To address this issue, an additional

training data can randomly be generated; however such data may not precisely represent the true nature of the original data. Creating realistic examples is a challenging task, especially in problem domains where little is known.

We believe that the RDC-ANNE approach can be useful in domains like medicine and biology where many incomplete and noisy examples are available, and yet the problem domain is still not well understood. The results produced by this approach represent a significant improvement in terms of comprehensibility of extracted rules from the trained ANN ensemble. The RDC-ANNE model can help in determining the possible interactions between the attributes and can be used to form new concepts.

Future work in this direction will look at designing a more efficient pre-processing technique that can guarantee preservation of more data for the rule extraction process.

References

1. Lisboa, P.J.G.: A review of evidence of health benefit from artificial neural networks in medical intervention. Neural Networks, Vol.15, No.1, January (2002)11-39
2. Murthy, S. K.: Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey. Data Mining and Knowledge Discovery, Vol. 2, No. 4, (1998) 345-389
3. Tickle, A., Orlowski M., and Diederich J.: DEDEC: Decision Detection by Rule Extraction from Neural Networks. Technical Report NRC QUT, Queensland University, September (1994)
4. Taha, I. A., Ghosh, J.: Symbolic Interpretation of Artificial Neural Networks. IEEE Trans. Knowl. Data Eng., Vol. 11, Issue 3, (1999) 448-463
5. Zhou, Z., Jiang, Y., Chen, S.: Extracting symbolic rules from trained neural network ensembles. AI Commun. Vol.16, No. 1, (2003) 3-15
6. Andrews, R., Diederich, J., Tickle, A.: Survey and critique of techniques for extracting rules from trained artificial neural networks. Knowledge-Based Systems, Vol.8, No.6, (1995) 373-389
7. Doyle H., et al.: Predicting outcomes after liver transplantation. A connectionist approach. Ann. Surg. Vol. 219, Issue 4, (1994) 408-415
8. Matis, S., Doyle, H., Marino, I., Murad, R, Uberbacher, E.: Use of Neural Networks for Prediction of Graft Failure following Liver Transplantation. Proceeding of the Eight Annual IEEE Symposium on Computer-Based Medical Systems (1995) 0133
9. Data Dictionary: ANZDATA Registry Database (2000)
10. Shadabi, F., Cox, R., Sharma, D., Petrovsky, N.: Use of Artificial Neural Networks in the Prediction of Kidney Transplant Outcomes. KES'04, the 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems (2004)
11. Shadabi, F., Cox, R., Sharma, D., Petrovsky, N.: Experiments with a Neural Network Ensemble to Predict Renal Transplantation Outcomes. AISAT 04 Proceeding, the 2nd International Conference on Artificial Intelligence in Science and Technology, November (2004) 271-276
12. 12. Witten, I.H. and Frank, E.: Data mining: Practical machine learning tools   and techniques   with Java implementations. Morgan Kaufmann, San Francisco (1999)
13. Brieman, L.: Bagging predictors. Machine Learning Vol. 24 Issue 2, (1996) 123-140
14. Mitchell, T. M.: *Machine Learning*. McGraw-Hill, New York (1997)
15. Quinlan, J. R.: C4.5: programs for machine learning. Morgan Kaufmann Publishers Inc., San Francisco, CA (1993)
16. Frank, E., Witten, I. H.: Generating Accurate Rule Sets Without Global Optimization. In Shavlik, J., ed., Machine Learning: Proceedings of the Fifteenth International Conference, Morgan Kaufmann Publishers (1998)

Performance Comparison for MLP Networks Using Various Back Propagation Algorithms for Breast Cancer Diagnosis

S. Esugasini[1], Mohd Yusoff Mashor[1], Nor Ashidi Mat Isa[1], and Nor Hayati Othman[2]

[1] Control and Electronic Intelligent System (CELIS) Research Group
School of Electrical & Electronic Engineering, Universiti Sains Malaysia
Engineering Campus, 14300 Nibong Tebal, Penang Malaysia
esu@hotmail.com, {yusof,ashidi}@eng.usm.my
[2] School Medical Sciences, Universiti Sains Malaysia, Health Campus
16150 Kubang Kerian, Kelantan, Malaysia

Abstract. This paper represents the performance comparison of the Multilayered Perceptron (MLP) networks using various back propagation (BP) algorithms for breast cancer diagnosis. The training algorithms used are gradient descent with momentum and adaptive learning, resilient back propagation, Quasi-Newton and Levenberg-Marquardt. The performances of these four algorithms are compared with the standard steepest descent back propagation algorithm. The current study investigates and compares the accuracy, sensitivity, specificity, false negative and false positive results of the selected four algorithms to train MLP networks. The Papinicolou image of breast cancer cells were captured via an image analyzer and thirteen morphological features were extracted to numerical scores. The feature scores are used as data sets to train the MLP network. The MLP network using the Levenberg-Marquardt algorithm displays the best performance for all the five measurement criteria's (accuracy, specificity, sensitivity, true positive and true negative) at a lower number of hidden nodes.

1 Introduction

A total of 26,089 patients were diagnosed with cancer disease among all residents in Peninsular Malaysia, in the year 2002, comprising 54% of the amount to be female. The crude rate in the same year was 148.4 per 100,000 populations for females, with the breast cancer being the number one killer. Breast cancer accounted 30.4% of newly diagnosed cancer cases in Malaysian women, with the possibility of 1 of every 19 woman in Malaysia has the risk to develop breast cancer in their lifetime [1]. Several risk factors have been identified by researchers over the year. Age and genetic are the major factors related to breast cancer. The women at risk age vary from 35 to 55 and women with first degree relatives with the history of breast cancer are at higher risk than the rest of the population. An organized mass screening is not yet practiced in Malaysia although the number of cases is increasing each year. The key factor for this situation is the number of specialist (pathology, etc) required for mass screening is very less in Malaysia. Malaysian Ministry of Health, 1999 states that the number of specialist to population ratio is about 1:20 000 in Malaysia [2].

As it is, researchers have studied many alternatives to assist the pathologists to conduct the diagnosis of breast cancer faster and more accurately. The application of neural networks in medical field seems to be a promising technique in developing intelligent systems to assist the medical practitioners in decision-making. However, the contribution of neural network in the diagnosis of breast cancer via FNA images is still very slow. Particularly, works on neural networks implementation in cancer diagnosis have been done by [3], [4] and [5] for cervical cancer and breast cancer by [6] and [7]. Yao & Liu [6] defined two neural network approaches for breast cancer diagnosis, the evolutionary approach and the ensemble approach. The evolutionary approach is used to design compact neural networks automatically by evolving network architectures and weights, while the ensemble approach is aimed at tackling large problems which may not dealt with efficiently by a monolithic neural network. Kates *et. al* [7] presents the potential contributions of neural network to a clinical decision support framework for the prediction of breast cancer therapy response.

2   Multilayered Perceptron Network (MLP)

MLP networks are feed forward neural networks with one or more hidden layers. Cybenko [8] and Funahashi [9] have proved that the MLP network is a general function approximator and that one hidden layer networks will always be sufficient to approximate any continuous function up to certain accuracy. A MLP network with two hidden layers is shown in Figure 1. The input layer acts as an input data holder that distributes the input to the first hidden layer. The outputs from the first hidden layer then become the inputs to the second layer and so on. The last layer acts as the network output layer. A hidden neuron performs two functions that are the combining function and the activation function. The output of the *j*-th neuron of the *k*-th hidden layer is given by

$$v_j^k(t) = F\left(\sum_{i=1}^{n_{k-1}} w_{ij}^k v_i^{k-1}(t) + b_j^k\right) \qquad \text{for } 1 \le j \le n_k \qquad (1)$$

and if the *m*-th layer is the output layer then the output of the *l*-th neuron $\hat{y}_l$ of the output layer is given by

$$\hat{y}_l(t) = \sum_{i=1}^{n_{m-1}} w_{ij}^m v_i^{m-1}(t) \qquad \text{for } 1 \le l \le n_o \qquad (2)$$

where n_k, n_o, w's, b's and $F(.)$ are the number of neurons in *k*-th layer, number of neurons in output layer, weights, thresholds and an activation function respectively. $F(.)$ is an activation function that is normally selected as sigmoidal function.

In the current study, the network with four output nodes and a single hidden layer is used. With these simplifications the network output for each output node are as follows:

$$\hat{y}_l(t) = \sum_{i=1}^{n_1} w_i^2 v_i^1(t) = \sum_{i=1}^{n_1} w_i^2 F\left(\sum_{j=1}^{n_r} w_{ij}^1 v_j(t) + b_j^1\right) \text{for } 1 < l < n_o \qquad (3)$$
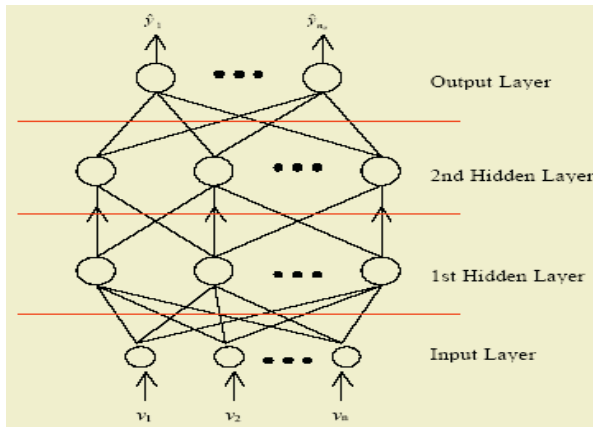
Fig. 1. Multilayered Perceptron networks

Where $m = 2$ and $n_o = 4$. The weights w_i and threshold b_j are unknown and should be selected to minimize the prediction errors defined as,

$$\varepsilon(t) = y(t) - \hat{y}(t) \tag{4}$$

where $y(t)$ is the actual output and $\hat{y}(t)$ is the network output.

3   Multilayer Perceptron Based Network Models for Diagnosis System

Four models of MLP network were used to compare the diagnosis performance of breast cancer cells. The gradient descent with momentum and adaptive learning (GDX) algorithm based MLP model refers to papers by [10] and [11]. The second and third network models are trained using the resilient back propagation (RPROP) algorithm based on [12], and Quasi-Newton (QN) algorithm by [13] respectively. While the Levenberg-Marquardt (LM) algorithm based network model is based on paper by [14].

The inputs to the MLP network are 13 features of the breast cells, which are cellularity, background, cohesiveness, cell in cluster, significant stromal component, clump thickness nuclear membrane, bare nuclei, normal nucleoli, mitosis, nucleus stain, uniformity of cell and fragility. The networks will receive these features as the inputs and classify the cells into four categories which are fibroadenoma, fibrocystic disease, other benign and malignant cells, as shown in Figures 2(a), 2(b), 2(c) and 2(d) respectively. The FNA image of the Pap smear slides are captured via an image analyzer and stored. The features of the samples are extracted and interpreted into numerical data sets by experienced cytotechnologists, as shown partly in Table 1, with a close supervision of experienced pathologists.

4   Methodology and Data

The performance of the MLP networks trained using the gradient descent with momentum; adaptive learning rate algorithm, resilient back propagation algorithm,

Quasi-Newton algorithm and Levenberg-Marquardt algorithm presented in the previous section were compared. A data set consists of 330 samples were used for this comparison. The number of each case consists in the samples are described in Table 2, below:

(a) Fibroadenoma

(b) Fibrocyctic

(c) Other benign

(d) Malignant
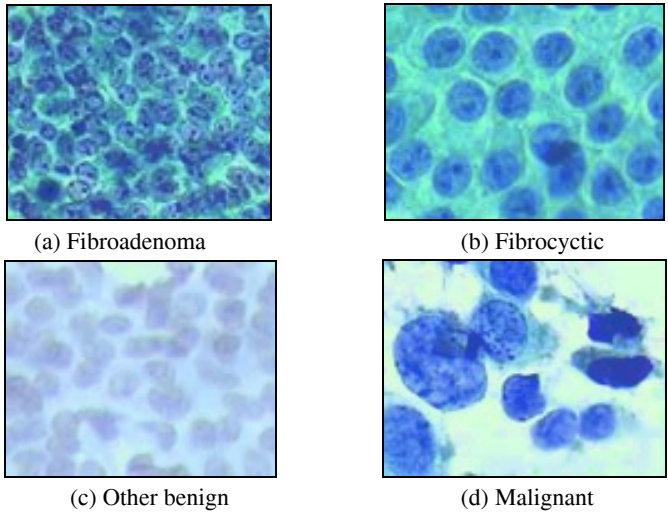
Fig. 2. Four categories of Breast Cell

Table 1. Scoring for breast cell features

Input Marker	Categories	NN Inputs
Cellularity	Scanty	1
	Moderate	2
	High	3
Background	Clean	1
	Slightly Dirty	2
	Dirty	3

Table 2. The number of data's/samples for each type of breast tumour cells

Type of tumour cells	No. of Training Data	No. of Testing Data	Total
Fibroadenoma	60	40	100
Fibrocystic	60	40	100
Other Benign	60	40	100
Malignant	20	10	30

The comparisons are based on the accuracy, sensitivity, specificity, false positive and false negative of the proposed networks. Accuracy determines the total performance of the network model, similarly to the ability of the network to classify the type of breast tumour cells. In this context, tumour cells include fibroadenoma, fibrocystic, malignant and other benign cases. Thus, accuracy is defined as:

$$\% \text{ Accuracy} = \frac{\text{Total number of tumour cells classified correctly}}{\text{Total number of cells in the data set}} \times 100 \tag{5}$$

Sensitivity describes the percentage of correctly classified malignant cells as opposed to the total number of malignant cells in the data set, as described in the equation below:

$$\% \text{ Sensitivity} = \frac{\text{Number of Malignant cells classified correctly}}{\text{Total number of Malignant cells in the data set}} \times 100 \tag{6}$$

While specificity describes the percentage of correctly classified benign cells, which are divided into three categories fibroadenoma, fibrocystic and other benign, as described in the equation below:

$$\% \text{ Specificity} = \frac{\text{Number of Benign cells classified correctly}}{\text{Total number of Benign cells in the data set}} \times 100 \tag{7}$$

The false negative percentage describes the percentage of the benign cells which are wrongly classified as malignant, whilst the false positive percentage describes the percentage of benign cells which are classified as malignant cells. The false negative and false positive percentage is described in the equation below:

$$\% \text{ False Negative} = \frac{\text{Number of Malignant cells classified as Benign}}{\text{Total number of Malignant cells in the data set}} \times 100 \tag{8}$$

$$\% \text{ False Positive} = \frac{\text{Number of Benign cells classified as Malignant}}{\text{Total number of Benign cells in the data set}} \times 100 \tag{9}$$

The Pap images of the cancer cells were obtained from the Pathology Department of University Sains Malaysia Hospital, Kelantan. The data sets were divided to 200 samples for training and 130 samples for testing the network model. All the network models have the same structure with different training algorithms. The parameters for gradient descent with momentum and adaptive learning rate algorithm, resilient back propagation algorithm, Quasi-Newton algorithm and Levenberg-Marquardt algorithm were set as follows:

Epoch = 300 (for all network models)

Gradient descent with momentum and adaptive learning rate algorithm:
Momentum constant, $\alpha_w = \alpha_b = 0.85$ Learning rates, $\eta_w = \eta_b = 0.005$

Resilient back propagation algorithm:
Learning rates, $\eta_w = \eta_b = 0.005$ Initial weight change, $\Delta_0 = 0.07$

Quasi-Newton algorithm:
Initial step size, $\lambda = 0.01$ Sufficiently large step size, $\beta = 0.1$

Levenberg-Marquardt algorithm:
Initial value, $\beta = 0.1$

5   Results and Discussion

After all the networks were trained properly they were tested using the training and testing data sets. The accuracy, sensitivity, specificity, false negative and false positive performance of all the network models and their optimum hidden nodes are shown in Table 3 and Table 4. Table 5 is displays the performance of the standard gradient descent back propagation algorithm which is used as a referral in this study.

Table 3. Diagnostic performance for training data sets

Network Model	Quasi-Newton	GDX	RPROP	LM
Hidden Node	25	45	50	20
% Accuracy	74.50	71.50	76.50	80.50
% Sensitivity	100.00	100.00	100.00	100.00
% Specificity	61.94	59.29	66.43	70.90
% False Negative	0.00	0.00	0.00	0.00
% False Positive	38.06	40.71	33.57	29.10

Table 4. Diagnostic performance for testing data sets

Network Model	Quasi-Newton	GDX	RPROP	LM
Hidden Node	25	45	50	20
% Accuracy	72.31	69.23	73.08	80.00
% Sensitivity	100.00	100.00	100.00	100.00
% Specificity	60.87	54.02	59.77	71.74
% False Negative	0.00	0.00	0.00	0.00
% False Positive	39.13	45.98	40.23	28.26

Table 5. Diagnostic performance for training and testing data sets with same network structure

Network Model	Quasi-Newton		GDX		RPROP		LM	
Hidden Node	20		20		20		20	
Data sets	Train	Test	Train	Test	Train	Test	Train	Test
% Accuracy	75.50	73.08	67.50	64.62	69.50	66.15	80.50	80.00
% Sensitivity	100	100	100	100	100	100	100	100
% Specificity	63.43	61.96	51.49	50.00	54.48	52.17	70.90	71.74
%False Negative	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% False Positive	36.57	38.04	48.51	50.00	45.52	47.83	29.10	28.26

The training and testing network performance of the four network models shows a great number of variations in the parameter selection to obtain an optimum output for the breast cancer cells classification. The GDX and RPROP network requires a larger number of hidden nodes to train the network for an optimum performance. While, the Quasi-Newton and LM network model was able to converge by using a smaller amount of hidden node only. However, overall comparison shows that the LM network model withstands the other network models, by performing high accuracy of 80.5% for training network and 80% for testing network. The LM network was able to classify all the malignant cases (severe cases) and 71.74% of the benign cases. However the false positive percentage of the network has to be overlooked, since 28.26% of the benign cases are wrongly classified. Table 4 shows the result of the networks with the same structure; however this is not a fair comparison. Each of the networks has its optimum or best result at their respective network structure (number of hidden nodes) as described in Table 3 and Table 4. The results in Table 5 clearly conclude that the LM network has the highest ability to classify the tumour cells into four respective categories with lower number of hidden nodes.

Table 6. Diagnostic performance of the standard gradient descent back propagation

Network Model	Standard Gradient Descent			
	Training Data		Testing Data	
Hidden Node	100	150	100	150
% Accuracy	71.00	73.50	72.31	72.31
% Sensitivity	100.00	100.00	100.00	100.00
% Specificity	56.72	61.03	59.55	59.55
% False Negative	0.00	0.00	0.00	0.00
% False Positive	43.28	38.97	40.45	40.45

Table 6 displays the performance of the standard gradient descent back propaga-
tion algorithm. The standard gradient descent back propagation algorithm is used as
an indicator to compare with the GDX, RPROP, Quasi-Newton and LM networks.
The standard back propagation network could not converge when the training epoch
was set to 300 as the other networks. The output of the network at 300 epochs was
only 2.67% accuracy for training and 2.3% for testing. A large number of hidden
nodes are used to further improve the network performance. Thus, the network results
a very slow convergence rate. The gradient descent network is designed to start at
some point on the error function defined over the weights and attempt to move to the
global minimum of the function. However in real problems, as it is in this case study
the error surfaces are typically complex and usually the data's are trapped in local
minima.

6 Conclusion

The Multilayer Perceptron (MLP) was used to perform a comparison study using four
back propagation algorithms, which are gradient descent with momentum and adap-
tive learning rate algorithm, resilient back propagation algorithm, Quasi-Newton
algorithm and Levenberg-Marquardt algorithm. Hence, four network models were
picked in this study to resolve the best network model which can perform optimally
for the diagnosis of breast cancer cells. The standard gradient descent network model
was used as a referral network to determine the brilliance of the proposed four net-
work models. The proposed neural network models are static models, with 13 inputs
and 4 outputs. The standard network could result an optimum performance of
73.50% accuracy, 100% sensitivity, 61.03% specificity, 0% false negative and
38.97% false positive, for training only with epoch training of 50 000 and above. A
large number of hidden nodes were used to train standard network which results a
very slow convergence rate. Among the four selected networks for this comparison
study, the LM network model shows a good performance in all measures, as shown in
Table 2. The gradient descent with momentum and adaptive learning rate algorithm
based network produces a very low accuracy and specificity, but very high false posi-
tive result. The Quasi-Newton and RPROP algorithm models results a moderate per-
formance. The LM network model can perform an optimum results output at a much
lower number of hidden nodes or adjustable parameters. The entire network however
produced 100% sensitivity. Therefore, LM does not only could result a better per-
formance but it also more efficient compared to the other networks. As a conclusion
the LM network model is able to show the best performance for breast cancer cell

classification with optimum network parameters, when compared with the standard back propagation network and other back propagation algorithms.

References

1. Lim G.C.C., Yahaya H., & Lim T.O. The First Report of the National Cancer Registry: Cancer Incidence in Malaysia 2002. (2002).
2. MOH (Ministry of Health Malaysia). Malaysia's Health: Technical report of the director-general of health, Malaysia. (1999).
3. Kok M. R., Schreiner-Kok P.G., Veen V.D., & Boon M.E., 'Potentially difficult smears of women with squamous cell carcinoma pose fewer problems when PAPNET is used for primary screening', Cytopathology, **10**(5), pp. 324-334. (1999).
4. Mitra P., Mitra S & Pal S.K., 'Staging of Cervical Cancer with Soft Computing', IEEE Trans. On Biomedical Eng., **47**(7), pp. 934-940. (2000).
5. Mashor M.Y., Mat Isa N.A. & Othman N.H., 'Classification of Pap Smear Tests Using Neural Networks', IFMBE Proceedings: International Conference on Biomedical Engineering, **7,** pp. 21-24. (2004)
6. Yao, X. & Liu, Y., 'Neural Networks for Breast Cancer Diagnosis', IEEE Trans., pp. 1760-1767 (1999).
7. Kates, R., Harbeck, N. & Schmitt M., 'Prospects for clinical decision support in breast cancer based on neural network analysis of clinical survival data', 4th International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies, pp. 764-767 (2000).
8. Cybenko, G., 'Approximations by superposition of a sigmoidal function', Mathematics of Control, Signal and Systems, **2**, 303-314. (1989).
9. Funahashi, K., 'On the approximate realisation of continuous mappings by neural networks', Neural Networks, **2**, 183-192. (1989).
10. Plagianakos, V.P., Sotiropoulos, D.G. & Vrahatis, M. N. 'An Improved Backpropagation Method with Adaptive Learnaing Rate', Technical Report: Dept. of Mathematics, University of Patras, **2**, pp.1-8. (1998).
11. Boukis, C.G. & Papoulis, E.V. (2003): ' A normalised adaptive amplitude nonlinear gradient descent algorithm for system identification', IEEE Trans. ICECS, pp.1042-1045.
12. Mastorocostas, P.A., 'Resilient back propagation learning algorithm for recurrent fuzzy neural networks', Electronics Letters, **40**(1), pp. iii-iv. (2004).
13. Ouyang, S., Ching, P.C. & Lee, T., ' Robust adaptive quasi-Newton algorithms for eigen-subspace estimation', IEEE Proceedings-Vis. Image Signal Process, **150**(5), pp. 321-331. (2003).
14. Chen, T., Han, D., Au, F.T.K. & Tham, L.G., 'Accelaration of Levenberg-Marquardt Training of Neural Networks with Variable Decay Rate', IEEE Trans., pp. 1873-1878. (2003).

Combining Machine Learned and Heuristic Rules Using GRDR for Detection of Honeycombing in HRCT Lung Images

Pramod K. Singh and Paul Compton

School of Computer Science and Engineering
University of New South Wales
Sydney, NSW - 2052, Australia
{pksingh,compton}@cse.unsw.edu.au

Abstract. A knowledge based system for detection of honeycombing patterns in HRCT lung images is described. In the system, rules generated by machine learning on low level image pixel-based features and heuristic rules from the domain expert on high level region-based features are combined using a generalized ripple down rules (GRDR) framework. Results demonstrate that the systems' performance can be incrementally improved.

1 Introduction

Automatic pattern recognition (APR) in medical images is one of the key issues in the development of a computer aided diagnosis (CAD) system. Certain disease patterns in medical images can be detected via the application of well-known machine learning and/or low level image processing but some patterns are more easily resolved using heuristic knowledge. Generally, APR algorithms do not use heuristic knowledge about the domain, because the machine generated rules can not be combined easily with the heuristic rules. An APR system which integrates machine learning and heuristic knowledge can take advantage of both kind of knowledges.

The proposed method is tested on a real world problem of pattern recognition in high resolution computed tomography (HRCT) images. HRCT is a widely used technique for the assessment of lung diseases. Honeycombing is a lesion belonging to the diffuse lung disease category. Amongst the entire diffuse lung disease family, honeycombing is the most difficult pattern to detect automatically as shown by Uppaluri and colleagues [8]. Most of the reported honeycombing detection algorithms are based on either low level image processing [10] or machine learning [8].

A ripple down rules (RDR) framework has been used for integrating machine learning (ML) and knowledge acquisition (KA) techniques in several studies [5],[9]. However, most of these systems use a single knowledge base built by ML and KA together using the same feature space. Its obvious that a human expert provides his/her opinion in the form of heuristic rules on object/region

whereas pixel based analysis is done on spectral and texture features of the image. Hence, the feature spaces of both expert and data based rules in such problem domain are different. In this study we use generalized ripple down rules (GRDR) framework for combining low level machine learning and high level heuristic rules. For simplicity, we first present how the system is constructed and report results without reference to the GRDR framework. This description and the results achieved stand by themselves, but we believe there is considerable value in describing the system as an example of the GRDR framework. This follows earlier work on image processing which also links low level ML classifiers and heuristic rules in a simpler RDR structure [3].

2   Machine Learning Rules Induction

The ML techniques used in other systems [3][5][9] develop a discriminative description of classes which discriminates between instances of one class and other classes with an assumption that the instances of all possible classes of the domain are included in the training set, whereas this condition is seldom satisfied in our application domain. Under such situations a characteristic descriptor for a class should enable a system to identify all instances of that class and reject all instances of the other (disjoint) classes. In order to get characteristic descriptor for detection of honeycombing (HC) patterns we use an interval based characteristic concept learner, IC^2 [7], summarized as follows.

Data Preparation: The dataset for characteristic rule learning is generated from the HRCT lung images by applying statistical operators to discrete cosine transforms (DCT) of the labeled regions. The regions of the images are labeled by the radiologists as regions of interest (ROI) for the diagnosis of lung diseases. The feature generation technique is reported elsewhere [6].

Learning Algorithm: The IC^2 constructs rules using instances of only one class as described below. The algorithm consists of two main steps - rule generation and feature selection. In the first step a function called **Interval_Rules** creates concept rules, see Table 1. It works on single feature for simple rules and generates composite rules using multiple features.

A feature selection method based on a *wrapper model* is used in this algorithm. The function **Feature_Select** uses a goodness measure for selection of promising features, see Table 2. For measuring the goodness of the interval rules generated from a feature or a set of features we use two measures called *precision* and *recall* of the rules. *Precision* is the ratio of the number of target class examples correctly covered by the rules to the total number of examples covered. *Recall* of the rules is ratio of the number of target class examples covered by rules to the total number of target class examples. An accuracy factor *f_accuracy* measure combines *recall* and *precision* statistics using following fomula. A feature or feature set with a larger *f_accuracy* is preferred. Thus, *f_accuracy* is used to rank the features and to guide the specificity factor σ in the interval rules generation process.

$$f\_accuracy = \frac{\beta + 1}{\frac{\beta}{recall} + \frac{1}{precision}} \tag{1}$$

where β is a generalization parameter. $f\_accuracy$, with a value of β greater than 1, favors *recall* and more generalization of concepts, whereas with a value smaller than 1, it favors *precision* and more specific rules.
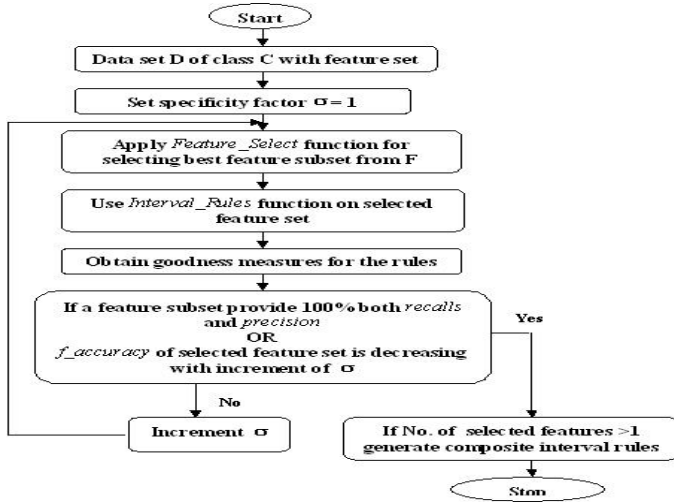
Fig. 1. IC^2 Algorithm

Here, the goodness measure needs negative examples for computation of the *precision* of rules. Based on Mitchell's density estimation concept [4] we synthetically generate uniformally distributed negative examples for the target class to be learned. For selecting the set of features we use the popular forward sequential selection (FSS) method in our algorithm. Fig.1 summarizes the IC^2. Due to space limitation we have omitted various details of the algorithm, for details readers are referred to [7].

3   Experiment

Experiments were conducted to investigate the effectiveness of the proposed method, we collected 29 labelled images of HC selected from 12 patients and 138 other images of normal and other tissue patterns selected from 52 patients. The labelling of images is done by a radiologist of our group. In a single image there may be one or more HC regions as shown in Fig.2(a). About 50% of the regions of the HC class are used for initial learning and all images including HC and other categories are used as test data.

Heuristic knowledge is acquired as an opinion of an expert on the results. An example of heuristic knowledge is: "honeycombing patterns appears only in

Table 1. Interval_Rules Function for Interval Rules Generation

Simple Interval_Rules	**Composite Interval_Rules**
Input: Training Dataset D, feature f.	**Input:** Training Dataset D, feature set F.
Output: Disjunct Interval Set.	**Output:** Cartesian Interval Set.
S_Interval Algorithm:	**Initialization:** Cart-intervals $= \phi$.
Get min and max values of f in D.	**C_Interval Algorithm:**
Compute $mean$, SD and no. of intervals k.	For each feature f $\in$ F do{
Compute lower and upper bounds for each interval k.	//operator X is cartesian product. Cart-interval = Cart-intervals X
Allocate instance $i \in D$ in one of the intervals k based on value of f.	S_interval(f).}
Merge adjacent intervals of non-zero coverage and define new bounds.	Allocate instance $i \in D$ to one of the rule $\in$ Cart-intrvals.
Delete intervals of zero coverage.	Remove rules $\in$ Cart-intervals have zero instance coverage.
Return Disjunct intervals.	**Return** Cart-intervals.

(a)   (b)   (c)
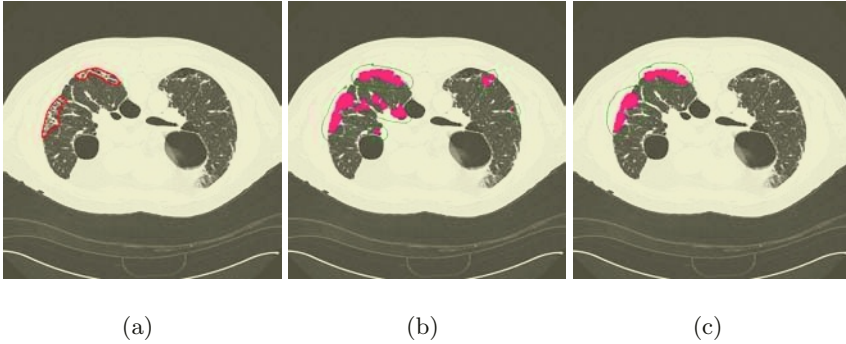
Fig. 2. Honey combing regions (a) labelled by radiologist;(b) detected by rules of conceptual layer; (c) validated by rules of heuristic layer

the peripheral regions of the lung". Such abstract knowledge is converted into rules as follows. Suppose ROI_1" is a detected HC region by ML rules; a typical heuristic rule is: if $Distance(ROI_1, \text{Boundary}) \leq \tau$ then ROI_1 is HC region". Where, $Distance$ is a function which produces a projected distance between the region and boundary in pixels, $Boundary$ is a peripheral position and τ is distance threshold in terms of pixels. In the present system we incrementally captured 13 set of heuristic rules based on experts' suggestions.

There are two separate knowledge bases systems(KBS): the ML and the heuristic system. In each the rules are linked in RDR fashion. The key feature of RDR is that the rules are evaluated in a strict order [1]. A rule that corrects an incorrect conclusion is linked as a refinement rule to the incorrect rule and is

Table 2. Feature_Select for Feature Selection

Feature_Select Function	Goodness of a feature
Input: Dataset D**, feature set** F**.**	**Input: Dataset** D**, feature set** F**.**
Output: Feature Subset S**.**	**Output: goodness measures.**
Initialization: $S = \phi$**.**	**Goodness Algorithm:**
Feature_Select Algorithm:	Generate uniformly distributed
// use forward sequential selection method.	negative dataset U for D.
While(*Recall* and *Precision*	Divide D evenly into n
both not 100% or *f_accuracy*	groups, $S_1, S_2, ..., S_n$.
increasing){	For i from 1 to n S_n and
For each $f \in F$ do {	U as testing data and rest
S = $S \cup f$.	of data as training data{
Calculate Goodness(S).	if F has only one element{
S = $S - f$.}	// generate simple interval rules.
Select best feature f.	Take feature f from F.
$S = S \cup f$.	enerate rules by S_Interval(f)
$F = F - f.$ }	using training data.}
Return S.	else{
	// generate composite interval rules.
	Generate rules by C_Interval(F)
	using training data.}
	Test the rules on the testing data.
	Count the total number of test
	examples correct and incorrect.}
	Calculate goodness measures:
	recall, *precision* and *f_accuracy*.
	Return goodness measures.

only evaluated when the parent rule fires. In our system, if a ML rule gives the wrong conclusion the expert labels the region and ML is applied to this data to generate a new rule which is attached as a refinement. Similar RDR refinements are made in the second system. RDR requires that a change should not affect previous saved cases. With two separate systems, any change made to the first system needs to be checked against the second.

3.1   Results and Discussion

Fig.2 shows one result of HC pattern detection in an image. The pattern detection results shown in Fig.2(b) and (c) clearly reveal the effect of heuristic rules in the detection of ROIs. In Fig.2(b) the number of detected honeycombing ROIs are 11 (small and big pink coloured regions), whereas there are only 2 ROIs resulted in Fig.2(c) by reducing number of false detected regions.

Fig.3 shows the acquisition of rules in machine and heuristic knowledge bases on incremental testing of cases. It clearly indicates that the errors in detection
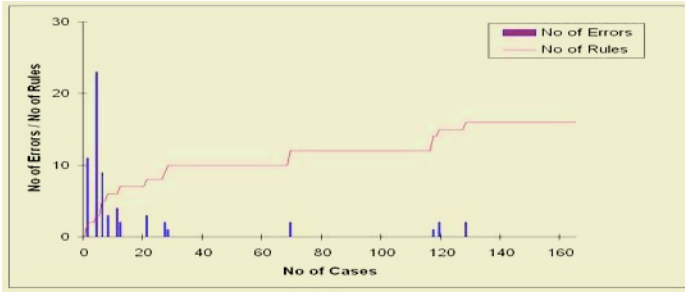
Fig. 3. Performance Graph of Integrated Rules in GRDR Framework

drop significantly after acquisition of some rules in the initial phase. The erroneous results, on some cases in the later stage, were fixed using exception rules. This shows that the system is able to incorporate ML and heuristic rules incremently. Although very small number of rules are involved, Fig.3 shows the standard RDR pattern of a decreasing rate of rule addition as domain coverage is achieved.

As mentioned above, in the system two separate ML and heuristic KBS are linked. In order to maintain the consistency of rules the linking between them facilitates cross checking. In our implementation the relations between these two KBS are defined based on the control mechanism of the GRDR [2]. In the GRDR control mechanism determines the sequence of evaluation of KBS is dependent on the sibling and correction relations among them. In sibling relations the output of first KBS must be passed through the second KBS with a restriction that the output of first KBS can only be supplemented with new results by second KBS. In correction relations the results of first KBS is passed through the second KBS and if the second KBS does not provide any output then the output from the first KBS is considered final, otherwise, output from the second KBS overrides the result of the first KBS.

In the GRDR structure used here, the errors in ML KBS are fixed by addition of new ML rules to KBS (sibling relations among ML rules). For these changes the performance of heuristic KBS must be checked against past cases and the only a ML rule which causes no change in heuristic KBS can be added. Otherwise rules in heuristic KBS can override conclusion from ML KBS as required (correction relations among ML and heuristic rules). But when these rules in heuristic KBS give incorrect conclusions they are refined (correction relations among heuristic rules). In the experiments here 2 sets of extra rules added for ML KBS and 13 heuristic rules captured in heuristic KBS.

4   Conclusion and Future Work

In the paper, a GRDR based multi-level knowledge based pattern detection system is presented which uses ML and heuristic rules to classify a complex pattern in HRCT images. The ML rules for pixel level classification are generated using

a characteristic concept learner on spectral and texture features of image patterns. The resulting regions from machine-level knowledge base are validated by heuristic rules acquired from the experts by inviting their opinions on classification results. In the present system heuristic rules are acquired with the help of a knowledge engineer because certain image processing procedures are needed for getting appropriate features to translate the heuristic knowledge into appropriate rules. Further research is needed to facilitate easy acquisition of heuristic rules, without involvement of knowledge engineer, in the system.

Acknowledgment

References

1. P. Compton and R. Jansen, *A Philosphical Basis for Knowledge Acquisition*, Knowledge Acquisition, vol.2, 1990, pp.241-257.
2. P. Compton, T.M. Cao and J. Kerr, *Generalizing Incremental Knowledge Acquisition*, Proceedings Pacific Knowledge Acquisition Workshop 2004, pp.44-53.
3. J. Kerr and P. Compton, *Toward Generic Model-Based Object Recognition by Knowledge Acquisition and Machine Learning*, Workshop on Mixed-Initiative Intelligent Systems, Int. Joint Conf. on AI, 2003.
4. A.R. Mitchell, *"Boosting" a Positive-Data-Only Learner*, Int. Conf. on Machine Learning, 2000, pp.607-714.
5. G.M. Shiraz and C. Sammut, *Combining Knowledge Acquisition and Machine Learning to Control Dynamic Systems*, Proc. of Int. Joint Conf. on AI, 1997, pp.908-913.
6. P.K. Singh, *Unsupervised Segmentation of HRCT Lung Images using FDK Clustering*, IEEE Int. Workshop on Biomedical Circuits and Systems, Singapore, 2004.
7. P.K. Singh, *IC^2: An Interval Based Characteristic Concept Learner*, Technical Report No. TR0442, School of Computer Science and Engineering, University of New South Wales, Australia, 2004.
8. R. Uppaluri, E.A. Hoffman, M. Sonka, P.G. Hartley, G.W. Hunninghake and G. McLennan, *Computer Recognition of Regional Lung Disease Patterns*, Am. J. Respir. Crit. Care Med., vol.160, no.2, 1999, pp.648-654.
9. T. Wada, T. Yoshida, H. Motoda and T. Washio, *Extension of the RDR Method That Can Adapt to Environmental Changes and Acquire Knowledge from Both Experts and Data*, Proc. of Pacific Rim Int. Conf. on AI, 2002, pp.218-227.
10. C.M.J. Wang, M. Rudrapatna and A. Sowmya, *Lung Disease Detection Using Frequency Spectrum Analysis*, Indian Conference on Computer Vision, Graphics and Image Processing, 2004.

Automatic Detection of Breast Tumours from Ultrasound Images Using the Modified Seed Based Region Growing Technique

Nor Ashidi Mat Isa, Shahrill Sabarudin,
Umi Kalthum Ngah, and Kamal Zuhairi Zamli

School of Electrical & Electronic Engineering, Universiti Sains Malaysia
Engineering Campus, 14300 Nibong Tebal, Penang, Malaysia
{ashidi,umi,eekamal}@eng.usm.my

Abstract. Past statistics have revealed that breast cancer is the world's leading cause of death among women. One popular method of screening breast cancer is ultrasound. However, reading an ultrasound image is not an easy task because it lacks spatial resolution, subject to image distortion, susceptible to noise and is highly operator dependant. Several image processing techniques have been introduced to enhance the detection of diagnostic features. This study proposes modified seed based region growing algorithm to detect the edges and segment the area of solid masses in an ultrasound image without having to specify the location of the seed and the grey level threshold value manually. Automatic seed selection is done by using moving *k*-means clustering. A performance analysis has been carried out towards 3 different ultrasound images. The results reveal that this algorithm can detect the edges of solid masses and segment it from the rest of the image effectively.

1 Introduction

An ultrasound image is an image of the inner breast obtained using sound wave emissions. High-frequency sound waves are passed through breast tissues and then converted into images on a viewing screen. Ultrasound complements other tests. If an abnormality is seen on mammography or felt by physical exam, ultrasound is the best way to find out if the abnormality is solid (such as a benign fibroadenoma or malignant lesion) or fluid-filled (such as a benign cyst). It cannot determine whether a solid lump is cancerous, nor can it detect calcifications. Ultrasound allows significant freedom in obtaining images of the breast from almost any orientation. Additionally, it can often quickly determine if a suspicious area is in fact a cyst or an increased density of solid tissue which may require a biopsy to determine if it is malignant.

Region Growing is an image processing technique that combines pixels into regions based on 2 criteria; proximity and homogeneity. This technique can be used for edge detection, features extraction and also image segmentation on a particular image. As of date, many robust and complex edge detection techniques have been proposed [1][2]. One of them is the Seed Based Region Growing (SBRG). The SBRG is an algorithm that utilizes this technique to detect edges and segment images into distinguishable features that are recognizable. This can be very useful especially to enhance diagnostic features in medical images.

However, the SBRG algorithm requires a location of a 'seed' point in an image and a grey level threshold value for region growing to take place, thus detecting the

edges of the image and for segmentation. The seed location is based on a priori information which requires some brief knowledge or experience of the image. These processes must be iterated until satisfactory results are obtained. Thus, this method is time consuming and is highly subjective and more dependable on the operator rather than the algorithm itself. The conventional SBRG algorithm finds the edges of the region of interest by using region growing concept from a seed point. Once the seed point does not fulfil certain conditions, the region growing will be stopped even though the whole image is not considered to be grown yet. Hence, the setback of the SBRG algorithm is the problem of a trapped seed point, which can cause incomplete edge detection process.

This study proposes a modified seed based region growing (MSBRG) as an automatic detection technique. The technique based on conventional algorithm with several enhancements will then be applied to ultrasound images as a case study for detecting solid mass edges.

2   The Modified Seed Based Region Growing Algorithm

In the proposed algorithm, a clustering algorithm is integrated together with MSBRG to automatically find the threshold value for classifying 2 cluster regions, i.e. the object of interest (the solid mass) and the background. Basically, clustering analysis uses non-parametric methods to find obvious regions on an image without locating a seed point. Clustering methods have been used throughout the years for several applications such as segmentation and for unsupervised learning [3].

Most of the previous studies on digital images use k-means and fuzzy c-means clustering algorithms as clustering techniques [4][5]. However, these clustering techniques do not always produce desirable results. In 2000, Mashor [6] proposed the moving k-means (MKM) clustering algorithm as a modification of the k- means clustering algorithm. A research that has been carried out by [5] towards pap smear image segmentation revealed that the MKM clustering algorithm provides several advantages compared to other clustering algorithms because MKM clustering can avoid the three most common problems associated with other clustering algorithms which are centre deaths, centre redundancy and centre trapped in local minima. For this reason, the moving k-means clustering algorithm is proposed in this study to find the clusters values.

Consider one ultrasound image with $N_b \times N_l$ pixels (where b and l are number of row and column of the image respectively) to be clustered into n_c regions. Let $p(x,y)$ is a pixel to be clustered and C_j is the j-th cluster (centre) ($x = 1, 2, …, N_b$, $y = 1, 2, …, N_l$ and $j = 1, 2, …, n_c$). Based on the considerations, the implementation of moving k-means clustering algorithm [6] to find the cluster values of the medical image is as follows:

1. Initialise the centres and α_0, and set $\alpha_a = \alpha_b = \alpha_0$ (where α_0 is a small constant value, $0 < \alpha_0 < \frac{1}{3}$ and should be chosen to be inversely proportional to the number of centres).

2. Assign all pixels to the nearest centre and calculate the centre positions using:

$$C_j = \frac{1}{n_j} \sum_{y \in c_j} \sum_{x \in c_j} p(x,y); \tag{1}$$

3. Check the fitness of each centre using equation:

$$f(C_j) = \sum_{y \in C_j} \sum_{x \in C_j} \left(\|p(x,y) - C_j\| \right)^2;$$

$\qquad$ (2)

where

$$j = 1,2,...,n_c$$
$$x = 1,2,..., N_r$$
$$y = 1,2,..., N_c$$

4. Find C_s and C_l, the centre that has the smallest and the largest value of $f(\bullet)$.
5. If $f(C_s) > \alpha_a f(C_l)$,

(5.1) Assign the members (pixels) of C_l to C_s if $p(x,y) < C_l$, where $x, y \in C_l$, and leave the rest of the members (pixels) to C_l.

(5.2) Recalculate the positions of C_s and C_l according to:

$$\left. \begin{array}{l} C_s = \dfrac{1}{n_s} \sum_{y \in c_s} \sum_{x \in c_s} p(x,y); \\[2mm] C_l = \dfrac{1}{n_l} \sum_{y \in c_l} \sum_{x \in c_l} p(x,y); \end{array} \right\}$$

$\qquad$ (3)

Note: C_s will give up its members (pixels) before step (5.1) and, n_s and n_l in Equation (3) are the number of the new members (pixels) of C_s and C_l respectively, after the reassigning process in step (5.1)

6. Update α_a according to $\alpha_a = \alpha_a - \alpha_a / n_c$ and repeat step (4) and (5) until $f(C_s) \geq \alpha_a f(C_l)$.
7. Reassign all pixels to the nearest centre and recalculate the centre positions using Equation (1).
8. Update α_a and α_b according to $\alpha_a = \alpha_0$ and $\alpha_b = \alpha_b - \alpha_b / n_c$ respectively, and repeat Step (3) to (7) until $f(C_s) \geq \alpha_b f(C_l)$.
9. Sort the centres in ascending order where $C_1 < C_2 < ..... < C_{n_c}$.

In this study, two cluster values, C_1 and C_2, were determined where each cluster value represents the tumour and background regions respectively.

3 Methodology for Automatic Edge Detection

Having found the threshold value from the MKM clustering algorithm, the MSBRG algorithm would then be applied on ultrasound images for 2 main purposes:

1. This algorithm should be able to detect the edges of the tumour and clearly outline the edges of the tumour so that the edges are distinguishable and easily seen from the whole image.
2. This algorithm should be able to segment the tumour from the rest of the image and clearly distinguish the tumour from the whole image effectively.

The algorithm of the proposed MSBRG can be implemented as follows [5]:

1. Implement three pre-processing algorithms to the image, namely median filter, histogram normalization and histogram equalization.
2. Choose $N \times N$ neighbourhood as shown in figure 1 (for $N = 5$). N must be an odd number equal to or higher than 3.
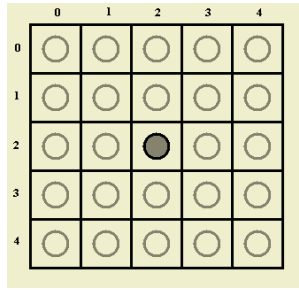
Fig. 1. Location of the seed pixel and it's 5 x 5 neighbourhood

3. Take the value of the tumour cluster, C_1 as a threshold value, β, which is determined by using the MKM clustering algorithm in the earlier process.
4. Examine all the pixels in the image. Set the first pixel with the grey level value higher than β as the initial seed location, $p_0(x,y)$. Make sure the initial seed location is located at the centre of all it's $N \times N$ neighbours.
5. Calculate the region meanl, $\bar{x}$ and the standard deviation, σ for $N \times N$ neighbourhood according to Equations (4) and (5) respectively.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \tag{4}$$

$$\sigma = +\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{5}$$

where x_i is the grey level value for the i-th pixel and n is the total of pixels in the image.
6. Compare each neighbour pixel with the initial seed pixel. Add a pixel to a region if it qualifies either one of the two conditions listed below [7][8]:
   a. If the gradient of the neighbour pixel is less than 95% of the equalized histogram and its grey level value is more or equal to the pre-selected threshold, β.
   b. If the gradient of the neighbour pixel is more than or equal to 95% of the equalized histogram and the grey level of the pixel is not more than or equal to one standard deviation away from the region mean.
7. Set the neighbour pixel, which is added to the region, as the new seed pixel.
8. Repeat step (5) to (7) until the region cannot be grown or all the pixels have been considered.
9. Change the grey level of the pixel that cannot be grown with the value of 0. Note: After this step, the tumour's edge will be detected as a black edge.
10. Set the next pixel with a grey level more than β as a new initial seed location, $p_0(x,y)$ if the pixel has not been grown yet.

4   Results and Discussions

The proposed MSBRG algorithm was applied on three ultrasound images. Each of which has its own distinctive cases and is chosen because they are the most common cases identified. The original images are given in Figure 2(a), 3(a) and 4(a). Figure 2(a) shows a homogenous, irregular mass which was palpable during the clinical physical exam. Applying the MSBRG algorithm towards the image reveals that the pixels within the growing region belongs to the same cluster family according to the MKM clustering algorithm which means that the density of mass in that particular area is the same and that the tumour detected is solid. The result is shown in figure 2(b) and 2(c) for tumour edge detection and tumour segmentation respectively.

The ultrasound in Figure 3(a) could be easily mistaken as another homogenous tumour. However, the edge detection and the tumour segmentation towards the image reveals that the density of the pixels within the area is not in the same cluster family, instead the pixels are mixed up between the pixels that belong to the tumour's cluster family and the pixels that belong to the background's cluster family. Hence, this shows that the tumour is a heterogeneous tumour and may comprise of frayed parts of fibro grandular tissues that are not connected to each other but clustered together in such a way that at a glance, they may seem to form a solid mass; hence may lead to a misleading diagnosis and improper treatment. The result of tumour edge detection and tumour segmentation is shown in Figure 3(b) and 3(c) respectively.

The ultrasound shown in Figure 4(a) shows an obvious breast lesion. Because shadowing was encountered, malignancy which is associated to the spread of cancer cells would be suspected. Traditionally, a radiologist will have to undergo several other tests and crucial verification before ultimately coming up with the correct diagnosis. Usually, such cases happen as an effect of incorrect degree of probe pressure or flattened tissues that could cause the anatomy to change the shape, location, affect the acoustic impedance, or change the appearance of the blood flow into superficial shadowing. After applying the MSBRG towards the ultrasound, the radiologist may get a clearer vision of the tumour where the acoustic shadowing is fairly dissipated.

Overall, the results obtained show that the proposed MSBRG algorithm which is powered by the integration of MKM clustering algorithm is capable of detecting edges of solid masses imaged by ultrasound images. For both edge detection and region segmentation, the result suits the desired outcome where the tumour is observed to be outlined perfectly and is distinguishable from the rest of the image.

5   Future Work

As the scope for future work, we are planning to evaluate the performance of the MSBRG algorithm further by comparing it with other similar algorithms. In particular, we are looking to assess our work against that of Cheng *et al* [9] whereby Fuzzy reasoning was used. Our preliminary assessment with three ultrasound images demonstrates promising results, although more assessments using a larger set of ultrasound images might be helpful to conclusively summarize the overall performance of the MSBRG algorithm.
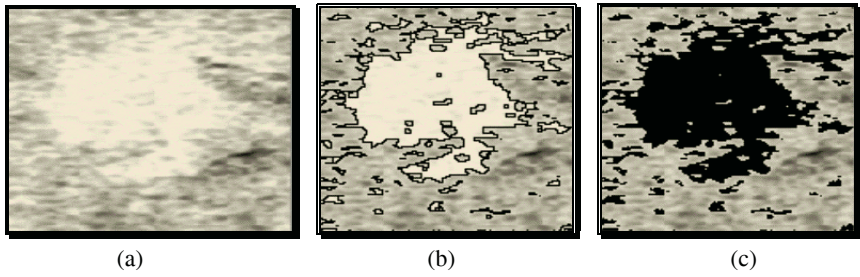
Fig. 2. (a) Original image of ultrasound A (b) Result for tumour edge detection (c) Result for tumour region segmentation
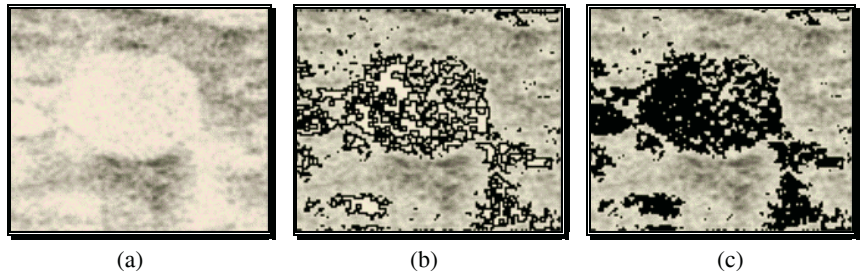
Fig. 3. (a) Original image of ultrasound B (b) Result for tumour edge detection (c) Result for tumour region segmentation
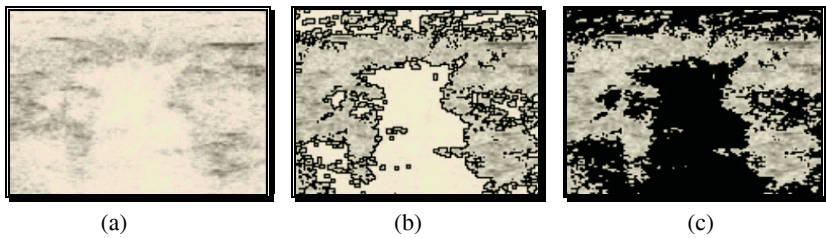
Fig. 4. (a) Original image of ultrasound C (b) Result for tumour edge detection (c) Result for tumour region segmentation

6   Conclusion

This study proposes an edge detection technique using the modified seed based region growing (MSBRG) algorithm that is powered automatically by moving k-means (MKM) clustering algorithm. Firstly, the MKM clustering algorithm is used to determine the clusters, which will then be used as thresholds. These threshold values will then be used in the proposed MSBRG algorithm to detect the edges of interested regions of digital images effectively. The suitability and capability of the proposed algorithm is tested on ultrasound images. The results show that the proposed technique has successfully detected the edges of the tumour automatically while maintaining the original size and shape of the tumour. This approach would prove valuable for

diagnostic purposes in assisting radiologists in interpreting an ultrasound image and eventually arrive at a correct diagnosis.

References

1. Fan, J. P., Yau, D. K. Y., Elgmagarmid, A. K. & Aref, W. G. Automatic Image Segmentation by Integrating Colour-Edge Extraction and Seeded Region Growing. *IEEE Transaction on Image Processing.* 10(10). (2001) 1454-1466.
2. Roghooputh, S. D. D. V. & Roghooputh, H. C. S. Image Segmentation of Living Cells. *Proceedings of International Conference on Robotics, Vision and Parallel Processing for Automation.* 1. (1999). 8-13
3. Jain, A. K., & Dubes, R. C. *Algorithms for Clustering Data*, Prentice Hall, Englewood Cliffs, New Jersey (1988).
4. Ghafar, R., Mashor, M. Y. & Othman, N. H. Segmentation of Pap Smear Slides Images Using K-Means Clustering. *Proc. of Kuala Lumpur Int. Conf. on Biomedical Engineering.* (2002). 41-43.
5. Mat-Isa, N. A., "Cervical Cancer Diagnosis System Based on Neural Networks" PhD Thesis, Universiti Sains Malaysia, (2003).
6. Mashor, M. Y. Hybrid Training Algorithm for RBF Network. *International Journal of The Computer, The Internet and Management.* 8(2). (2000) 50-65.
7. Romberg, J., Akram, W. & Gamiz, J., (1997), Image Segmentation Using Region Growing http://www.owlnet.rice.edu/elec539/projects97/WDEKnow/inde(2002)
8. Ngah, U. K., Ooi, T. H., Sulaiman, S. N. & Venkatachalam, P. A. Embedded Enhancement Image Processing Techniques on a demarcated Seed Based Grown Region. *Proceedings of Kuala Lumpur International Conference on Biomedical Engineering*. (2002) 170-172.
9. Cheng, X. Y., Akiyama, I., Itoh, K., Wang, Y., Tananiguchi, N., & Nakajima, M. Automated Detection of Breast Tumours in Ultrasonic Images. *Proceedings of IEEE International Conference on Image Processing (ICIP' 97)*, (1997). 420-423.

An Automatic Body ROI Determination
for 3D Visualization
of a Fetal Ultrasound Volume

Tien Dung Nguyen[1], Sang Hyun Kim[2], and Nam Chul Kim[1]

[1] Laboratory for Visual Communications, Department of Electronic Engineering
Kyungpook National University, Daegu, 702-701 Korea
[2] Department of Multimedia Engineering, Youngsan University
Yangsan, 626-847 Korea

Abstract. This paper presents an efficient method to determine a body region-of-interest (ROI) enclosing a fetus in a two-dimensional (2D) key frame that removes some irrelevant matters such as the abdominal area in front of a fetus to visualize a fetal ultrasound volume along with the key frame. In the body ROI determination, a clear frontal view of a fetus lying down floating in amniotic fluid mainly depends on the successful determination of the top bound among the four bounds of an ROI. The key idea of our top-bound setting is to locate it in amniotic fluid areas between a fetus and its mother's abdomen, which are dark so as to typically induce local minima of the vertical projection of a key frame. The support vector machines (SVM) classifier, known as an effective tool for classification, tests whether the candidate top bound, located at each of the local minima which are sorted in an increasing order, is valid or not. The test, using textural characterization of neighbor regions around each candidate bound, determines the first valid one as the top bound. The body ROI determination rate as well as resulting 3D images demonstrate that our system could replace a user in allocation of a fetus for its 3D visualization.

Keywords: fetal ultrasound volume, visualization, region-of-interest, significant local minimum, support vector machines.

1   Introduction

Recently, modern three-dimensional (3D) obstetrical ultrasound imaging systems have been capable of generating a transparent view that depicts a sculpture-like reconstruction of a fetus based on 3D visualization, which has a significant impact, especially on diagnostic applications, in the obstetrical field. Typical 3D fetal ultrasound imaging systems allow users to define ROIs manually on 2D slices from a volume using some editing tools and then concatenate these regions to form a volume-of-interest (VOI) [1]. Since a 3D fetal volume generally contains a fetus including the wall of the uterus and floating matters in the amniotic fluid areas, some means of separating the fetus from the surrounding

data is necessary to obtain an uncovered view of the fetus while reducing the volume size for manipulation of 3D images. A manual determination of the size and position of an ROI enclosing an object to be visualized might be tedious and time-consuming tasks even for obstetrical experts, especially for visualization of a fetus directly from an acquired volume data in recent medical imaging systems.

In our previous work [2], an algorithm for setting a face ROI enclosing a fetal face in a predefined body ROI is presented where the SVM using facial shape features tests which is valid among the two candidate ROIs left and right around a fetal neck. Since the accuracy of a face ROI setting depends on how accurately its body ROI is determined, we further develop an efficient algorithm for automatic determination of a body ROI enclosing a whole fetus for visualizing and manipulating all its available anatomical organs. In this body ROI determination, since the top bound is most important in achieving a clear frontal view of a fetus, the SVM, using texture features, tests whether candidate top bounds of a body ROI are in amniotic areas or not. The candidate top bounds are given at the local minima of the vertical projection of a key frame, which are induced by amniotic fluid areas. Among them sorted in an increasing order, the SVM selects the first valid candidate for the body ROI top bound. The left and right bounds are found in the most left and right boundaries of bright regions in the binarized key frame, which typically corresponds to bones of a fetus reflecting ultrasound signals. The bottom bound is simply determined to the first local minimum of the vertical projection of a key frame corresponding to an amniotic area just under the fetus. In order to investigate whether each candidate top bound is between a fetus and its mother's abdomen or not, the dissimilarity in textural activities of two neighbors around the bound is measured by BDIP, BVLC moments [3], and DLPV (difference of local projection values). The reason why the SVM approach is selected for finding a valid top bound is because it can effectively overcome the difficult problem of searching for a general model to fit databases such as fetal volumes which have variation of amniotic fluid levels according to gestation [4]. For real-time visualization of a 3D fetus, its VOI is then simply created by extending the body ROI of the key frame centered in its volume to all frames and rendered to show a fetus surface in the frontal view.

2   Proposed Body ROI Determination

Fig. 1 shows the block diagram of a 3D fetal ultrasound visualization system with the proposed body ROI determination. In the proposed method, the 2D frame centered in a 3D fetal volume is first extracted as a key frame for body ROI determination. Fig. 2a shows the 2D key frame and Fig. 2b shows its vertical projection curve that is obtained as

$$\mu_i = \frac{1}{W} \sum_{j=0}^{W-1} I_{ij}, \quad 0 \leq i \leq H-1 \tag{1}$$

where I_{ij} is the pixel intensity of the frame of size $H \times W$. As shown in Fig. 2b, the bright area of the mother's abdomen induces the first local maximum
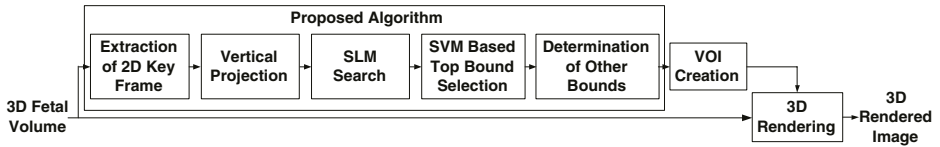
Fig. 1. Block diagram of a 3D fetal ultrasound visualization system with the proposed body ROI determination.

(a)   (b)   (c)   (d)
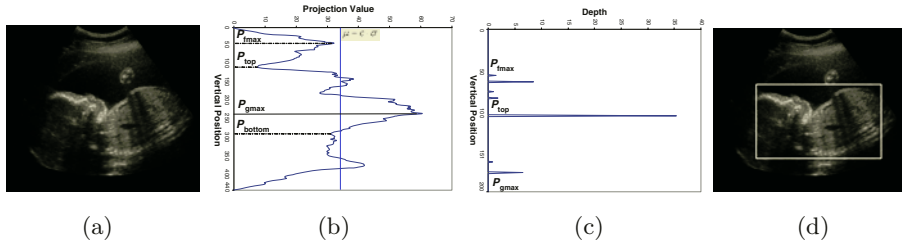
Fig. 2. Intermediate results of the proposed body ROI determination. (a) 2D key frame. (b) Vertical projection curve of (a). (c) Depths of SLMs calculated on the interval $[P_{fmax}, P_{gmax}]$. (d) Body ROI determined finally.

of the projection curve (at position P_{fmax}) and the very high intensity area of the fetus induces the global maximum (at position P_{gmax}). Between these two maxima, there typically exist local minima corresponding to the amniotic fluid areas with very low intensities. Among these local minima, significant ones considered as dark and deep valleys are searched by an algorithm that is called here significant local minimum (SLM) search. This is for reducing the influence of speckle noise and some irrelevant matters in dark amniotic areas on top bound determination. That is, a SLM is defined as the one whose elevation (projection value) is less than a threshold and where the minimum of its relative elevations of two adjacent maxima is less than a value that is defined experimentally. Fig. 2c shows the depths of significant local minima which are computed as

$$\Delta_k = \frac{1}{2} \left[(M_{k-1} - m_k) + (M_k - m_k) \right] \tag{2}$$

where m_k denotes the projection value at the kth SLM, M_{k-1} the projection value at the largest maximum between the $(k-1)$th and the kth SLM, and M_k the projection value at the largest maximum between the kth and the $(k+1)$th local minima. The SVM tests whether each SLM sorted in an increasing order of depth is valid or not using the features extracted at the SLM, which will be presented in Chapter III. Finally the top bound (denoted as P_{top}) is located at the first valid SLM.

The bottom bound of the ROI, denoted as P_{Bottom} in Fig. 2b, is simply determined as the first local minimum on the interval $[P_{gmax}, H]$. The key frame is next binarized by a recursive Otsu threshold [5] and the left and right bounds

of the ROI are determined in the most left and right boundaries of bright regions that typically correspond to bones of a fetus.

3   SVM Based Top Bound Selection

The SVM have been recently proposed as a new technique for patterns recognition. Intuitively, given a set of pattern or feature points which belongs to either of two classes, the SVM finds the hyperplane leaving the largest possible fraction of feature points of the same class on the same side, while maximizing the distance of either class from the hyperplane [4]. Prior to starting the classification, the SVM classifier is trained from learning feature examples in a training set. As a result, a feature example outside the training set can be classified well by the SVM classifier. Whereas linearly separable data may be analyzed with a hyperplane, linearly non-separable data, which are usually encountered in practical classification problems, are analyzed with an appropriate nonlinear operator for mapping input feature points into a higher dimensional feature space by using some proper kernels. Based on our test with all available kernels, a nonlinear SVM classifier with Gaussian RBF kernel is selected to classify the SLMs from which candidate ROI top bounds are given.

3.1   Feature Extraction for SVM

a. DLPV. As depicted in Fig. 3, the DLPV is defined as the averaged relative elevation of the projection curve around a candidate top bound, which is expressed as

$$\text{DLPV} = \frac{\sum\limits_{l \in [-\Delta, +\Delta]} [p(P_{top}) - p(l + P_{top})]}{2\Delta} \tag{3}$$

where $p(l + P_{top})$ are projection values in the interval $[p(P_{top}) - \Delta, p(P_{top}) + \Delta]$. Generally for a valid top bound, DLPV takes a large negative value due to a SLM as shown in Fig. 3a. However for an invalid top bound, DLPV takes a positive or small negative value due to a local minimum of small depth as shown in Fig. 3b. As a result the less the DLPV is, the deeper the valley is.

b. Differences of BDIP and BVLC Moments. Another feature in SVM is based on BDIP and BVLC moments [3]. While BDIP measuring the variation of intensities in an image block, BVLC measures the degree of roughness in an image block that is defined as the variation of local correlation coefficients according to four orientations $(-90°, 0°, -45°, 45°)$. Fig. 4 shows the BVLC and BDIP values computed in the two neighbors around the top bound in the key frame. Since BDIP and BVLC features are utilized to characterize the activity of these neighbors also, the height of the two neighbors in this calculation takes the same value of as that in Eq. 3.

Each of the two neighbors is divided into non-overlapping blocks of size $B \times B$, and then the BDIP and BVLC are computed in each block. In order to reflect
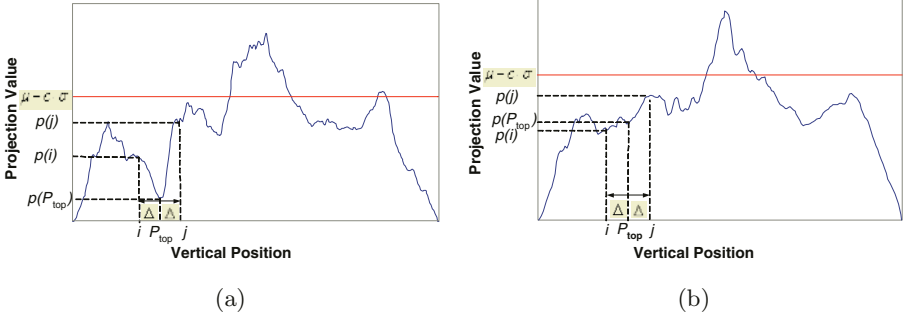
Fig. 3. Computation of DLPV on interval of size 2Δ in case of (a) valid and (b) invalid top bound.
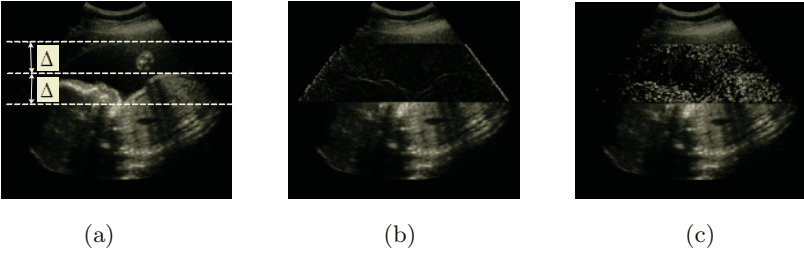
Fig. 4. BDIP and BVLC values computed in the two neighbor regions around a top bound in a 2D key frame. (a) Original frame. (b) BDIP values. (c) BVLC values.

the activity characteristics in the proposed feature vector, all blocks are classified into four classes. All the blocks are first classified into two groups by using the average of BDIPs over all blocks as a threshold. Next all blocks in each group are classified again into two subgroups by using the average of BDIPs over all blocks in each group as a threshold. The first moments of BDIPs are then computed for each class. Similarly, the first moments of BVLCs are computed for each class; however, each class is classified by using the average of BVLCs as a threshold. The regions outside the fan-shape areas, which generate BDIP and BVLC values of zero, are excluded from the calculation of BDIP and BVLC thresholds. These moments are combined to form a vector $\mathbf{h}$, which can be written in general case as

$$\mathbf{h} = [\mu_1(D),\ \mu_2(D), \cdots,\ \mu_n(D),\ \mu_1(V),\ \mu_2(V), \cdots,\ \mu_n(V)] \qquad (4)$$

Here $\mu_i(D)$ and $\mu_i(V)$ denote the BDIP and BVLC means for the ith class, respectively. Let $\mathbf{h}^{Up}$ and $\mathbf{h}^{Down}$ represent the vectors, which are formed by Eq. 4 for the two neighbors above and below the top bound, respectively. We define the difference vector $\mathbf{E}$ as

$$\mathbf{E} = \frac{\mathbf{h}^{Up}}{\sigma^{Up}} - \frac{\mathbf{h}^{Down}}{\sigma^{Down}} \qquad (5)$$

where the division means a component-wise division and $\boldsymbol{\sigma}^{Up}$ and $\boldsymbol{\sigma}^{Down}$ stand for vectors of standard deviations, which are calculated from the training DB, for each component in $\mathbf{h}^{Up}$ and $\mathbf{h}^{Down}$, respectively. These vectors are used as normalization factors to prevent feature vector components of great variances from dominating in the SVM calculation. The difference vector $\mathbf{E}$, which represents the block activities for each class in the two neighboring areas is then associated with the DLPV to form the $(2n+1)$ dimensional feature vector $\mathbf{x}$ for the SVM, which takes a form of

$$\mathbf{x} = [\mathrm{DLPV}/\sigma_{\mathrm{DLPV}}, \mathbf{E}] \tag{6}$$

where σ_{DLPV} denotes the standard deviation of the DLPV computed from the training DB for the same reason. This vector is used as an input of the SVM classifier to classify a given top bound into valid or invalid one. The significant difference of feature vector components calculated from the training DB for the two groups of top bounds implies that the proposed feature vector is effective in the classification.

3.2   Top Bound Selection

If the candidate top bound is classified by the SVM as invalid, the next candidate one is set up by successive replacement of the current candidate with a new candidate according to the increasing order of their depth (Eq. 2), and the feature vector is computed again for the new candidate. The replacement stops when the validity of this new candidate is reached. Occasionally, if no new valid candidate is found, the actual top bound is set to the one which generates the largest value of the decision function.

4   Experimental Results

4.1   Preliminary for Experiments

A set of 116 real fetal volumes at different gestation stages, which has been collected by Medison Co., Ltd Korea, is used to evaluate the performance of our automatic body ROI determination algorithm applied in a visualization system. The five frames of size 440×512 around the center frame sliced in the coronal view of each fetal volume are extracted and the set of all 580 frames extracted is arbitrarily divided into two subsets, each of which consists of 290 frames. One of the two subsets is used as a training set in training the SVM classifier and the other one as a test set in evaluating the performance of the proposed algorithm. In each of 150 frames chosen at random from the training set, a valid top bound is manually set to a local minimum with the largest depth, which is close enough for the frontal view of a fetus so that fluids and matters are completely removed. In each of the other 140 frames, an invalid top bound is set to one of the local minima of its projection curve except the deepest one.

The threshold for elevation of SLMs is adaptively computed from the mean μ and standard deviation σ of the key frame as shown in Fig. 3. In the calculation of BDIP and BVLC, the block size is chosen as 2×2. The height of the two neighbors around a candidate top bound is chosen as the average of amniotic fluid levels in the data set, $\Delta = 45$. As selecting the SVM classifier with Gaussian RBF kernel, a grid search for its best parameters of width σ and regularization factor C is performed by a five-fold cross-validation procedure on the training set. With the best cross-validation accuracy, the SVM classifier with $\sigma = 4$ and $C = 10$ is configured for classification of candidate top bounds. The number of classes for block classification in the computation of BDIP and BVLC moments is determined as 4 because there is no significant improvement for further increase in the number of classes.

4.2   Performance Evaluation

The receiver operating characteristic (ROC) analysis is used for providing a comprehensive summary of the trade-off between the sensitivity and the specificity of the SVMs. To investigate the contribution of DLPV, BDIP and BVLC moments to the role of the proposed feature vector, the ROC curve of the SVM trained by using this vector is compared with those of the SVMs trained by using the other combinations of DLPV, BDIP and BVLC moments. As shown in Fig. 5a, the participation of all components yields the highest SVM performance (the area under the ROI curve $A_z = 0.952$) over the remaining combinations. The SVM performance comparison in terms of accuracy, sensitivity, and specificity in Fig. 5b shows the effectiveness of the proposed feature vector in classification process. Regarding to insights of the SVMs, the proposed feature vector also leads to the smallest number of support vectors (about 28.3% of the trained samples).

The performance of the SVM in the proposed body ROI determination for a test set of 290 frames is analyzed in Tab. 1. The accuracy of the top bounds given at the first SLMs is 241 out of 290 (about 83.1%). Among these SLMs, 235 are

(a)   (b)
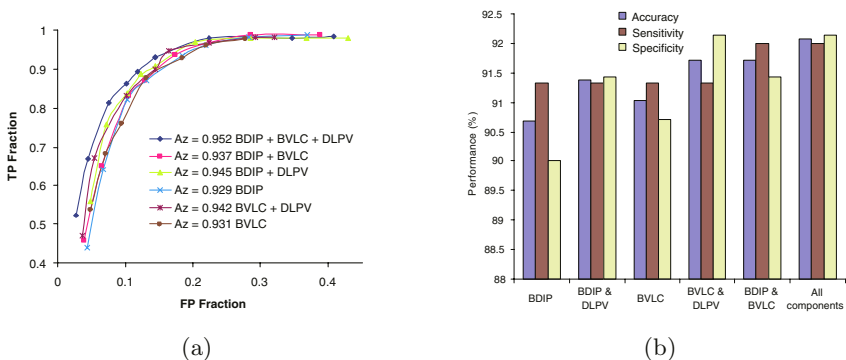
Fig. 5. Comparison of (a) ROC curves and (b) performance of SVM classifiers with various combinations of DLPV, BDIP and BVLC moments.

correctly classified as valid (about 97.51%) and 6 are classified as invalid (about 2.49%). Among the other 49 incorrect SLMs, 42 that are classified as invalid are to be successively replaced with a next SLM for SVM test and finally 35 out of 42 are correctly determined, which leads to the total determination rate of 270/290 (about 93.1%). That means the performance is improved by about 10% compared to the case without introducing the SVM based testing.

Table 1. Analysis of performance of the SVM in the proposed ROI determination for a test set of 290 frames.

Accuracy of first SLMs		Classification results for first SLMs being correct				Determination rate	
		Valid		Invalid			
241/290	83.1%	235/241	97.51%	6/241	2.49%	270/290	93.1%

4.3   Visualization

In SONOACE 9000, a 3D imaging system of Medison Co. Ltd., an input volume data is preprocessed by scan-conversion before applying the proposed algorithm. The visualization here is based on ray-casting method [6] to produce 3D images by directly processing the VOI without an intermediate geometrical representation. Since only the data within the VOI are rendered, the system offers in real time actual 3D images of fetuses on display devices as depicted in Fig.6.
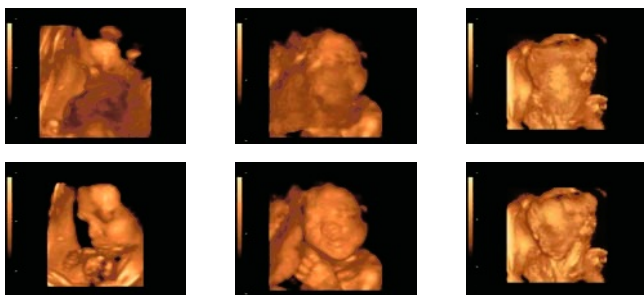
Fig. 6. Visualization of raw fetal volumes (first row) and their visualization with the proposed ROI determination (second row).

5   Conclusion

An automatic body ROI determination using SVM based top bound selection for visualization of a fetal ultrasound volume has been proposed. The experimental results have proved the appropriateness of the proposed feature vector in the validity test for an improvement of body ROI determination rate.

Acknowledgement

References

1. D. R. Ney and E. K. Fishman. "Editing tools for 3D medical imaging," *IEEE Computer Graphics and Applications*, Vol. 11, No. 6, pp. 63-71, Nov. 1991.
2. T. D. Nguyen, J. I. Kwak, S. H. Kim, and N. C. Kim, "Automatic selection and validation of face ROI for 3D fetal ultrasound imaging", in *Proc. Int. 17th EURASIP Conf. Biosignal 2004*, Brno, Vol. 17, pp. 334-336, June 2004.
3. Y. D. Chun, S. Y. Seo, and N. C. Kim, "Image retrieval using BDIP and BVLC moments," *IEEE Trans. on CSVT*, Vol. 13, No. 9, Sep. 2003.
4. J. C. Platt, "Fast training of SVMs using SMO," *Adv. in Kernel Method - S. V. Learning*, B. Scholkopf, C. Burges and A. Smola, eds. MIT Press, Cambridge, 1998.
5. M. Cheriet and J. N. Said, "A recursive thresholding technique for image segmentation," *IEEE Transactions on Image Processing*, Vol. 7, No. 6, June 1998.
6. E. Bullitt and S. R. Aylward, "Volume rendering of segmented image objects," *IEEE Trans. on Medical Imaging*, Vol. 21, No. 8, pp. 998 -1002, Aug. 2002.

Swarm Intelligence and the Holonic Paradigm: A Promising Symbiosis for a Medical Diagnostic System

Rainer Unland[1] and Mihaela Ulieru[2]

[1] University of Duisburg-Essen
Institute for Computer Science and Business Information Systems (ICB)
45117 Essen, Germany
unlandR@cs.uni-essen.de
[2] Canada Research Chair, The University of New Brunswick,
Faculty of Computer Science, Fredericton, Canada
http://eis.enel.ucalgary.ca

Abstract. A self-organizing medical diagnosis system, mirroring swarm intelligence to structure knowledge in holonic patterns is presented. The system sets up on an alliance of medical experts - realized by agents - that stigmergically self-organize in order to provide a viable medical diagnosis. Starting point is always a flat set of autonomous agents that spontaneously and temporarily organize on a case-based basis into a holarchy (without hierarchical control flow). Despite their sophisticated task the proposed agents, like ants in an ant colony, exhibit a comparatively simple architecture built on reactive behavior. The real power of the system stems from the fact that a large number of those simple agents collaborate in order to come to a reliable diagnosis.

Introduction

Medical diagnostic is fundamentally the process of identifying the disease a patient is suffering from in order to determine how it can be treated best. Although research on medical diagnostic has been a hot topic in computer science for quite a while, especially in the field of artificial intelligence, computer-based medical diagnostic has not yet reached a breakthrough. Reasons for that are manifold. Most proposals published so far are monolithic systems that rely on a sophisticated and ever-growing set of rules whose interplay and "firing" cannot be controlled easily. For diagnosis in general and medical diagnostic in particular a seamless, fast and easy integration of new and an abortion of outdated knowledge is of the essence. Multi-agent systems are a first major step in this direction since they provide decentralized control and, thus, can be adapted and extended much easier. A next step is to rely on a huge number of relatively unsophisticated (reactive) agents instead of a smaller number of sophisticated (deliberative) ones. This avoids the modeling of complex domain knowledge, thus, avoids sophisticated architectures that can hardly be controlled. To implement such a system we propose to combine swarm intelligence techniques with the holonic paradigm (see [1], [2] for the first papers proposing the idea of a holonic medical diagnostic system). Swarming systems enable a system to satisfy requirements that would be surprising to its original designer as stated in [3]. The holonic paradigm adds the nested hierarchical structure in which the swarms may self-organize [4]. Altogether,

such a medical diagnostic system will exhibit all the key properties of these paradigms, most of all self-organization, emergent behavior [5], simplicity, fast and unbounded learning, flexibility, robustness, adaptivity, self-sustainability, and easy extensibility. Nevertheless, the system will still exhibit a relatively simple architecture that can be maintained and controlled comparatively easily. This is because the sophisticated behavior of such a system stems from the fact that a huge number of simple components on the so-called micro-level exhibit a sophisticated behavior at their interface, the so-called macro-level, due to their (implicit) collaboration and interplay.

Emergence, self-organization, and the holonic paradigm have some similarities [6]. In fact, holonic systems can be seen as one way to implement emergent systems. Basic building blocks of holonic systems are holons (cf. [7]). On one level of abstraction a holon can be seen as a unit with a unique interface hat provides some services. On the next lower level of abstraction a holon (recursively) consists of a set of lower-level holons that cooperate in order to solve the problem at hand. The unique and sole interface to a holon is realized by a so-called mediator or head agent. All the other agents of a holon are called body agents (cf. [8]). Holons form a nested structure, called holarchy, where each holon is defined on the next lower level of abstraction by a set of body holons. This recursively continues until, finally, the leaf level, consisting of atomic holons, is reached. Holonic multi-agent systems consist of autonomous agents with holonic properties that recursively group together to form holons and, by this, holarchies (cf. [6], [9]).

Functionality of the Diagnostic System

When designing a complex system like the proposed holonic medical diagnostic system (HMDS) a number of questions concerning the architecture of the system have to be answered. After a thorough investigation, where several other alternatives [10] were explored, we concluded that an integrated approach with relevant concepts from the holonic paradigm, multi-agent systems technology and swarm intelligence is the most promising approach. What results is an open, highly adaptable architecture. It consists of many comparatively simple components that collaborate in order to solve complex problems. Basis is the autonomous or flat agents' approach of holonic systems [11] where all agents are initially on the same level. Communication between agents is solely done via blackboards. Each non-leaf agent is assumed to possess such a blackboard on which messages can be published that are of interest for its (potential) body agents. Due to space limitations we can not discuss the system in detail. The interested reader is referred to [4].

The proposed system mirrors stigmergy (for our parallel we will use the ant colony as an example for stigmergy) as follows: When a request for a medical diagnosis is sent to the HMDS it is dealt with by its single entry point, the mediator agent. Based on the classification of the disease (e. g., eye disease, skin disease, inner organ disease, etc.) the request is then assigned to those holons that are assumed to be specialists in the field on this level of expertise. This procedure is recursively repeated until the request reaches the final level, the leaf level of the holarchy. Results that are obtained by individual holons now flow the other way round, from the bottom to the top of the holarchy. On their way up the results are examined, assessed, put in context and, finally, sorted according to their estimated likelihood and quality.

Individual, Case-Based Holarchy Formation
In HMDS the nested holarchy will emerge dynamically and individually on a case-based basis. For each diagnosis request a new, individual holarchy evolves. Thus, holarchies are formed in an ad-hoc way and exist only temporarily. As soon as a diagnosis request is sent to the HMDS the holarchy begins to evolve out of the flat set of agents. It emerges in a combination of top-down and bottom-up approach. The general direction is top-down while each holon is formed bottom-up. More specifically, first one or more agents agree to the job of working on the diagnosis request. In order to do so they must have some confidence that they will be able to provide a reasonable diagnosis. Each of them becomes the head of a new holon. In order to find appropriate body agents that are willing to help in the diagnosis finding process each head agent posts its request for support on its blackboard. Other agents can read this request and decide to cooperate. By that they become body agents of the holon - a new holon evolves. From now on body agents of the same holon are called buddy agents. Since communication is solely done via blackboards, requests as well as replies are posted on the appropriate blackboard. Thus, starting with the mediator agent of HMDS the holarchy grows from the top, level by level and holon by holon, till it finally reaches the leaves. A holarchy evolves where each holon can be seen as a swarm or colony of ants, thus, forming a nested ant colony. An ant on one level of abstraction is realized by an ant colony/swarm of ants on the next lower level of abstraction. Note that this temporary holarchy only comprises holons that are actively involved in the diagnosis finding process. Dead branches will not evolve.

Pheromone Trails
The main purpose of ants in an ant colony is to gather food to be able to survive. The main goal of an ant in HMDS is to solve a diagnosis request. On this basis there is obvious conformity between the knowledge about the existence of a blackboard and the knowledge of an ant about the existence of a food source due to the pheromone trail. For this reason, this concept was adopted. Whenever there is a relationship between two agents in the sense that one agent reacts to requests posted on the blackboard of a head agent or, more general, uses the blackboard of the head agent, a (the) pheromone trail from the body agent to (the blackboard of) the head agent is established, respectively strengthened. What emerges is a system whose structure is defined by blackboards and pheromone trails. The downward directed relationships from agents with broader knowledge to agents with deeper knowledge is emulated by the blackboard, thus, not encoded anywhere. The reverse relationship, the upward one, is expressed by the pheromone trail between agents. This implied set of acquaintances is not only very simple but also self-maintaining.

Emergence and Self-organization by Random Search and Feedback Messages
In ant colonies pheromone trails are the means for communication. In HMDS the blackboard plays the role of the pheromone trail and the information on it emulates the scent of the trail.

Scent Strength
In ant colonies ants will more likely follow a trail with strong scent. Since the scent evaporates or gets stronger in dependence of the usage of the trail by ants, scent is one

of the means for self-organization. In HMDS scent is emulated by the blackboard and the information on it. Each blackboard is partitioned in specific areas that are solely dedicated to those agents that have a pheromone trail to this blackboard, one partition for each agent. Among other information, each partition provides information about the strength of its pheromone trail. Partitions are sorted according to the strength of their pheromone trail. Thus, an agent that is on a quest and accidentally passes by (see the discussion about random search below) can immediately identify the more successful agents with respect to this blackboard and can compare itself to it. In addition to this body agent-specific scent value the blackboard maintains a holon-specific scent value that discloses information about the overall success of the head agent. It is essentially the sum of the individual scent values of all body agents.

Scent Type
In ant colonies pheromone trails are open for all agents; i. e., ants will follow it because of its strength but not because of the type of food they can expect at its end. In HMDS the latter plays a role since the ability to provide a successful diagnosis depends on whether the input matches the field of expertise of an agent. If not it is worthless; i. e., the trail is not worth to follow. In comparison to an ant colony this means that each ant can only carry a specific type of food. Thus, only trails to this kind of food source are of interest for them. HMDS emulates this in the following way: for every pheromone trail of an agent information about it is provided in their partition on each blackboard to which they have a pheromone trail. This information comprises the scent value of the trail, recent results from using this trail, the DDPs that were used in solution processes related to this trail, etc. If an agent wants to try new trails/holons it can check this information for compatibility to its own data. If there is a great overlap this indicates that the pheromone trails of this agent are probably of the right type. Thus, it may be worth to test them.

Explorer Agents
Ants in an ant colony either are involved in food carrying or are on a quest to find new food sources. Similarly, agents in HMDS either are engaged in the solution process of a request or are on a quest through the information space of the whole system (random search) in order to find better holons. Agents on a quest are called explorer agents. Such agents are very important for an uncontrolled and unpredictable, however, positive development. Due to the randomization accidental hits are possible which may lead to developments that could not be anticipated beforehand. An agent that wants to explore new possibilities has several ways to do so. The easiest way is to check the blackboards of all holons it is member of for success stories about involvement from other body agents in other holons, called buddy check in the following. As discussed above this is not exploring in its original sense but following an already existing, however, from the point of view of the agent, new pheromone trail that was laid by a buddy agent. Seen from the likelihood of success such a proceeding seems to be the most promising since it can be assumed that there is a lot of overlap between buddy agents. An explorer agent explores its environment by walking around randomly.

While explorer agents will walk randomly, it seems to be more logical that they will find new promising holons especially in their direct neighborhood. By direct neighborhood those holons are meant that share the same higher level (parent) holon

than the original holon of the explorer agent. This guarantees that the new holon is at least a representative of a (class of) disease(s) from the same super-class of diseases than the original holon. However, in order to make accidental positive development happen a random walk is a prerequisite.

Pheromone Trail Selection

As long as agents are in a prospering symbiosis with one or several holons they probably will decide not to worry about other success stories on blackboards they have access to. However, especially if all pheromone trail of an agent are on a comparatively low level then the likelihood that the agent may try new possibilities increases substantially. As discussed above there are two alternatives for a change. One is to read success stories on accessible blackboards and the other is to start a quest. In both cases a decision has to be made whether to join a visited (quest) or proposed holon (buddy check). In both cases the information about agents has to be analyzed. Decisive factors can be, on the one hand, the scent strength and, on the other hand, the overlap between the set of symptoms that was used by the "advertising" agent and the set of symptoms that is used by the searching agent.

Reinforcement Learning as Basis for Self-organization

Knowledge about pheromone trails can be seen as the life elixir that allows the agents to stay alive. Thus, it has to be their main goal to strengthen their pheromone trails as much as possible, respectively to add new, successful relationships to their list of acquaintances/holons.

Strengthening of Pheromone Trails by Providing Sound Solutions

An agent that provided something positive to the solution process of a request (a diagnosis that turned out to be valid) will strengthen its pheromone trail to the according blackboard.

Evaporation of Pheromone Trails by Negligence

Like with real ant colonies, a pheromone trail will evaporate over time if it is not replenished. Thus, if an agent does not want to work for a holon any longer (due to more attractive possibilities with other holons) the trail to it will slowly evaporate and finally diminish. Relationships that are cultivated will become stronger and stronger, especially, if the contribution of the delivering agent is of high quality and the competition is not too high; i. e., not too many other body agents deliver the same diagnosis to a request.

General Evaporation of Pheromone Trails

As with real ant colonies as well a pheromone trail that is not used enough any longer will weaken its scent and finally diminish. In order to prevent successful agents from losing their knowledge about blackboards just because their advice has not been asked for for a while (because, e. g., they represent a very seldom disease) the evaporation of the pheromone trail depends on the frequency in which requests are posted on a blackboard to which a pheromone trail exists and not on the time. Whenever a request is posted on the blackboard of a head agent of a holon the pheromone trail of all body agents will be evaporate a little bit. If an agent decides to contribute to the solution process, the delivery of a response will once again weaken the trail. This ensures that

a trail will evaporate if a body agent does not use it any more. It also ensures that it will evaporate if the body agent delivers a poor response. By how much the trail is weakened overall depends on the influence of the response on the diagnosis process. Exceptionally poor performances of agents lead to a faster evaporation of their pheromone trails since they have to pay the entrance fee, however, will not get a reward. In general, each agent that reacts to a diagnosis finding request will first weaken its pheromone trail. Thus, only when its contribution was strong enough the difference between this "entrance fee" and the reward will be positive, thus, strengthening its pheromone trail. This proceeding will filter out all those agents that only submit average or below average diagnoses in comparison to other agents.

When a new pheromone trail is created, i. e., an agent has joined a new holon, this trail needs some "seed capital" to allow it to participate in solution processes. This capital/scent can be withdrawn from existing pheromone trails, e. g., from those ones that will most likely not be used any longer. They will evaporate faster or even diminish immediately, however, finance the exploration of new trails.

Emergence Through Birth and Death of Agents
Less successful agents will lose knowledge about their pheromone trails. They may have to search for better holons in order to survive. If they do not find at least one proper holon, they will finally die. The other way round, agents that are very successful, what implies that there is a high demand of their services, can clone themselves. The "value" of an agent is decided by the number of pheromone trail it has knowledge of and the strength of each of these trails. If a certain threshold is exceeded an agent clones itself. This supports the natural selection process in the system. Successful agents will survive and proliferate while luckless ones finally will die (principle of survival of the fittest). Since new agents first need to find their place in the system, they will be equipped with an asset high enough to allow them to survive for some time even if they are not successful. They will not only inherit the knowledge of their generator about blackboards but also the strength of the scent that is related to each of these pheromone trails. However, they also exhibit a strong impulse to explore new possibilities. More specifically, the threshold for abandoning an existing pheromone trail and searching for new ones is much higher than that of their generators. This guarantees that successful agents do not flood their environment with, in principal, identical buddies that do not contribute to the quality of the solution process at all.

Conclusions

In today's ever-changing world that grows together ever faster, a quick, ubiquitous, and reliable diagnosis generation is of eminent importance as can be seen from the recent problems with SARS. In addition, the cost explosion problem in the health care sector asks for a highly efficient, cost-effective and reliable health care system. This paper presented a holonic medical diagnostic system that unifies the advantages of multi-agent system technology with self-organizing, emergent behavior, the simplicity, the capability of fast and unbounded learning, the flexibility, the robustness, the adaptivity, and the reliability of swarm intelligence and holonic systems. The resulting system supports the medical institution in its diagnosis generation, especially in cases in which the kind of disease the patient is suffering from can only be assumed.

References

1. Ulieru, M. and Geras, A.: *Emergent Holarchies for e-Health Applications: A Case in Glaucoma Diagnosis*; Proc. IECON '02; HOL - Agent-based Intelligent Automation and Holonic Control Systems; Sevilla, Nov. 2002
2. Ulieru, M.: *Internet-Enabled Soft Computing Holarchies for e-Health Applications,* Invited Chapter in: Fuzzy Logic and the Internet; Zadeh, L.A.; M. Nikravesh (Eds.), Physica Verlag; 2003
3. Van Dyke Parunak, H.; and Brueckner, Sven A.: *Engineering Swarming Systems*; in F. Bergenti, M.-P. Gleizes, F. Zambonelli (eds.): Methodologies and Software Engineering for Agent Systems; Kluwer Publishing Company; 2004
4. Ulieru, M.; Hadzic, M.; Chang, E,: *Soft Computing Agents for e-Health Applied to the Research and Control of Unknown Diseases*, Information Science (in print - to appear 2005).
5. De Wolf, T., Holvoet, T.: *Emergence and Self-Organisation: A Statement of Similarities and Differences*; Workshop on Engineering Self-Organizing Applications (ESOA); July 21-24, New York, USA; 2004
6. Ulieru, M.: *Emergence of Holonic Organizations from Multi-Agent Systems: A Fuzzy-Evolutionary Approach*, Invited Chapter in: Loia, V. (ed.): Soft Computing Agents, IOS Press; 2002
7. Koestler, A: *The Ghost In The Machine*, Arkana Press; 1967
8. Schillo, M: *Self-Organization and Adjustable Autonomy: Two Sides of the Same Coin?;* In Connection Science, 14 (4), pp. 345-359, 2003
9. Ulieru, Mihaela: *The Holonic Enterprise: Modeling Holarchies as MAS to Enable Global Collaboration*. Third Int. Workshop on Industrial Applications of Holonic and Multi-agent Systems - HoloMAS, 2-6 Sept. 2002, Aix-en-Provence, France. IEEE Computer Society, ISBN 0-7695-1668-8, pp 603-607; 2002
10. Ulieru, M.; Stoica, C.; Klüver; J.; Unland, R.: *A Holonic Medical Diagnosis System Based on MAS Technology and Neural Networks*; submitted for publication
11. Fischer, K.; Funk, P.; Ruß, C.: *Specialized Agent Applications*; in Luck, M.; Marik, V.; Stepankova, O.; Trappl, R. (eds.): Multi-Agent Systems and Applications; Springer Pub. Company; Lecture Notes in Artificial Intelligence, Vol. 2086; pp.365-382; 2001

Analysis Between Lifestyle, Family Medical History and Medical Abnormalities Using Data Mining Method – Association Rule Analysis

Mitsuhiro Ogasawara[1], Hiroki Sugimori[2], Yukiyasu Iida[1], and Katsumi Yoshida[2]

[1] Department of Medical Engineering, Toin Yokohama University, 1614 Kuroganemachi
Aoba-ku, Yokohama city, Japan
m.ogasawara@k8.dion.ne.jp
[2] St. Mariannna Univercity School of Medicine, 2-16-1 Sugo, Miyamae-ku
Kawasaki-city, Japan

Abstract. We conducted data mining method (association rule analysis) to elucidate the relationship between 6 lifestyles (overweight, drinking, smoking, meals, physical exercise, sleeping time, and meals), 5 family medical histories (hypertension, diabetes, cardiovascular disease, cerebrovascular disease, and liver disease), and 6 medical abnormalities (high blood pressure, hypercholesterolemia, hypertrigriceridemia, high blood sugar, hyperuricemia, and liver dysfunction) in examination data using the medical examination data of 7 years, obtained from 5,350 male employees in the age group of 40-49 years. We found that number of combinations derived from data mining (association rule method) was greater than that derived from conventional method (logistic regression analysis). Moreover, values of both "confidence" and "odds ratio" derived from association rule were greater than that derived from logistic regression. We found that "the association rule method" was more and useful to elucidate effective combinations of risk factors in terms of lifestyle diseases.

Keywords: health checkups, medical abnormality, lifestyle, family medical history, data mining.

1 Introduction

It is widely known that the top three primary causes of death in Japan are malignant neoplasm, cardiac diseases, and cerebrovascular disorder so-called. These diseases are referred to as "adult diseases" due to the fact that it because they develops and progresses increasingly from the age of 40's years and above with ages. However, it has been elucidated that an individual lifestyle habits, such as diet, smoking, and drinking, worse deeply strongly related to the onset of these diseases.; Therefore, recently, the "adult diseases" has been instead called are being referred to as "lifestyle diseases". For the onset of lifestyle diseases, the influence of lifestyle habits is considered to be significant. Historically, the principle of secondary prevention of "early recognition and treatment" was emphasized on during the medical examination; however, the guiding principle has recently been modified based on the concept that a good health fitting status is necessary for the primary prevention of the onset of the diseases. Accordingly this regard, the reexamination of the lifestyle of people at working in industries and providing health-related guidance of to the local community is considered

important;. Thus, the prevention of lifestyle diseases before happens, which also contributes to the positive effect on the medical economics, is anticipated through the individual improvement of health-related issues that are indicated by the evaluation of the individual lifestyles at an early stage

However, the health-related guidance that is provided on the based on the results of medical examinations at industries and local communities is generally uniform in general, and the individualized guidance that relies on the individual based on specific lifestyle conditions is rarely provided in reality. At the same time, medium-to-long term health-related guidance may be provided by the elucidation of how elucidating the manner in which the combination of transitions over a period of time and lifestyle affects the lifestyle disease, the medium-to-long term health guidance may be provided.

Previous studies that analyzed the relationship between lifestyle and lifestyle diseases mainly focused on the analyses of short-term (one to two years) effects of lifestyle on the examination data. This also can be attributed to the fact that only a few researches have been conducted due to the enormous number of combinations number. Recently, an analytical method to extract useful information from a huge amount of data—data mining method—has been attracted attention, and it has been utilized for analyzing a massive amount of medical data that has been accumulated due to the widespread use of electronic medical records, etc.

In this study, using long-term (seven years) medical examination data of company employees, genetic factors that greatly affect on the unhealthy lifestyle and the onset of lifestyle disease were picked up as the family medical history from the long-term (seven years) medical examination data of company employees., and the association rule method, which is one of the data mining methods, was employed to quantitatively analyze the relationship of the abnormalities in the examination data.[1-4]

2   Subjects and Methods

2.1   Subjects

Company employees who had undergone medical examinations regularly for seven years from 1991 to 1997 at the healthcare administration of the company were selected. The distribution of subjects, in terms of generations, based on their ages in the fiscal year 1992 (the reason for the fiscal year 1992 to be selected as the base is described in section 3-1) are shown in Table 1. The distribution of people on the basis of their ages is dispersed. The number of subjects required for higher reliability of the analysis should be large. Thus, subjects between the age group of 40-49 years, which included most of the employees, were specifically chosen. These employees had undergone the same set of examinations.

A total of 5,350 males were included in this study. Of these, healthy subjects in whom no abnormality was detected and no medication was prescribed after examination during the fiscal years 1991 and 1992 were selected and subjected to each examination item. The mean age of the selected group (age, 40-49 years) was 44.5 ± 2.5 years. The 40-44-years age group included 2,687 subjects and the 45-49-years age group included 2,663 subjects; significant bias was not observed with regard to age distribution.

Table 1. The distribution at age

	male	female	total
less than 40 years old	18	1	19
40-49years old	5350	632	5972
50 years old or more	642	98	740
total	6010	731	6741

2.2 Lifestyle of Subjects and Definition of Unhealthiness

The lifestyle was defined as the total of six items–"obesity," "alcoholism," "cigarette smoking," "exercising," "consuming three meals," and "sleeping." These items of lifestyle were determined based on the health index by Breslow. The health index proposed by Breslow, an American scientist, consists of seven healthy habits.

The seven healthy lifestyle habits are as follows:

1. Having adequate sleep
2. No smoking
3. Maintaining adequate weight
4. Moderating alcohol consumption
5. Exercising regularly
6. Taking breakfast every day
7. No eating between meals

The seven healthy lifestyle habits proposed by Breslow signify that people who implemented more of these habits showed a greater tolerance to diseases and lived longer. This was introduced in the Annual Report on Health and Welfare, published in 1997, and the Breslow's health index is considered as the fundamental contemporary health management method. The definition of unhealthiness, based on the six lifestyle items considered in this study, is shown in Table 2. The body mass index (BMI) is indicated as the weight (kg)/height $(m)^2$, and the exercising habit was defined as exercising continuously for more than 20 min at least twice a week.

Although Breslow's health index is composed of seven items, six items were included in this study because two items, namely, taking breakfast and no eating between meals, could be considered simultaneously when the healthcare administration questioned the subjects whether three meals were taken regularly everyday. [5-7]

Table 2. The definition of unhealty lifestyle

item	the criteria of unhealty lifestyle
overweight	BMI≧25
drinking (japanese sake)	1.8 liters or more for a week
smoking	yes
exercise custom	no
sleeping time	less than 6 hours during a day
three meals during a day	irregular

BMI:Body Mass Index
exercise custom:doing exercise of contunuance for 20 min. or more
2 times or more during a week

2.3   Definition of Examination Items and Abnormality

The six examination items that were included are as follows:

1. Blood pressure (SBP: systolic blood pressure, DBP: diastolic blood pressure)
2. TC: total cholesterol
3. TG: triglyceride
4. FBS: fasting blood sugar
5. UA: uric acid
6. Liver function (GOT: glutamic oxalacetic transaminase, GPT: glutamic pyruvic transaminase, γ-GTP: gamma glutamyl transpeptidase)

   Table 3 shows the definition of abnormality detected in these examination items; these followed the criteria of our healthcare administration.

Table 3. Medical check-up items and the definition of abnormality

items	criteria
Blood Pressure	SBP≧140mmHg or DBP≧90mmHg
Total Cholesterol	TC≧220mg/dl
Triglyceride	TG≧150mg/dl
Fast Blood Sugar	FBS≧110mg/dl
Uric Acid	UA≧7.0mg/dl
Liver Function	GOT≧40IU or GPT≧40IU or γ-GTP≧50IU

SBP∶Systolic Blood Pressure、DBP∶Diastolic Blood Pressure、TC∶Total Cholesterol
TG∶Triglyceride、FBS∶Fast Blood Sugar、UA∶Uric Acid、GOT∶Glutamic Oxalacetic
Transaminase、GPT∶Glutamic Pyruvic Transaminase、γ-GTP∶Gamma Glutamyl
Transpeptidase

2.4   Family Medical History

The involvement of genetic factors in the onset of diseases, such as hypertension and diabetes, has been known from previous studies. In this study, five diseases that were selected by our healthcare administration as diseases with family medical history are listed below.

1. Hypertension
2. Diabetes
3. Cardiovascular disease
4. Cerebrovascular disease
5. Liver disease

3   Analytical Method

3.1   DataBase

The data from the fiscal years 1991 to 1997 (37,450 data = 5,350 × 7) were imported into the database software Access (Microsoft Co.). The database containing the life-style information for seven years, data on nine laboratory items (SBP, DBP, TC, TG, FBS, UA, GPT, GOT, γ-GTP), information about medication, and family medical history of five diseases, which were considered for analysis, was constructed for each

subject using Access and imported to Excel (Microsoft Co.). Using Excel, fiscal year 1992 was set as the base and the subjects in whom no abnormality in laboratory data was detected and no medication was prescribed in fiscal years 1991 and 1992 were selected as subjects with normal data for each examination item. The reason for setting this condition was to exclude those subjects in whom any abnormality was detected during the years prior to the fiscal year that was set as the base.

The number of people with normal examination data for each item is shown in Table 4.

Table 4. The number of people with normal examination data

items	number of people
Blood Pressure	3239 （61%）
Total Cholesterol	4154 （78%）
Triglyceride	3321 （62%）
Fast Blood Sugar	4627 （86%）
Uric Acid	3981 （74%）
Liver Function	2860 （53%）

With regard to six lifestyle habits of each subject, the number 1 was assigned if an unhealthy lifestyle had been followed for four years or more during the six years between 1992 and 1997 or the number 0 was assigned otherwise. Similarly, with regard to the family medical history, the number 1 or 0 was assigned based on the presence or absence of the history of each disease, respectively. Further, with regard to medical abnormality, the number 1 or 0 was assigned based on the presence or absence of any medical abnormality, respectively, during the five fiscal years between 1993 and 1997.

3.2 Association Rule Analysis

The association rule describes the combination of items: $A = A_1, A_2, \ldots A_n$, $B = B_1, B_2, \ldots B_m$, as $A \rightarrow B$. Here, A and B are called the antecedent and the consequent, respectively. In the case in which the consequent B is true (all the items, $B_1, B_2, \ldots B_m$ are true) and if the antecedent A is also true (all the items, $A_1, A_2, \ldots A_n$ are true), the association rule is evaluated by the ratio of the number of events in which B is true to the number of events in which A is true (confidence) and the ratio of the number of events in which A and B occur simultaneously to the total number of events (support). Confidence corresponds to the conditional probability that the consequent occurs, given the antecedent.

In general, association rule analysis is considered effective when the confidence and support are high; however, it is not effective in the case in which the rule $A \rightarrow B$ gives high confidence and support and the rule $\overline{A} \rightarrow B$ gives a similar confidence, which indicates that B occurs with the same probability irrespective of the occurrence of A. $\overline{A}$ means complementary set of A.

Thus, by using the 2×2 contingency table as shown in Fig. 1, the significance of the rule can be evaluated by the degree of how much the value of the odds ratio, $(a \times d)/(b \times c)$, is significantly higher than 1, where a is the number of simultaneous occurrences of A and B, b is that of A and $\overline{B}$, c is that of $\overline{A}$ and B, d is that of $\overline{A}$ and $\overline{B}$, and $N = (a + b + c + d)$.

Then, the significance can be assessed by the following equation:

$$N(|ad - bc|- N/2)^2/(a + b)(a + c)(b + d)(c + d) \qquad (1)$$

This is assumed to follow the Chi-squared χ^2 distribution with one degree of freedom.

	A	$\overline{A}$
B	a	b
$\overline{B}$	c	d

Fig. 1. The 2 x 2 contigency table

With regard to the A→B rule, the significance can be evaluated by the χ^2 test based on the odds ratio $(a \times d)/(b \times c)$ in the 2×2 contingency table below. However, if any of the expected numbers, $(a + b)*(a + c)/N$, $(a + b)*(b + d)/N$, $(a + c)*(c +d)/N$, $(b + d)*(c + d)/N$, is less than or equal to five in the 2×2 contingency table, the significance probability of the χ^2 will be evaluated as low. Therefore, Fisher's exact test was employed for the assessment.

The occurrence of abnormality in the examination data was highly dependent on the first-year data. Accordingly, for further analysis the data for each examination item were separated into two classes—class 1 and class 2—depending on whether the values were greater or lower than the median values. Based on this method of classification, for the items such as blood pressure and liver function in which abnormality was evaluated from multiple examination values, the data were categorized as class 2 if any of the multiple examination data exceeded the median value and as class 1 if all of them were lower than the median. The blood pressure and liver function were considered abnormal if any of examination data exceeded the normal range; thus, we made the classification criteria consistent with it.

3.3 Odds Ratio-adjusted for Confounding in Association Rule Analysis

Odds ratio mentioned at 3-2 was the ratio of A1, A2, ..., An→B compared with A1, A2, ..., An→B. To elucidate odds ratio of each risk factor adjusted for confounding, we should remove and exclude influences of other risk factors in the combinations of rules (i.e. excluding the effect of Aj $(j \neq i)$ onto Ai) Therefore, to calculate odds ratio of each risk factor adjusted for confounding in the rules can analyze as follows,

 [1] method that keeping constant with proportion of elements of risk factors (proportional allotment)

 [2] method that remove other elements of risk factors from relevant risk factors.

In this paper we analyzed by using the latter method [2] which was relatively easier to calculate. By letting S be the combinations of the negations Aj of the items excluding Ai, *p* be the confidence of the association rule Ai • S→B, and p' be the confidence of the rule Ai • S→B, the odds ratio of Ai excluding the effect of others (i.e., removing the effects of confounding factors) onto Ai can be evaluated by the equation (Formula 1),:

$$(p / (1 - p)) /(p' / (1 - p')).\qquad\qquad\text{(Formula 1)}$$

Using this odds ratio adjustment for confounding, we might be able to compare with the conventional logistic regression methods, which is one of the conventional statistical methods used to deal the occurrence of abnormality in terms of risk factors. Namely, the confidence p of the association rules, $A_1, A_2, ..., A_n \rightarrow B$, can be expressed by the equation, $p = 1 / [1 + \exp (- (\alpha_0 + \alpha_1 A_1 + \alpha_2 A_2 + \bullet\bullet\bullet ... + \alpha_n A_n))]$, where $\alpha_0, \alpha_1, \alpha_2, ...,$ and α_n are the logistic regression coefficients. In this equation here, α_0 and α_i can be approximated constantly and the coefficient corresponding to Ai, respectively, and the exponential of the regression coefficient αi, $\exp(\alpha i)$, would represent be the odds ratio of Ai excluding the effect of A_j $(j \neq i)$ onto A_i. In logistic regression analysis, explanatory variables that significantly affect response variables are automatically selected according to the statistical evaluation criteria. We analyzed the same database using logistic regression analysis (SAS program Version 8.02 SAS Institute Inc., Cary, NC, USA) based on the stepwise method.

4   Results and Discussion

4.1   Average Abnormality Rate in Total Subjects

The average occurrence rate of an abnormality in the examination data of the two classes that were divided based on the values from the first fiscal year for each examination item was evaluated before evaluating the occurrence rate of an abnormality with regard to the combination of lifestyle and family medical history. The average occurrence rate of an abnormality in the examination data is the ratio of the number of subjects in each class in whom any abnormality was observed once or more during the five years to the total number of subjects in each class.

Table 5 shows the number of subjects categorized into the two classes based on the examination data in the first fiscal year, average occurrence rate of an abnormality in the examination data, and the 95% confidence interval (95% CI) of the abnormality occurrence rate. Table 5 shows the total number of healthy subjects in classes 1 and 2 for each examination item in the first fiscal year. The subjects that are classified by single laboratory examination item, such as TC, TG, FBS, and UA, are categorized into two classes based on whether the data values are more or less than the median value. As a result, the number of subjects in each class is approximately equal. On the other hand, the subjects that are classified by multiple laboratory examination items, such as blood pressure and liver function, are classified as class 2 when any of the examination data exceed the median value. Therefore, the number of subjects in class 2 was relatively higher than that in class 1. For the comparison of classes with regard to the occurrence rate of abnormality, the rate in class 2 was three to eight times higher than that in class 1.

4.2   Abnormality Rate in Subjects Without the Risk
in Lifestyle/Family Medical History

Next, the subjects in whom no abnormality was detected in the lifestyle and family medical history were selected, and the abnormality rates in their examination data were evaluated. Table 6 shows the number of subjects who did not have any risk

factor, but showed abnormal examination values, the occurrence rate of abnormality, and the 95% CI for each examination and class. For every examination item and every class, the number of subjects with no abnormality in lifestyle and family medical history was significantly low—only 2% to 4% of the subjects.

Table 5. The number of subjects categorized into the two classes based on the examination data in the first fiscal year, average occurrence rate of abnormality

items	class	number of people	number of abnormality	rate of abnormality	(95%CI		)
Blood Pressure	1	1406	140	0.10 (	0.08 -	0.12	)
	2	1833	762	0.42 (	0.39 -	0.44	)
Total Cholesterol	1	2085	133	0.06 (	0.05 -	0.07	)
	2	2069	921	0.45 (	0.42 -	0.47	)
Triglyceride	1	1651	221	0.13 (	0.12 -	0.15	)
	2	1670	736	0.44 (	0.42 -	0.46	)
Fast Blood Sugar	1	2512	145	0.06 (	0.05 -	0.07	)
	2	2115	524	0.25 (	0.23 -	0.27	)
Uric Acid	1	2094	80	0.04 (	0.03 -	0.05	)
	2	1887	586	0.31 (	0.29 -	0.33	)
Liver Function	1	781	55	0.07 (	0.05 -	0.09	)
	2	2079	729	0.35 (	0.33 -	0.37	)

The data values were highly variable; therefore, simple comparison with the average occurrence rates of abnormality was impossible. However, the averages of each examination item and class were in the range between a third and one of the average abnormality rates, and the values were significantly different depending on the examination item and class.

For example, almost no difference was found in the blood pressure, the class 2 of TC, UA, and class 1 of liver function.

The χ^2 test was performed on the groups with and without any abnormality in lifestyle and family medical history. The differences in ratios in the two groups were statistically significant with regard to TG and class 2 liver function ($p < 0.05$).

Table 6. The number of sujects who did not have any risk factor, number of abnormality and the occurrence rate of abnormality

items	class	number of people	number of abnormality	rate of abnormality	(95%CI		)
Blood Pressure	1	43	2	0.05 (	0.00 -	0.11	)
	2	60	24	0.40 (	0.28 -	0.52	)
Total Cholesterol	1	53	1	0.02 (	0.00 -	0.06	)
	2	54	23	0.43 (	0.29 -	0.56	)
Triglyceride	1	72	5	0.07 (	0.01 -	0.13	)
	2	39	7	0.18 (	0.06 -	0.30	)
Fast Blood Sugar	1	75	3	0.04 (	0.00 -	0.08	)
	2	48	6	0.13 (	0.03 -	0.22	)
Uric Acid	1	65	3	0.05 (	0.00 -	0.10	)
	2	48	10	0.21 (	0.09 -	0.32	)
Liver Function	1	29	2	0.07 (	0.00 -	0.16	)
	2	72	16	0.22 (	0.13 -	0.32	)

4.3 Extracted Rules

The antecedent contains 6 items of lifestyle and 5 items of family medical history, that is, a total of 11 items producing 2,048 ($2^{11} = 2,048$) combinations. The total num-

ber of extracted rules was 4,371. The odds ratio was more than or equal to one in 3,327 (76%) of the 4,371 extracted rules. Table 7 shows the number of rules, the averages of confidence, support, and odds ratio for each examination item and class. The odds ratios were statistically significant in 489 rules; this was approximately 11% of the total rules. However, the number of effective rules varied among the examination items. The confidence was in the range of 0.17 to 0.83, and the odds ratio was in the range of 3.12 to 12.02. The average support of rules was 0.017, and the number of events in which the rule was effective was low—less than 2% of the total events. The confidence of the association rule corresponded to the occurrence rate of abnormality, where the combination of risk factors in the antecedent was the condition. When the average occurrence rates of abnormality presented in Table 5 were compared for every examination and class, the rate of increase in the abnormality occurrence rate due to the combination of risk factors with respect to the average occurrence rate of abnormality was 1.4- to 7.4- fold. The increase in the rate due to the combination of risk factors was small in comparison to the difference in the occurrence rate of abnormality in classes 1 and 2. Excluding TC, the difference was 0.02 to 0.04, which was almost one tenth smaller in comparison with the difference of the average occurrence rate of abnormality of classes 1 and 2 that was 0.19 to 0.38.

Table 7. The number of rules and statistically significant rules, the average of statistically significant confidence, statistically significant support, and statiscally significant odds ration for each examination parameter and class

items	class	number of rules	number of statistically significant rules	confidence of rules		support of rules		odds ratio of rules	
Blood Pressure	1	209	30 (14%)	0.37 ±	0.20	0.01 ±	0.01	7.36 ±	6.37
	2	312	56 (18%)	0.66 ±	0.14	0.03 ±	0.03	4.23 ±	3.81
Total Cholesterol	1	174	15 (9%)	0.17 ±	0.05	0.00 ±	0.00	3.12 ±	1.02
	2	245	7 (3%)	0.78 ±	0.10	0.01 ±	0.00	5.56 ±	2.45
Triglyceride	1	223	46 (21%)	0.29 ±	0.12	0.02 ±	0.02	3.16 ±	2.49
	2	373	28 (8%)	0.63 ±	0.11	0.06 ±	0.06	2.64 ±	1.26
Fast Blood Sugar	1	365	99 (27%)	0.31 ±	0.20	0.00 ±	0.00	10.73 ±	11.72
	2	365	114 (31%)	0.54 ±	0.17	0.01 ±	0.02	5.34 ±	4.74
Uric Acid	1	271	9 (3%)	0.28 ±	0.16	0.00 ±	0.01	12.02 ±	9.17
	2	357	35 (10%)	0.57 ±	0.15	0.02 ±	0.03	3.97 ±	2.91
Liver Function	1	84	1 (1%)	0.27 ±	-	0.00 ±	-	5.18 ±	-
	2	349	49 (14%)	0.59 ±	0.13	0.02 ±	0.03	3.32 ±	2.45

Table 8 shows the association rules in which the confidence was highest for each examination item and class. The confidences of rules for each examination and class were in the range of 0.24 to 0.92 and highly dispersed. This is due to the fact that the rules with a high confidence include many combinations of risk factors, and the number of such events is low. Odds ratios were in the range of 4.7 to 50.0, and the rule with the maximum odds ratio was of the subjects who showed an abnormality in lifestyle habit—cigarette smoking, exercising, and consuming three meals—and at the same time, belonged to class 1 with a family medical history of hypertension and cerebrovascular diseases. The occurrence rate of abnormality in the blood sugar level for these subjects is drastically higher than that in the subjects without such combinations, and their odds ratios are more than five times higher than the others even in the 95% CI. The common tendency of the rules with a high confidence was that most of them included obesity and alcohol consumption, and for the class 2, the family history that was related to the risk factors of abnormality in examination data was history of blood pressure abnormality and hypertension, cholesterol abnormality and cardiovascular disease, and blood sugar level abnormality and diabetes.

Table 8. The association rules in which the confidence was highest for each examination parameter and class

items	class	antecedent					confidence	support	odds ratio	
Blood Pressure	1	over weight	drinking	smoking	exercise custom	three meals	0.67	0.38	18.59	**
	2	over weight	drinking	smoking	*hypertension*		0.92	0.14	17.12	**
Total Cholesterol	1	over weight	smoking	exercise custom	sleeping time	three meals	0.24	0.18	4.73	**
	2	over weight	drinking	sleeping time	*cardiovascular*		0.86	0.26	7.52	*
Triglyceride	1	over weight	drinking	smoking	exercise custom	*cerebrovascular*	0.67	0.35	13.05	*
	2	over weight	drinking	sleeping time	three meals		0.80	0.25	5.12	*
Fast Blood Sugar	1	smoking	exercise custom	three meals	*hypertension*	*cerebrovascular*	0.75	0.43	49.99	**
	2	over weight	three meals	*diabetes*			0.88	0.23	21.53	**
Uric Acid	1	drinking	smoking	exercise custom	three meals	*diabetes*	0.50	0.43	25.79	*
	2	over weight	exercise custom	*hypertension*	*cardiovascular*		0.83	0.30	11.19	*
Liver Function	1	exercise custom	sleeping time	*diabetes*			0.27	0.26	5.18	*
	2	drinking	smoking	exercise custom	sleeping time	*diabetes*	0.88	0.23	13.08	**

itaric letter : family history   **:significant by χ^2 test $(P<0.01)$ *:significant by χ^2 test $(P<0.05)$

4.4 Comparison of Association Rule Analysis with Conventional Modeling (The Logistic Regression Analysis)

All combinations of risk factors that were selected by the logistic regression modeling were also extracted from association rule analysis. Table 9 shows results from adjusted odds ratio and its 95% confidence intervals (95%CI) that were derived from the stepwised logistic regression model, and from the association rule analysis mentioned in 3-2. In the logistic regression modeling, we could not find significant combination of risk factors for Total Cholesterol abnormality and liver dysfunction in class 1. Moreover, the number and odds ratios of combination were larger in results from association rules analysis than that from logistic modeling (Table 8). Odds ratio adjusted for confounding mentioned 3-3 (association rule analysis) was similar to that of logistic modeling, other than blood pressure and triglyceride.

Table 9. Comparison between association rule analysis and logistic regression analysis

items	class	risk factors	Logistic regression adjusted odds ratio (95% IC)	Association rulue analysis adjusted odds ratio	odds ratio of rules
Blood Pressure	1	over weight	2.09 (1.33 - 3.20) **	1.52	3.79 **
		drinking	1.76 (1.22 - 2.52) **	1.49	
	2	Overweight	1.84 (1.46 - 2.31) **	1.72 **	3.89 **
		drinking	1.74 (1.42 - 2.12) **	1.54 **	
		hypertension	1.32 (1.05 - 1.66) *	1.26	
Total Cholesterol	1	-			
	2	-			
Triglyceride	1	over weight	2.42 (1.65 - 3.52) **	2.79 **	2.89 **
		drinking	1.87 (1.40 - 2.49) **	1.97 **	
	2	over weight	1.63 (1.29 - 2.04) **	1.36	2.10 **
		smoking	1.36 (1.12 - 1.66) **	1.25	
Fast Blood Sugar	1	over weight	2.58 (1.80 - 3.71) **	2.61 **	3.10 **
		drinking	1.71 (1.22 - 2.39) **	1.72 **	
	2	over weight	2.22 (1.80 - 2.73) **	2.33 **	2.59 **
		drinking	1.23 (1.00 - 1.50) *	1.35 *	
		three meals	1.34 (1.06 - 1.70) *	1.47	
Uric Acid	1	drinking	1.62 (1.02 - 2.54) *	1.62 *	1.62 *
	2	over weight	1.25 (1.00 - 1.56) *	1.36 *	1.87 *
		derinking	1.37 (1.12 - 1.67) **	1.48 **	
		cardiovascular	1.44 (1.01 - 2.04) *	1.50	
Liver Function	1	-			
	2	over weight	1.65 (1.32 - 2.07) **	1.84 **	1.86 **
		drinking	1.72 (1.42 - 2.08) **	1.84 **	

itatic letter : family history   **:significant by χ^2 test $P<0.01$) *:significant by χ^2 test $(P<0.05)$

5 Conclusion

We conducted data mining method to elucidate the relationship between the lifestyle, family medical history, and abnormality in examination data using the medical examination data of 7 years, obtained from 5,350 male employees in the age group of 40-49 years. As a result, we could extract more significant combinations of risk factors in the association rule analysis, and these odds ratios were higher than those obtained from the conventional epidemiological methods (logistic modeling), which treat each risk factors as an independent valuable. Consequently, we found that "the association rule method" (e.g., the data mining method) was more and useful to elucidate effective combinations of risk factors in terms of lifestyle diseases.

References

1. Bigus, Joseph P: Data Mining with Neural Networks: Solving Business Problems - From Application Development to Decision Support. New York: McGraw-Hill. 1996.
2. Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J.: Classification and Regression Trees. Routledge. 1983
3. Fayyad, U. M. et al. eds.: Advances in Knowledge Discovery and Data Mining. AAAI Press.1996
4. Heckerman, D. et al. eds: Proceeding Third International Conference on Knowledge Discovery and Data Mining. AAAI Press. 1997
5. H. Sugimori,Y. Yamada, M. Suka, T. Tanaka, T. Izuno, A. Kiyota, Y. Iida, K. Yoshida "Multiple Risk Factor Syndrome in Japanese Male Subjects Using Automatic Multiphasic Health Testing and Service Data A Work-site Cohort Study", Health Evaluation And Promotion, Vol. 30, No. 4, pp.48-51,2003
6. H. Sugimori, Y. Iida, M. Suka, T. Izuno, T. Kato, K. Yoshida. " Analysis of factors associated with improvement of hypertension considering behavioral modification in Japanese worker - the Pooling Repeated Observation (PRO) method ?",International Conference on Health Promotion 2003
7. Y. Iida, E. Takahashi, Y. Suzuki,T. Izuno, S. Ohtsuka,K. Yoshida : "Contribution of past life- styles and medical indicators in medical indicator value prediction", 18th Biennial Conference of International Health Evaluation Association,1998

A Moving-Mass Control System for Spinning Vehicle Based on Neural Networks and Genetic Algorithm

Song-yan Wang, Ming Yang, and Zi-cai Wang

Control & Simulation Center, Harbin Institute of Technology, 150080 Harbin, P.R. China
wangsyhit@yahoo.com.cn

Abstract. The ability of a moving-mass control system to control a spinning vehicle using two internal moving mass actuators is investigated. The nonlinear equations of motion are provided, and the influence to the system of moving masses' motion with respect to the vehicle's shell is described. For the self-learning capacity of the neural networks and the optimum ability of the genetic algorithm, the hybrid trajectory PID control scheme based on the neural networks and genetic algorithm is produced to improve the dynamic qualities and the adaptive capacity of the system. A nonlinear simulation of a typical mission profile demonstrates the ability of the controller to effectively control the vehicle's trajectory.

1 Introduction

Over the years, techniques of controlling the flight of missiles have gravitated to systems that deliver relatively large amounts of control authority. Several studies have suggested that Moving-Mass Control System (MMCS) appears to offer the greater design and cost advantages [1~4].

The actuators of an MMCS are the internal moving masses, which are moved within the Maneuvering Spinning Vehicle (MaSV) to offset the center of mass (c. m.) of system. The control torque resulting from the product of the aerodynamic force and the distance the c. m. has moved from centerline is used to control the vehicle. It has several design advantages since an MMCS operates within a vehicle and is fully protected from the environment. Such as an MMCS does not directly interact with the hypersonic flow field surrounding the vehicle unlike flaps or surface-mounted thrusters. The earliest analysis of an MMCS found in the literature was done in 1967 by Nelson et al [5]. This report contained a classical roll autopilot design and a six-degree-of-freedom (DOF) dynamic simulation. Some assumptions were taken to simplify the design of control system, however it did not agree with the fact. And further studies of the MMRCS with a moving mass were performed by Regan et al at the U. S. Naval Systems Warfare Center. Regan and his co-workers established the governing equations with the relative motion and devised a Linear Quadratic Regulator (LQR) controller [6].

Neural Networks Control System (NNCS), as a branch of the artificial intelligent control method, are used to construct learning control systems for a class of unknown nonlinear systems. For such kind of adaptive learning control system used in this paper, the weights of the network need to be updated using Genetic Algorithm (GA) [7]. So the stability of the closed-loop control system is improved.

This paper derives the control ability of the MMCS with two moving masses. The nonlinear equations of motion of the coupled MaSV-moving-mass system, as well as the influence of the masses' motion to the system, are provided and discussed. The trajectory control algorithm is presented by using hybrid PID controller based on the NN and GA. A simulation demonstrates the ability of the controller to effectively control the vehicle's motion.

2   Equations of Motion

The basic principle by which an MMCS is able to control the vehicle's motion is to produce the control torque by using the aerodynamic forces and moving the masses within the MaSV to offset the c. m. of system.
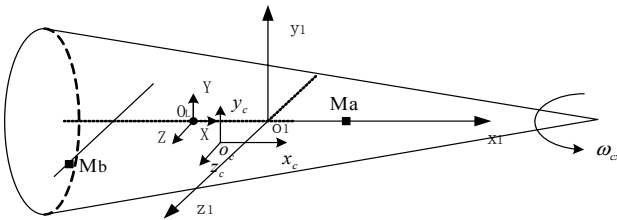
Fig. 1. Spinning Vehicle with Two Moving Masses

The MMCS of the spinning vehicle is shown in Fig. 1. Missile's shell of mass M and two moving masses (M_a and M_b) combine to form system of mass Mt, where $Mt = M + M_a + M_b$. Points of interest are O_1, the mass center of the shell, and O_c, the mass center of system. Coordinate frames are the local geodetic inertial frame d, the body-fixed frame 1, and the system frame c.

This section derives the equations of motion fully accounted for the dynamic coupling between the three bodies. The moving masses were allowed to translate with respect to the MaSV's shell, but were not allowed to rotate with respect to the MaSV's shell. Both the MaSV and the moving masses were assumed to be rigid bodies. The governing equations of motion were derived in the c frame, whose origin was located at the instantaneous c. m. of the three-body system. The c frame rotates with the MaSV and is aligned with MaSV shell's c. m. 1 frame.

Then the vector force equation, from Newton's second law, is

$$\sum \vec{F} = M_t \cdot \dot{\vec{v}}^c_{c/d} \tag{1}$$

where $\bar{v}^c_{c/d}$ means velocity of c frame with respect to d frame, and the left-hand side of the force equation are the external forces acting on the MaSV system.

And the absolute momentum torque about the instantaneous c. m. of system is

$$H^c = I^c_{M_T} \cdot \omega^c_{c/d} + \mu_a \cdot r^c_{a/1} \times v^c_{a/1} + \mu_b \cdot r^c_{b/1} \times v^c_{b/1}$$
$$- \mu_{ab} (r^c_{a/1} \times v^c_{b/1} + r^c_{b/1} \times v^c_{a/1}) \tag{2}$$

Then the following angular moment equation of system motion about the c frame is obtained

$$M^c = I_{M_T}^c \cdot \dot{\omega}_{c/d}^c + \dot{I}_{M_T}^c \cdot \omega_{c/d}^c + \mu_a r_{a/1}^c \times \dot{v}_{a/1}^c + \mu_b r_{b/1}^c \times \dot{v}_{b/1}^c$$
$$+ \omega_{c/d}^c \times I_{M_T}^c \cdot \omega_{c/d}^c + \mu_a \omega_{c/d}^c \times r_{a/1}^c \times v_{a/1}^c + \mu_b \omega_{c/d}^c \times r_{b/1}^c \times v_{b/1}^c \qquad (3)$$
$$- \mu_{ab}(r_{a/1}^c \times \dot{v}_{b/1}^c + r_{b/1}^c \times \dot{v}_{a/1}^c - \omega_{c/d}^c \times r_{a/1}^c \times v_{b/1}^c - \omega_{c/d}^c \times r_{b/1}^c \times v_{a/1}^c)$$

where $I_{M_t}^c$ means instantaneous inertia tensor of system, $\omega_{c/d}^c$ is the angular velocity, and $\bar{r}_{a/1}^c$ is the relative position of mass *Ma* with respect to 1 frame resolved in *c* frame. The right-hand side of the equation (3) contains eleven terms in all, which manifest the dynamic coupling that exists between the moving masses and the MaSV except the first and the fifth terms. This dynamic coupling also represents the influence, which includes the disturbances caused by the accelerations and velocities of the masses, to the angle velocities of the system. The mass ratios are defined as

$$\mu_a = \frac{M_b + M}{M_t} M_a, \quad \mu_b = \frac{M_a + M}{M_t} M_b, \quad \mu_{ab} = \frac{M_a M_b}{M_T} \qquad (4)$$

And the system inertia tensor is given by

$$I_{M_T}^c = B_1^c I_M^1 + B_a^c I_{m_a}^a + B_b^c I_{m_b}^b + M(r_{1/c}^{cT} \cdot r_{1/c}^c \cdot [1] - r_{1/c}^c \cdot r_{1/c}^{cT})$$
$$+ M_a(r_{a/c}^{cT} \cdot r_{a/c}^c \cdot [1] - r_{a/c}^c \cdot r_{a/c}^{cT}) + M_b(r_{b/c}^{cT} \cdot r_{b/c}^c \cdot [1] - r_{b/c}^c \cdot r_{b/c}^{cT}) \qquad (5)$$

where I_M^1 is the inertia tensor of the missile's shell, $I_{m_a}^a$ and $I_{m_b}^b$ are the inertia tensor of the mass elements, $r_{1/c}^c$ means the relative position of 1 frame with respect to *c* frame resolved in *c* frame, and [1] is the identity matrix.

The moments acting on the system about the system frame as *c* are defined

$$M^c = B_1^c \cdot M_R^1 + r_{c/1}^c \times B_1^c \cdot F_R^1 \qquad (6)$$

where B_1^c is the direction-cosine matrix from 1 to *c* frame, M_R^1 is the aerodynamic moment about the shell's c. m. resolved in 1 frame, and F_R^1 means the aerodynamic forces on multi-body system resolved in 1 frame.

Furthermore, the equations of motion of MaSV system also include some relative movement functions and nonlinear aerodynamic functions. The equations of motion clearly indicate that the MMCS is a complex nonlinear system which has the variable coefficients and large disturbances caused by the accelerations and velocities of masses.

3 Moving Mass Control Algorithm

Since the PID controller is simple and easy to realize, and the NNCS is not dependent on the accuracy of the model completely, the hybrid method is produced to improve the dynamic qualities of system. The control coefficients of PID are calculated and updated by NNC, and the weights of the network are trained by GA. With the nonlinear-mapping and self-learning characteristics of NNCS, the adaptive ability of system is increased.

A simplified block diagram of the hybrid control system is provided in Fig. 2. The expected sight angles determined by the proportional navigation scheme are com-

pared with the actual ones, and the error signal e and its differential $\dot{e}$ are sent to the NNCS to adjust the control coefficients of PID controller. GA, as a training method, is used to optimize the weights of the network to reduce the training time. The controller calculates the commanded position of c. m. of the system. And then the actuator realizes the mass movement within the MaSV. The actual masses' positions are used by the plant model to calculate the new sight angles, and the guidance control loop continues.
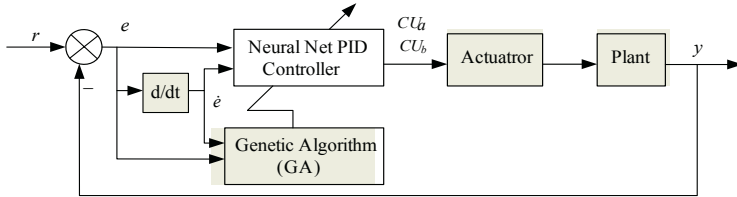
Fig. 2. Hybrid PID Controller Based on NNCS and GA

A formal linear PID controller is designed to confirm the value of the position of c. m. needed by the system. For the limitation of parameters offered by the missile, the compact linear control function of the system is

$$u = k_p \Delta q + k_d \Delta \dot{q} + \zeta \tag{7}$$

where Δq is the error of the sight angle, k_p and k_d are the control coefficients, and ζ is the corrected term. Since the control coefficients are adjusted by NNC, the hybrid controller has the characteristics of nonlinear controller.

Two three-layered NNC with the net structure as $N_{2,3,2}$ are used to control the trajectory in vertical and horizontal plane, respectively. Using the controller in vertical plane as an example, the inputs of input-layer are given by

$$\begin{cases} c_1(k) = \Delta q(k) \\ c_2(k) = \Delta q(k) - \Delta q(k-1) \end{cases} \tag{8}$$

And the input of the ith nerve cell of hidden-layer is

$$x_i(k) = \sum_{j=1}^{2} w_{ij} c_j(k) \tag{9}$$

where w_{ij} is the weight of the input-layer to the hidden-layer, and c_j is the output of the jth nerve cell of input-layer. The output of the hidden-layer is

$$d_i(k) = f[x_i(k)] \tag{10}$$

where $f(\bullet)$ is the nonlinear function. And the output of the output-layer is

$$u_p(k) = \sum_{i=1}^{3} {}^2 w_{ip} d_i(k) \tag{11}$$

where $^2 w_{ip}$ is the weight of the hidden-layer to the output-layer.

The performance index or cost function is

$$E = \sigma_1 \Delta e + \sigma_2 t_s + \sigma_3 \xi \tag{12}$$

where Δe, t_s and ξ represent the absolute value of the steady-state error, adjustment time and the maximum peak overshoot respectively. And σ_i $(i = 1,2,3)$ means the coefficient of the performance index, while $\sum \sigma_i = 1$.

In GA training block, the weight of NNCS is coded into a 10-bit binary unsigned integer, so the chromosome (or string) in GA population will have 170-bit in length. And the error signal e and its differential $\dot{e}$ sent to GA block are used to determine the actual range of the value of the weight (namely $(w_{ij})_{\min} \sim (w_{ij})_{\max}$) to adjust the coding precision.

And the fitness function is defined as

$$f = \frac{1}{E + \Delta} \tag{13}$$

where Δ is a function with a little value. So it can be seen that the fitness of a set of control coefficients (named an individual of the chromosomes in population) will decrease when the value of performance index becomes greater.

The proportional operator is used in GA. And the fitness of individual is adjusted linearly for avoiding the premature convergence. To quicken up the convergence velocity, the optimized individual is left. And to improve the search efficiency, the adaptive crossover and mutation probability (p_c, p_m) are performed as follows

$$p_c = \begin{cases} k_1 - \dfrac{(k_1 - k_2)(f' - \bar{f})}{f_{\max} - \bar{f}}, & f' \geq \bar{f} \\ k_1, & f' < \bar{f} \end{cases} \tag{14}$$

$$p_m = \begin{cases} k_3 - \dfrac{(k_3 - k_4)(f_{\max} - f)}{f_{\max} - \bar{f}}, & f \geq \bar{f} \\ k_3, & f < \bar{f} \end{cases} \tag{15}$$

where f' is the maximum fitness of the two parental chromosome individuals, f is the individual fitness that needs to be mutated, and $k_1 \sim k_4$ are the constant coefficients.

After the above operations, the desired c. m. of system, which is sent to the mass position algorithm to calculate the change of position of each mass, is gained.

4   Simulation Results

A nonlinear simulation of a typical mission is shown in following figures, where the results of the normal linear PID controller in controlling the nonlinear system are shown in Fig. 3 and the results of hybrid PID controller are shown in Fig. 4.

It can be seen from the plots that the hybrid controller is more effective than the normal one, and the control error of the system of hybrid controller is smaller than the ordinary one.

5   Conclusions

In this paper an investigation was conducted to determine the ability of a moving-mass control system to control a spinning vehicle. The general equations of motion for a rigid body system moving with two internal moving mass actuators were derived, and the influence of the mass movement to the system was discussed. The

trajectory control algorithm was arrived at using the hybrid PID control method based on the neural networks and genetic algorithm. The adaptive ability of control system was obtained through the self-learning character of NNCS. And then the optimal solution was gained by optimizing the weights of network via GA training, to improve the dynamic performances and achieve the control requirements of system. Combing the proportional navigational scheme and the moving masses movement characters, the trajectory control algorithm were produced. A simulation proves the ability of the controller to effectively control the spinning system.
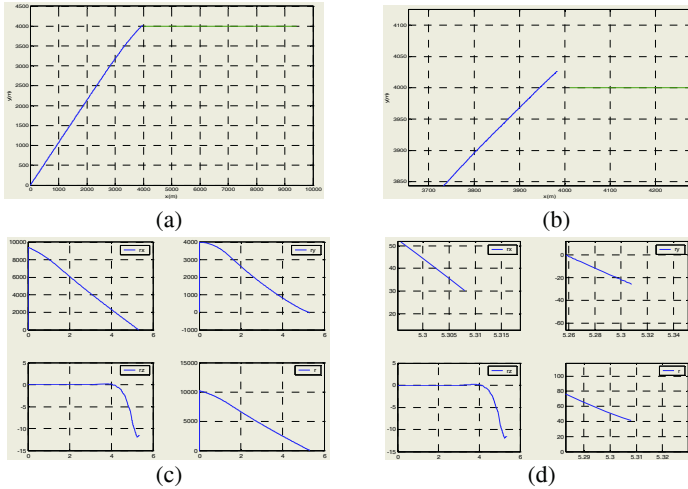
Fig. 3. (a) Mission Trajectory in Vertical Plane (b) Part of Junction (c) Relative Position of Missile and Target in 3-Direction of *d* Frame (d) Part of Relative Position
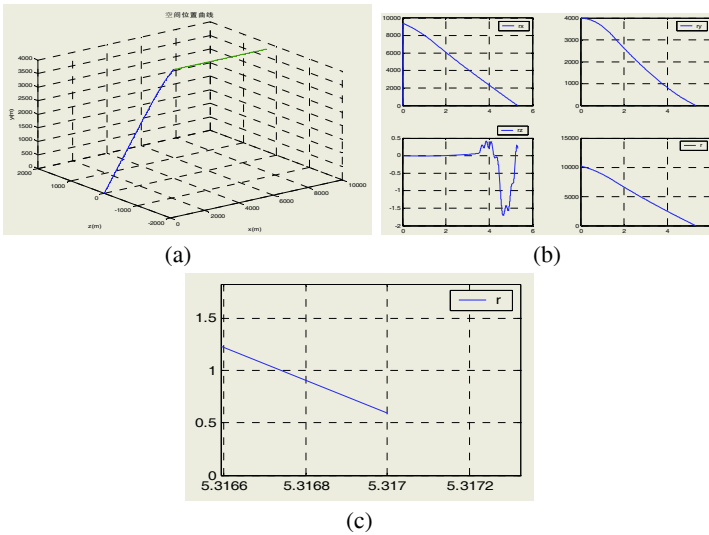
Fig. 4. (a) Mission Trajectory in 3-D Space (b) Relative Position with Missile and Target (c) Part of Relative Position

References

1. Raymond H. Byrne, Beverly Rainwater Sturgis, Rush D. Robinett.: A Moving Mass Trim Control System for Reentry Vehicle Guidance. AIAA-96-3438-CP
2. Awad EI-Gohary.: Global stability of the rotational motion of a rigid body containing moving mass. International Journal of Non-Linear Mechanics, Vol. 36, (2001) 663-669
3. Raymond H. Byrne, Rush D. Robinett, Beverly Rainwater Sturgis.: Moving Mass Trim Control System Design. AIAA-96-3926
4. William Whitacre, Dr. Craig Woolsey.: Rigid Body Control Using Internal Moving Mass Actuators.
   URL www.aoe.vt.edu/~cwoolsey/Advisees/IndependentStudy/WhitacreVSGCPaper.pdf
5. Nelson R. L., Price D. A., and Delpino F. H..: A New Concept for Controlled Lifting Entry Flight Experiments. NASA-CR-66718
6. Thomas Petsopoulos, and Frank J. Regan.: Moving-Mass Roll Control System for Fixed-Trim Re-Entry Vehicle. Journal of Spacecraft and Rockets, Vol. 33, No. 1, (1996)54-60
7. Xu lina.: Neural Net Control. Publishing House of Electronics Industry (2003)

Appendix

Two-Dimensional Fitting of Brightness Profiles in Galaxy Images with a Hybrid Algorithm

Juan Carlos Gomez, Olac Fuentes, and Ivanio Puerari

Instituto Nacional de Astrofísica Óptica y Electrónica
Luis Enrique Erro # 1, Tonantzintla, Puebla, 72840, México
{jcgc,fuentes,puerari}@inaoep.mx

Abstract. Fitting brightness profiles of galaxies in one dimension is frequently done because it suffices for some applications and is simple to implement, but many studies now resort to two-dimensional fitting, because many well-resolved, nearby galaxies are often poorly fitted by standard one-dimensional models. For the fitting we use a model based on de Vaucoleurs and exponential functions that is represented as a set of concentric generalized ellipses that fit the brightness profile of the image. In the end, we have an artificial image that represents the light distribution in the real image, then we make a comparison between such artificial image and the original to measure how close the model is to the real image. The problem can be seen as an optimization problem because we need to minimize the difference between the original optical image and the model, following a normalized Euclidean distance.

In this work we present a solution to such problem from a point of view of optimization using a hybrid algorithm, based on the combination of Evolution Strategies and the Quasi-Newton method. Results presented here show that the hybrid algorithm is very well suited to solve the problem, because it can find the solutions in almost all the cases and with a relatively low cost.

1 Introduction

Galaxies span a wide range of morphology and luminosity, and a very useful way to quantify them is to fit their light distribution. Fitting profiles for galaxies in one dimension is frequently done because it suffices for some applications and is simple to implement [5], but many studies now resort to two-dimensional fitting, because many well-resolved, nearby galaxies are often poorly fitted by standard one-dimensional models. Although empirical techniques for galaxy fitting and decomposition have led to a number of notable advances in understanding galaxy formation and evolution, many galaxies with complex isophotes, ellipticity changes, and position angle twists can be modelled accurately in two dimensions. We illustrated this by 5 examples, which include elliptical and spiral galaxies displaying various levels of complexities. In one dimension, the galaxy bulge and disk may appear to merge smoothly, which causes non-uniqueness in the decompositions, while in two dimensions ishophote twists and ellipticity change provide additional constraints to break those degeneracies.
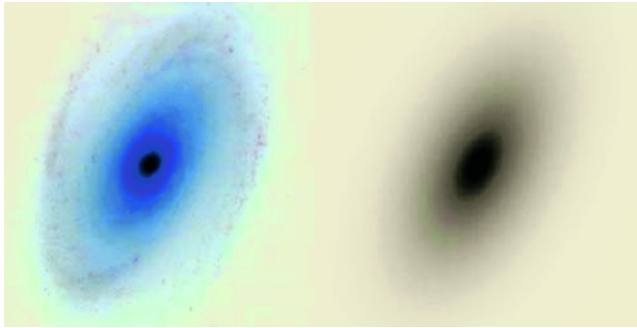
Fig. 1. Galaxy image and its modelled brightness profile

The algorithms we are using here are: Evolution Strategies (ES) [1, 6], Quasi-Newton (QN) [2, 3] and a hybrid algorithm, merging both of them. The hybrid algorithm takes advantage of the main features from the two previous algorithms: the global exploration of ES and the fast convergence of QN when the search is near the solution.

The rest of the paper is structured as follows: in Section 2 a brief description of theory of brightness profiling and a description of the problem are presented, the optimization methods are shown in Section 3, Section 4 includes the general description of the implementation, results are presented in Section 5 and Section 6 includes conclusions and future work.

2   Brightness Profile

Surface brightness in a galaxy is literally defined as how much light the galaxy emits[5], and luminosity is defined as the total energy received per unit of area per unit of time. Then, the surface brightness of an astronomical source is the ratio of the source's luminosity F and the solid angle (Ω) subtended by the source:

$$B = \frac{F}{\Omega} \tag{1}$$

The surface brightness profile in an elliptical galaxy follows the de Vaucoulers Law $r^{1/4}$ [4]:

$$I_b = I_e \exp\left[-7.67\left[\left(\frac{r^{1/4}}{r_e}\right) - 1\right]\right] \tag{2}$$

where r is the distance from the galactic center, r_e is the mean ratio of the galaxy brightness (the radius where half of the total brightness lies), and I_e is the surface brightness for $r = r_e$.

Also, the surface brightness profile for a disc galaxy has an exponential distribution:

$$I_d = I_0 \exp\left(-\frac{r}{r_d}\right) \tag{3}$$

where I_0 is the central surface brightness and r_d is the radial scale length.

Finally, surface brightness distribution in elliptical and spiral galaxies can be described as the sum of equations (2) and (3), which is an approximation of the profile using concentric ellipses [5].

$$I_d = I_b + I_d \qquad (4)$$

In fact, it is not expected that equations (2) and (3) fit all the profiles measured in the radial range of the galaxy, because sometimes sky substraction errors in external regions of galaxy can distort the profile. An example of a galaxy image and its corresponding generated brightness profile using the previous equations are shown in Figure 1.

The problem of two-dimensional fitting can be described in brief as follows: given an image, taken from photometric observations, of a spiral or elliptical galaxy, an exploration of search space will be done to estimate a set of parameters that define the surface brightness profile of the galaxy. Parameters to be determined by the algorithms are: r_e, mean ratio of the galaxy brightness; I_e, surface brightness in $r = r_e$; I_0, central surface brightness; r_d, radial scale length and two angles i_1 and i_2, which are the rotation angles about the x and z axes, with the x axis being horizontal and the z axis pointing towards the observer.

3   Optimization Methods

3.1   Evolution Strategies

Evolution Strategies (ES) [1, 6] is concerned with finding the global minimum of a function with a large number of variables in a continuous space. The algorithm starts by choosing k, $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_k\}$ individuals, each characterized by an object parameter vector $\mathbf{q}$ and a corresponding strategy parameter vector $\mathbf{s}$:

$$\mathbf{q}_i = \langle q_{1,i}, q_{2,i}, \cdots, q_{L,i} \rangle \ i = 1, \cdots, k$$
$$\mathbf{s}_i = \langle \sigma_{1,i}, \sigma_{2,i}, \cdots, \sigma_{L,i} \rangle \ i = 1, \cdots, k$$

In the first generation the elements of the $\mathbf{q}$ and $\mathbf{s}$ vectors can be chosen totally at random. Each of the k individuals must be evaluated according to a fitness function. The fitness function is what we need to minimize (or maximize, depending on the point of view), and also is called *target function*.

The next step is to produce a new population applying the genetic operators cross-over and mutation. For cross-over, two individuals (parents) are chosen at random, and then we create two new individuals (offspring) by combining the parameters of the two parents.

Mutation is applied to the individuals resulting from the cross-over operation. Each element of the new individual is calculated from the old individual using the simple operation: $\mathbf{q}_{j,\mathrm{mut}} = \mathbf{q}_j + N(0, \sigma_j)$, where $N(0, \sigma_j)$ is a random number obtained from a normal distribution with zero mean and standard deviation σ_j, which is given by the strategy parameter vector. The process of cross-over and mutation is repeated until the population converges to a suitable solution.

3.2   Quasi-Newton

To solve a system of non-linear equations Newton's method has to estimate the Jacobian in each iteration, which implies computing partial derivatives; that results in a very high computational cost. To avoid that complexity, Quasi-Newton (QN) method [2, 3] substitutes the Jacobian Matrix in Newton's method with an approximate matrix that is recalculated in each iteration.

In the beginning, the Jacobian $J(x_1)$ is substituted by a matrix A_1 in New-ton's method:

$$A_1 = J(x_0) + \frac{[f(x_1) - f(x_0) - J(x_0)(x_1 - x_0)] (x_1 - x_0)^t}{\|x_1 - x_0\|^2} \tag{5}$$

Then, this matrix is used to calculate x_2:

$$x_2 = x_1 - A_1^{-1} f(x_1) \tag{6}$$

Quasi-Newton works with the two previous equations, substituting the corresponding matrix for each iteration. It is possible to calculate the inverse matrix in (9) with the following equation, avoiding the inversion process in each iteration:

$$A_i^{-1} = A_{i-1}^{-1} + \frac{(s_i - A_{i-1}^{-1} y_i) s_i^t A_{i-1}^{-1}}{s_i^t A_{i-1}^{-1} y_i} \tag{7}$$

where $y_i = f(x_i) - f(x_{i-1})$ and $s_i = x_i - x_{i-1}$.

The previous equation uses only matrix multiplication, then the total number of operations in the whole process has a complexity $O(n^2)$.

3.3   Hybrid Algorithm
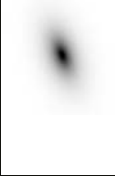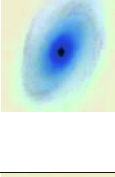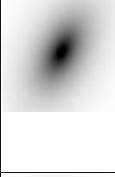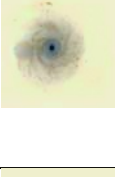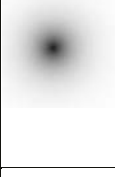
The hybrid algorithm used here is ES+QN. This hybrid algorithm is implemented using ES as a global method to identify promising regions in the search space, and then, once the region is located, we switch to a QN as local method to refine the best solution found by ES.

The employed metric to determine where the interesting region lies and when to change between algorithms is based on progress of the global algorithm, that is, if the ES algorithm has found a stable region, it is a good sign indicating that the optimum is near. So when the algorithm ES has not experimented an improvement of 10% in its best individual during 10 generations, we switch to the QN algorithm.

4   Implementation

ES were implemented using a population of 10 individuals and was evolved until the target function reached a value greater than 0.95. QN was implemented using the function FMINUNC included in the Optimization Toolbox of Matlab, using all the default parameters.

Table 1. Results for a set of galaxy images

Galaxy	Original Image	Best Model	Type	Algorithm	Function Evaluations	$\frac{1}{1+\frac{\|A-B\|^2}{B+\epsilon}}$
NGC2768			Elliptical	QN ES Hybrid	N/C 5139 **2970**	N/C 0.9824 0.9824
NGC2903			Spiral	QN ES Hybrid	N/C **3348** N/C	N/C 0.9719 N/C
NGC3031			Spiral	QN ES Hybrid	N/C 5737 **4716**	N/C 0.9616 0.9664
NGC3344			Spiral	QN ES Hybrid	**159** 2430 1588	0.9514 0.9514 0.9514
NGC4564			Elliptical	QN ES Hybrid	**302** 3671 693	0.9918 0.9917 0.9918

In the beginning we have an image of a spiral or elliptical galaxy. By convention we chose working with 256x256 images in grey scale. In the image we first determine the galactic center by finding the brightest pixels in a box of 10x10 pixels, in this box we take the central pixel and use its coordinates as the center of the galaxy in the image.

The process starts by generating randomly a vector (or set of vectors in the case of ES) of numbers which represents the set of brightness parameters described in Section 2: $[i_{1,j}, i_{2,j}, I_{e,j}, r_{e,j}, I_{0,j}, r_{d,j}]$

The parameter vector is used as input for a program which creates an artificial image that represents the light distribution in the galaxy, following equations (2) and (3). The obtained image is also 256x256 pixels in grey scale. Then, the artificial image is evaluated with the following target function:

$$f = \frac{1}{1 + \frac{\|A-B\|^2}{B+\epsilon}} \tag{8}$$

where A is the artificial image, B is the original optical image and ϵ is a very small constant to prevent division by 0. The target function represents the similarity between both images, and its value range is between 0 and 1, with 1 as a perfect match and 0 as totally different images. At the end, the simulated image that maximizes this equation is the one that was produced by the set of brightness parameters we are looking for. We established empirically that with a minimum value of 0.95 for the target function, the difference between images is almost imperceptible.

If the target function has not reached the tolerance, then the next step is to make a modification to the parameter vector. This modification is done according to the optimization method (ES or QN), following particular features of each technique, as seen in Section 3. Once the modification is done, the process is repeated again.

The number of times the process can be repeated is determined in different ways for each algorithm: for ES we established that the process stops if during 10 generations the target function has not been enhanced in a value greater than 0.3, and Matlab determines the total number of evaluations QN can do.

In the case of the hybrid algorithm, we start with ES and evolved the population until the target function has a value greater than 0.9, or if during 10 generations the target function has not been improved in a value greater than 0.1, then we change to QN, with the best individual reached until that moment as the starting point for the local algorithm.

5 Results

Table 1 shows the results obtained after applying the algorithms to a set of galaxy images. The first column indicates the name of the galaxy, the second and third columns show the original and best model reached (by one of the three algorithms) respectively, the fourth one the kind of algorithm used, the fifth one the total number function calls needed by the algorithm to reach convergence and the last one shows the value for the target function for the maximum found (in a normalized quantity between 0 and 1, with 1 as a perfect match).

In the fifth column the smallest number of function calls is marked with bold text, which indicates the algorithm with the best behavior in each example. We choose the number of function calls as the measure of efficiency because, in most cases, all the algorithms have very similar accuracies at the end.

From the table we can see that the hybrid algorithm has a very good behavior, as was expected, it found four of five solutions, and the cost is lower than ES (and in some cases similar to QN). QN is the algorithm with the most deficient behavior, since it presents difficulties to reach convergence in three of the five examples, but, on the other hand, when the algorithm was able to find a solution it did so with only a few iterations. Also from the figure we see that ES is the algorithm with the most stable behavior, it was able to find good solutions for all

the cases, but the cost to reach an acceptable model is very high in comparison with QN and the hybrid algorithm.

6   Conclusions

In this work we have solved the problem of two-dimensional fitting of brightness profiles for spiral and elliptical galaxies using a hybrid algorithm, based on a global optimization algorithm and a local optimization traditional algorithm. The hybrid method was compared with the two optimization techniques separately. The hybrid algorithm achieved the best results considering the total number of iterations and the number of solutions found. QN was the worst because it was not able to find solutions in most of the cases, as was expected, because the problem of fitting profiles is a complex problem with real noise, and this kind of algorithm are not very well suited to work with this. ES is the most reliable algorithm to find a solution, but the cost of finding a model can be very high. Thus, it is possible to conclude that the hybrid algorithm outperforms QN and ES in most of the examples.

References

1. Back, T., Hoffmeister, F., Schwefel, H. P., A Survey of Evolution Strategies, Proceedings of the Fourth International Conference on Genetic Algorithms. (1991), 2-9
2. Coleman, T. F., Li, Y., An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds, SIAM Journal on Optimization. (1996), 6, 418-445.
3. Coleman, T. F., Li, Y., 1994, On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds, Mathematical Programming. (1994), 67, 2, 189-224.
4. de Vaucouleurs, G., Recherches sur les Nebuleuses Extragalactiques, Ann. d'Astrophys. (1948), 11, 247
5. Karttunen, H., Fundamental Astronomy, Springer-Verlag. (1996)
6. Rechenberg, I., Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution, Stuttgart: Fromman-Holzboog. (1973)

A Hybrid Tabu Search Based Clustering Algorithm

Yongguo Liu[1,2], Yan Liu[3], Libin Wang[1], and Kefei Chen[1]

[1] Department of Computer Science and Engineering
Shanghai Jiaotong University, Shanghai 200030, P.R. China
{liu-yg,wang-lb,chen-kf}@cs.sjtu.edu.cn
[2] State Key Laboratory for Novel Software Technology
Nanjing University, Nanjing 210093, P.R. China
[3] School of Applied Mathematics
University of Electronic Science and Technology of China, Chengdu 610054, P.R. China
liuyan@uestc.edu.cn

Abstract. The clustering problem under the criterion of minimum sum of squares clustering is a nonconvex program which possesses many locally optimal values, resulting that its solution often falls into these traps. In this paper, a hybrid tabu search based clustering algorithm called KT-Clustering is developed to explore the proper clustering of data sets. Based on the framework of tabu search, KT-Clustering gathers the optimization property of tabu search and the local search capability of K-means algorithm together. Moreover, mutation operation is adopted to establish the neighborhood of KT-Clustering. Its superiority over K-means algorithm, a genetic clustering algorithm and another tabu search based clustering algorithm is extensively demonstrated for experimental data sets.

1 Introduction

The clustering problem is a fundamental problem that frequently arises in a great variety of fields such as pattern recognition, machine learning, and data mining. In this paper, we consider this problem stated as follows: Given N objects in R^m, allocate each object to one of K clusters such that the sum of squared Euclidean distances between each object and the center of its belonging cluster for every such allocated object is minimized. The clustering problem can be mathematically described as follows:

$$\min_{W,C} F(W,C) = \sum_{i=1}^{N} \sum_{j=1}^{K} w_{ij} \parallel x_i - c_j \parallel^2 \tag{1}$$

where $\sum_{j=1}^{K} w_{ij} = 1$, $i = 1, \ldots, N$. If object x_i is allocated to cluster C_j, then w_{ij} is equal to 1; otherwise w_{ij} is equal to 0. In Equation 1, N denotes the number of objects, K denotes the number of clusters, $X = \{x_1, \ldots, x_N\}$ denotes the set of N objects, $C = \{C_1, \ldots, C_K\}$ denotes the set of K clusters, and W denotes the $N \times K$ $0-1$ matrix. Cluster center c_j is calculated as follows:

$$c_j = \frac{1}{n_j} \sum_{x_i \in C_j} x_i \qquad (2)$$

where n_j denotes the number of objects belonging to cluster C_j. It is known that this clustering problem is a nonconvex program which possesses many locally optimal values, resulting that its solution often falls into these traps. Many clustering approaches have been developed [1]. Among them, K-means algorithm is a very popular one, but it is proved to fail to converge to a local minimum under certain conditions [2]. In [3], genetic algorithms are used to deal with the clustering problem, called GA-Clustering in this paper. GA-Clustering encodes the clustering partition as a chromosome. After a specified number of evolutions, the best individual obtained is viewed as the final solution. Tabu search is a metaheuristic technique that guides the local heuristic search procedures to explore the solution space beyond the local optimality [4], which has been successfully applied to computer networks, image processing, pattern recognition, etc. In [5], tabu search is proposed to deal with the clustering problem, called TABU-Clustering in this paper. It encodes the solution as a string similar to that in GA-Clustering. After a specified number of iterations, the best solution obtained is viewed as the clustering result. To efficiently use tabu search in various kinds of applications, researchers often combine it with local descent approaches. In [6], Nelder-Mead simplex algorithm, a classical local descent algorithm, and tabu search are hybridized to solve the global optimization problem of multiminima functions. It is known that K-means algorithm is simple and computationally attractive, in this paper, we introduce it into tabu search to establish a new clustering method called KT-Clustering to explore the proper clustering. Moreover, mutation operation, a genetic operator used in genetic algorithms, is adopted to set up the neighborhood of tabu search. The motivation is to accelerate the convergence speed of the clustering method based on tabu search and to attain the globally optimal result as much as possible. By computer simulations, its superiority over K-means algorithm, GA-Clustering, and TABU-Clustering is extensively demonstrated. The remaining part of this paper is organized in the following. In Section 2, KT-Clustering and its main components are described in detail. Results of computer simulations are extensively analyzed in Section 3. Finally, some conclusions are drawn in Section 4.

2   KT-Clustering Algorithm

Figure 1 gives the general description of KT-Clustering. Its most procedures observe the architecture of tabu search. Two new operations are developed, K-means operation and mutation operation. The detail introduction to tabu search can be found in [4].

Based on the structure of tabu search, KT-Clustering gathers the optimization property of tabu search and the local search capability of K-means algorithm together. Besides main procedures of tabu search, KT-Clustering introduces two new operations: K-means operation and mutation operation. K-means operation adopted here is one-step K-means algorithm which is used to modulate the distribution of objects belonging to different clusters, improve the similarity between objects and their cen-

```
Begin
    set parameters and the current solution Xc at random
    while (not termination-condition) do
        perform K-means operation
        use mutation operation to generate the neighborhood
        select the proper neighbor of Xc as the new current solution
        update the tabu list and the best solution
    end
    output the best solution
end
```

Fig. 1. General description of KT-Clustering

troids, and to accelerate the convergence speed of the clustering method. It is described as follows: Given a solution $X = x_1, \ldots, x_i, \ldots, x_N$, reassign object x_i to cluster C_k, $k = 1, \ldots, K$, iff

$$\| x_i - c_k \| \le \| x_i - c_l \|, \; l = 1, \ldots, K \; \text{and} \; k \ne l \tag{3}$$

Then new cluster centers, $c_1', \ldots, c_K'$, is calculated as follows:

$$c_k' = \frac{1}{n_k} \sum_{x_i \in C_k} x_i \tag{4}$$

where n_k denotes the number of objects belonging to cluster C_k. After this operation, the modified solution is viewed as the current solution X_c.

Mutation operation, in this paper, is applied to provide a proper neighborhood for tabu search to avoid getting stuck in local optima and find the optimal result. It is described as follows: Given the current solution $X_c = x_1, \ldots, x_i, \ldots, x_N$, $x_i = j$, $j = 1, \ldots, K$, the mutation probability p_m, and the size of the neighborhood N_t, for $i = 1, \ldots, N$, draw a random number $c_1', \ldots, c_K'$. If $p_i < p_m$, then $x_i^t = x_i$, $t = 1, \ldots, N_t$; otherwise $x_i^t = k, k = 1, \ldots, K, k \ne j$. Here, $X^t \ne X_c$, X^t denotes the tth neighboring solution. In this article, the solution is defined the same as that in GA-Clustering and TABU-Clustering, which is suitable for computing the objective function value and comparing KT-Clustering with these two methods.

In order to obtain the good performance of KT-Clustering, we need discuss the choice of several parameters including the size of the tabu list, the size of the neighborhood, and the mutation probability. The average and standard deviation values of iterations where the best result is firstly attained for different parameters are shown as Figure 2. Each experiment includes 20 independent trials and 1000 iterations.

It is found that with the increase of the size of the tabu list, the capability of the global search of KT-Clustering is enhanced but its capability of the local search is

restricted because many promising solutions may be tabooed. Therefore, this parameter is overlarge or over small will reduce the performance of KT-Clustering. When the size of the tabu list is equal to 20, the best performance is attained as shown in Figure 2(a). So, we choose this parameter to be 20. On one hand, the high quality and stable result can be achieved, and on the other hand, a great amount of computational resource can be saved.

(a)   (b)   (c)
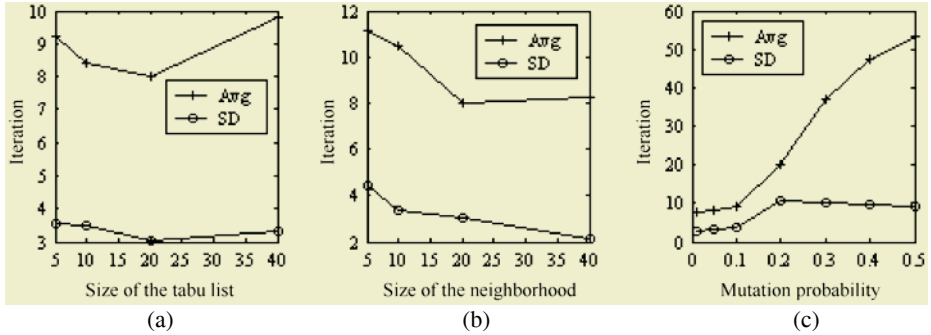
Fig. 2. Comparison of different parameters of KT-Clustering

It is found that, in Figure 2(b), the larger the size of the neighborhood the better, because KT-Clustering has more choices to select from. However, this is done at the expense of more computational effort. Therefore, deciding a proper value of this parameter is the process of exploring a balance between quality and cost. In this article, the case that the size of the neighborhood is equal to 20 can reach our goal. In this case, KT-Clustering can obtain the high quality and stable result at the expense of proper computational cost.

In KT-Clustering, the mutation probability is used to moderate the shake-up on the current solution and create a neighbor. The higher the value of this parameter, the less shake-up is allowed and, in consequence, the more similar the neighbor to the current solution, and vice versa. In this paper, the variable mutation probability reported in [3] is introduced. That is, the mutation probability reduces with the increase of the number of iterations. In [7], extensive researches have been conducted and the best parameter settings for genetic algorithms are given. Among them, the mutation probability is recommended to be in the range [0.005, 0.01]. We choose the terminal probability to be 0.005. Different original probabilities are compared as shown in Figure 2(c). It is seen that, as the increase of the orginal mutation probability, stability and performance of the algorithm decrease. In this article, we choose the original mutation probability to be 0.05.

3   Experiment Analysis

Before conducting simulation experiments, we analyze the time complexity of algorithms used in this paper. The time complexity of TABU-Clustering is $O(GN_tmN)$, where G is the number of iterations. The time complexity of GA-Clustering is $O(GPmN)$, where P denotes the population size. In KT-Clustering,

the time complexities of K-means operation and neighborhood creation are $O(KmN)$ and $O(N_t mN)$, respectively. Hence, the time complexity of KT-Clustering is $O(GN_t mN + GKmN)$. In most cases, K is a small constant, so its time complexity is $O(GN_t mN)$ which is the same as that of TABU-Clustering.

Performance comparisons between KT-Clustering and other techniques are conducted in Matlab on an Intel Pentium III processor running at 800MHz with 128MB real memory. Five data sets representing different distribution of objects are chosen to test the adaptability of KT-Clustering: two artificial data sets (Data set 1 and Data set 2) and three real life data sets (Ruspini, Iris, and Crude Oil). Data set 1 is a two dimensional data set having 200 nonoverlapping objects where the number of clusters is three. Data set 2 is a two dimensional data set having 250 overlapping objects where the number of clusters is five. Ruspini has 75 objects described by means of two attributes and four classes [8]. Iris represents different categories of irises having four feature values. The four feature values represent sepal length, sepal width, petal length, and petal width in centimeters. It has three classes with 50 samples per class [9]. Crude Oil has 56 objects, five features, and three classes [10].

In computer simulations, experimental results of K-means algorithm, GA-Clustering, TABU-Clustering, and KT-Clustering are obtained after a fixed number of iterations. For all experiments, a maximum of 1000 iterations is fixed. Each experiment includes 20 independent trials. However, in all experiments, K-means algorithm terminates much before 1000 iterations. The average and best values of the clustering results obtained by four methods, respectively, are shown as Table 1.

Table 1. Results of four clustering algorithms for five data sets

Algorithm	Data set 1 Avg(min)	Data set 2 Avg(min)	Ruspini Avg(min)	Iris Avg(min)	Crude Oil Avg(min)
K-means	2071.1(827.1)	488.8(488.1)	34793.0(12881.1)	91.8(78.9)	1653.2(1647.4)
GA-Clustering	9292.2(10465.4)	2532.8(2693.1)	75796.0(49333.5)	298.4(252.3)	2105.34(1722.4)
TABU-Clustering	8021.7(9789.9)	2526.3(2603.9)	70567.01(60774.4)	281.9(228.9)	2194.3(1747.5)
KT-Clustering	827.1(827.1)	488.0(488.0)	12881.1(12881.1)	78.9(78.9)	1647.2(1647.2)

For Data set 1, the optimal value is 827.1. K-means algorithm is found to attain this value in some trials. But, in most cases, it gets stuck in suboptimal values. KT-Clustering achieves the optimal value in all runs. Noticeably, GA-Clustering and TABU-Clustering fail to attain this value even once and their best values obtained are far worse than the optimal one. For Data set 2, KT-Clustering is obviously superior to other three approaches. Its optimal value is 488.0, which is found by only KT-Clustering in each trial. Since objects of this data set are overlapping, other three approaches cannot reach the optimal result in all runs.

For Ruspini and Iris, their optimal values are 12881.1 and 78.9, respectively. K-Means algorithm and KT-Clustering can attain their best values, but the average values of K-Means algorithm are far inferior to those of KT-Clustering, which shows K-means algorithm is not stable and liable to get stuck at locally optimal values. For Crude Oil, its optimal value is 647.2. Among four methods, only KT-Clustering can achieve the optimal value even in each trial. In face of different data sets, both GA-Clustering and TABU-Clustering cannot provide meaningful results within the pre-

scribed iterations. However, we find that these two algorithms can still obtain improved results if more iterations are executed. In order to understand the performance of four methods better, we use Crude Oil to show the iteration process. The clustering results of KT-Clustering, GA-Clustering, and TABU-Clustering are shown as Figure 3.
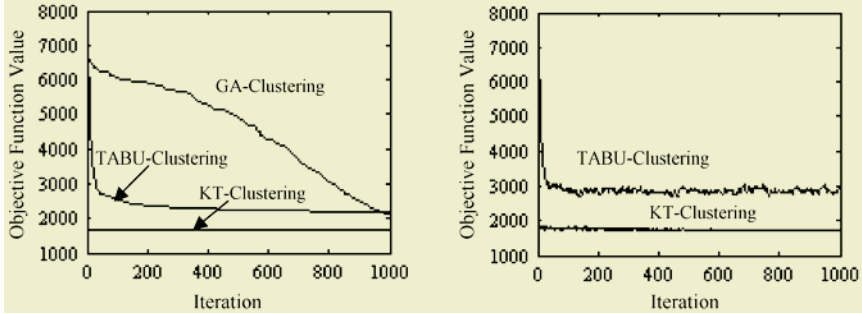
Fig. 3. The best value (L) and the current value (R) obtained in the process of iterations

From Figure 3, the best and current values of the clustering results obtained by KT-Clustering in the process of iterations are obviously the best. The reason is that, on one hand, K-means operation fine-tune the current solution so as to accelerate the convergence speed of the clustering algorithm, on the other hand, tabu search equipped with mutation operation can avoid the greedy charateristic of K-means algorithm and stably attain the optimal result. Therefore, KT-Clustering can obtain the optimal result faster and more stably than GA-Clustering and TABU-Clustering. For Crude Oil, with the same computational cost as TABU-Clustering, KT-Clustering spends average twelve iterations in attaining the best value, while TABU-Clustering cannot touch this value even once within 1000 iterations. In [3], the specified number of iterations where GA-Clustering obtains the best result for Crude Oil is up to 10000, which spends much more computational resources than KT-Clustering. Even so, there are still 22% of trials where it cannot obtain the optimal result. So, it is seen that it is not a good way to attain the optimal result by only adding the number of iterations. By computer simulations, it is shown that the combination of tabu search and K-means algorithm for clustering is meaningful and two new operations, K-means operation and mutation operation, greatly improve the performance of KT-Clustering.

4 Conclusions

In this paper, a hybrid tabu search based clustering algorithm called KT-clustering is proposed by integrating two new operations, K-means operation and mutation operation. The choice of algorithm parameters is extensively discussed. Performance comparisons between KT-Clustering and other three algorithms are conducted on experimental data sets. As a result, with the same computational cost as that of TABU-Clustering, KT-Clustering can get much better clustering results sooner than TABU-Clustering and GA-Clustering. Moreover, KT-Clustering obtains the optimal per-

formance in all trials, which shows it is more stable than other three algorithms. In future, the estimation of the number of clusters has to be incorporated in KT-Clustering, and different local search procedures and K-means variants should be tested within the tabu search framework.

Acknowledgements

References

1. Jain, A.K., Dubes, R.: Algorithms for clustering data. Prentice-Hall, New Jersey (1988)
2. Selim, S.Z., Ismail, M.A.: K-means-type algorithm: generalized convergence theorem and characterization of local optimality. IEEE Transactions on Pattern Analysis and Machine Intelligence. 6 (1984) 81-87
3. Murthy, C.A., Chowdhury, N.: In search of optimal clusters using genetic algorithms. Pattern Recognition Letters. 17 (1996) 825-832
4. Glover, F., Laguna, M.: Tabu search. Kluwer Academic Publishers, Boston (1997).
5. Al-sultan, K.S.: A tabu search approach to the clustering problem. Pattern Recognition. 28 (1995) 1443-1451
6. Chelouah, R., Siarry, P.: A hybrid method combining continuous tabu search and Nelder–Mead simplex algorithms for the global optimization of multiminima functions. European Journal of Operational Research. 161 (2005) 636-654
7. Schaffer, J.D., Caruana, R.A., Eshelman, L.J., Das, R.: A study of control parameters for genetic algorithms. In: Proceedings of the 3rd International Conference on Genetic Algorithms. Morgan Kaufmann (1989) 51-60
8. Kaufman, L., Rousseeuw, P.J.: Finding groups in data – An introduction to cluster analysis. John Wiley & Sons, New York (1990)
9. Fisher, R.A.: The use of multiple measurements in taxonomic problem. Annals of Eugenics. 7 (1936) 179-188
10. Johnson, R.A., Wichern, D.W.: Applied multivariate statistical analysis. Prentice-Hall, New Jersey (1982)

Neural Network Based Feedback Scheduling of Multitasking Control Systems

Feng Xia and Youxian Sun

National Laboratory of Industrial Control Technology,
Institute of Modern Control Engineering
Zhejiang University, Hangzhou 310027, China
{xia,yxsun}@iipc.zju.edu.cn

Abstract. To cope with resource constraints in multitasking control systems, feedback scheduling is emerging as an important technique for integrating control and scheduling. The ability of neural networks (NNs) as good and robust nonlinear function approximators, reducing the computation time as compared against algorithmic solutions, suggests applying them to the feedback scheduling problem. A novel, simple and intelligent feedback scheduler is presented using a feedforward NN. The algorithmic optimizer is utilized as a teacher to generate data samples for NN training. The role of the NN based feedback scheduler is to provide a good approximation to the optimal solution and online adjust the sampling period of each control task so that the overall system performance is maximized in the face of workload variations. The performance of the NN approach is evaluated through co-simulations of the scheduler, controllers and process dynamics.

1   Introduction

Recently, there is a trend towards integrated control and scheduling [1,2], where both the real-time computing community and the control community are involved. With rapid evolution of embedded computers for real-time control, it is increasingly common that some independent control tasks reside in the same processor, resulting in multitasking control systems. In these cases, the CPU time, which is always limited in an embedded environment, becomes the shared resource that several control tasks compete for. Consequently, the performance of multitasking control systems depends not only on the controller design, but also on the efficient scheduling of the shared computing resources. Traditional control systems are developed without much attention to the resource limitations, which only motivate real-time researchers. As these systems move from the research laboratories to real-world applications, they will inevitably suffer from resource constraints, which give rise to overload conditions in the shared processor. As a result, the performance of the control loops may be greatly degraded, and even destabilized sometimes.

   To address these requirements, a few efforts (e.g. [3-5]) have been made in the emerging field of feedback scheduling. When encountering workload variations, the feedback scheduler typically adjusts the sampling periods of concurrent control tasks so that the system remains schedulable and the overall control performance is optimized. However, the system performance is generally nonlinear as a function of the sampling period of each control loop. The optimization routine, which may corre-

sponds to constrained Newton optimization, inside a feedback scheduler can be very time consuming [5]. This is expensive for real-time control, and a computationally cheaper algorithm is desirable. Although approximation versions [3] of the optimal feedback schedulers are suggested in terms of linear proportional rescaling of the nominal sampling periods, they are somewhat empirical, and high accuracy of the resulting solutions cannot always be guaranteed.

In this paper, we present a neural network (NN) approach to the problem of feedback scheduling. Since new tasks can be activated or terminated, and computation time can vary due to the non-deterministic behavior of underlying platforms, the CPU workload may change over time. As one of the advanced technologies of recent times, neural networks [6,7] offer an exciting computational model. The ability of neural networks as good and robust nonlinear function approximators, reducing the computation time as compared with algorithmic solutions, suggests applying them to the feedback scheduling problem. We intend to exploit a simple and effective structure that uses a feedforward NN for feedback schedulers, demonstrating a novel application of NN at the same time. A back propagation (BP) NN trained with the Levenberg-Marquardt (LM) algorithm [6,8] is employed to dynamically adjust the sampling periods of control tasks. The existing algorithmic approach is used as a teacher to label the data samples for the NN training. Once well-trained offline, the NN is used to deliver almost optimal solution to the feedback scheduling problem so that the overall performance of all control loops is maximized whenever the CPU workload changes.

The rest of this paper is structured as follows. The problem of feedback scheduling is formulated in Section 2. We present the NN based feedback scheduler in Section 3. And its performance is evaluated via preliminary simulations in Section 4. In Section 5, we conclude this paper.

2   Problem Formulation

The system we consider is a multitasking control system where a set of controllers shares a CPU with limited computing capacity. As mentioned above, the role of feedback scheduling is to maximize the total control performance under task schedulability constraints. The feedback scheduler should control the CPU utilization at a desired level by dynamically adjusting the controllers' sampling periods. It is assumed that the execution times of all tasks and the workload are available from the real-time kernel at all times. From the control perspective, the scheduler can be viewed as a controller. The CPU utilization corresponds to the controlled variable and the manipulated variables are the sampling periods.

It has been pointed out in [3,5] that the feedback scheduling can be formulated as an optimization problem. Let the number of control tasks within the CPU be n, and let each task execute with a period h_i (sampling period) and have the execution time C_i. Each task is associated with a cost function $J_i(h_i)$, which gives the control cost as a function of sampling period h_i. Then the feedback scheduler should solve the problem

$$\min_h \ J = \sum_{i=1}^{n} J_i(h_i) \quad \text{s. t.} \ \sum_{i=1}^{n} C_i / h_i \le U_{ref} \tag{1}$$

where U_{ref} is the desired utilization level. It can be seen that the minimization routine of feedback scheduling constitutes a nonlinear programming (NLP) problem. To get linear constraints, the costs are recast as functions of sampling frequencies $\mathbf{f} = [f_1,...,f_n]^T$ instead of the sampling periods. Let $V_i(f_i) = J_i(1/h_i)$. According to the Kuhn-Tucker conditions [3,9], if $\mathbf{f}^* = [f_1^*, ...,f_n^*]^T$ is the optimal solution, then

$$\nabla \mathbf{V}(\mathbf{f}^*) + \lambda \mathbf{c} = 0$$
$$\lambda[U_{ref} - \mathbf{c}^T \mathbf{f}^*] = 0 \qquad\qquad (2)$$
$$\lambda \geq 0$$

where $\nabla \mathbf{V}$ is the gradient vector, $\mathbf{c} = [C_1, ..., C_n]^T$, and λ is the Lagrange multiplier.

Although there are several techniques to solving such a constrained NLP problem [9], e.g. Zoutendijk feasible direction algorithm and Frank-Wolfe method, solving the optimization problem exactly can be very time-consuming [5]. An algorithmic optimizer representatively involves calculation of the derivatives (gradients) or the Jacobian matrix, and a significant number of iterations are often needed. This will cause large computing overhead and hence holdbacks the application of existing NLP algorithms in feedback scheduling. In order for feedback scheduling to be practically useful, it is crucial that the scheduling overhead is relatively small. To address these problems, a simple neural network based approach is proposed in the next section.

3   Neural Network Approach to Feedback Scheduling

The reasons for the use of feedforward NN in feedback scheduling are twofold. First, in contrast to the algorithmic approach, simple feedforward NN based solutions can be delivered in a real-time fashion, reducing the scheduling overhead. Second, the neural networks offer very accurate solutions, e.g., for nonlinear programming problems like feedback scheduling, thanks to their powerful nonlinear approximation capacities. Therefore, we use a feedforward neural network to approximate the algorithmic optimizer and replace it in practical applications. In this case, training data is not a problem since it can be easily created off-line using existing algorithmic techniques. In this section, we describe the network architecture and training methods utilized.

A three-layer feedforward NN (with one hidden layer) is selected as a universal approximator to the NLP optimizer. The execution time C_i of each controller task and the desired utilization level U_{ref} are chosen as the network inputs, while the sampling periods are network outputs, see Fig.1. Accordingly, the numbers of network inputs and outputs are $n+1$ and n respectively. The principle behind the selection of network inputs/outputs is that all these variables are tightly associated with the CPU utilization level (according to basic real-time scheduling theory), which impacts greatly on the overall control performance of a multitasking control system (according to time-delay systems theory). The relationship between network inputs and outputs is given by

$$\mathbf{Y} = \mathbf{W}_2(\sigma(\mathbf{W}_1\mathbf{X} + \mathbf{B}_1)) + \mathbf{B}_2 \qquad\qquad (3)$$

where $\mathbf{W}_i$, $\mathbf{B}_i$ are the weight matrices of appropriate dimensions and bias vectors respectively, input vector $\mathbf{X}=[C_1, ..., C_n, U_{ref}]$, and output vector $\mathbf{Y}=[h_1, ..., h_n]$.
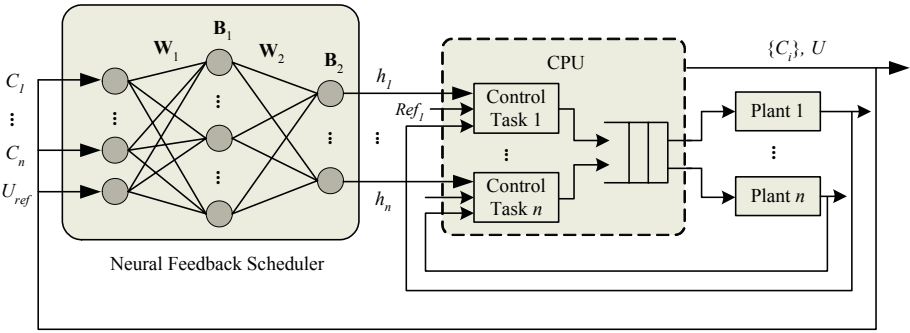
Fig. 1. Neural network based feedback scheduling

As we can see from equation (3), the computational complexity of the NN feedback scheduler mainly depends on the number of hidden neurons and the number of control tasks. In situations with large n, the NN may take a considerable time to yield an output. Fortunately, real-world applications always have limited number of concurrent control tasks that reside in one CPU. Therefore, as long as the number of hidden neurons is chosen properly, small scheduling overhead can be achieved. The activation functions used are the hyperbolic tangent sigmoid transfer function $\sigma(\bullet)$ in the hidden layer and the linear transfer function in the output layer.

To collect data samples for NN training, we solve the feedback scheduling problem with different C_i and U_{ref} as a NLP task. The cost functions are approximated as $V_i(f_i) = \alpha_i + \beta_i / f_i^2$. It has been shown in [3,5] that this is feasible especially for small sampling periods. The algorithmic optimization process may then be conducted using Matlab function *fmincon*. Ideally, the data set for NN training should cover the whole range of the variable execution time of each control task and that of the desired utilization level. After normalized, these data pairs are used as optimal solutions to train and validate the BP neural network. In this work, the neural network is trained offline using the LM algorithm, which appears to be the fastest method for training feedforward networks up to several hundred weights. Since the LM algorithm has an efficient Matlab implementation, we carry out the NN training using the Matlab function *trainlm*.

After it is well trained in supervised mode, the BP neural network can be used to provide a good approximation to the optimal solution for the feedback scheduling problem. Optimal sampling periods, which result in an optimal overall control performance, will be produced with respect to a set of given inputs. In this way, when the allowable CPU utilization or the execution time of a control task changes, the NN feedback scheduler will dynamically adjust the control periods, thus maximizing the overall control performance.

4   Performance Evaluation

We consider simultaneously control two inverted pendulums given by equation (4) with a single CPU, i.e., $n = 2$. The number of hidden neurons of the NN feedback scheduler is chosen to be 5. Every pendulum is controlled by one LQG controller

[10], which is designed using Matlab and remains the same for different simulations. For comparison, three different cases are considered during simulation studies based on TrueTime toolbox [11]: I) ideal case, where the execution of all tasks are disregarded, II) our solution with NN feedback scheduler, and III) traditional approach without any feedback scheduler. The execution times of both control tasks are varying. In addition to this type of workload variations, we also consider the fluctuation of the desired CPU utilization, which indicates executing or terminating other tasks within the same CPU. What the feedback scheduler concerns is to effectively manage these two types of workload uncertainties and to maximize the overall control performance through optimally adjusting the sampling periods of the two control loops.

$$G_1(s) = \frac{33.6}{s^2 - 33.6}, \quad G_2(s) = \frac{98}{s^2 - 98} \tag{4}$$

In all simulations, the control tasks are assigned fixed priorities. Furthermore, control task 1 is given the lowest priority and will thus suffer most during an overload. In case II, the feedback scheduler is given the highest priority. The execution times of the two control tasks vary in the same range of [8 10] ms, according to the uniform distribution. The nominal sampling periods are 33 and 20 ms respectively. At time t = 2s, U_{ref} changes from 0.8 to 0.5. The data samples for training the neural network is collected as follows: C_1 and C_2 range from 8 to 10 ms with a step of 0.2 ms respectively, while U_{ref} ranges from 0.4 to 1 with a step of 0.025. The corresponding optimal sampling periods are calculated with the help of Matlab. And then the feedforward NN is trained offline. The NN feedback scheduler is implemented as a periodic task with a period of 100 ms and the scheduling overhead is neglected.

Fig. 2 and 3 show the control performance of the two pendulums in different simulations. It is clear that the control performance is degraded when taking into account the execution of the control algorithms, in contrast to the ideal case. From t = 0 to 2s, as we can see from Fig. 2 and 3, two inverted pendulums perform well in all cases. During this time segment, the maximum requested CPU utilization is U_{max} = 10/33 + 10/20 = 0.803 ≈ 0.8. Since the available CPU utilization U_{ref} = 0.8, there is sufficient CPU time available for executing both control tasks. This explains the good perform-
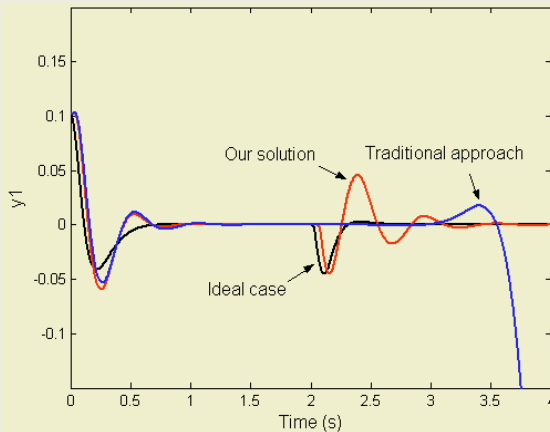
Fig. 2. Transient response of the first inverted pendulum
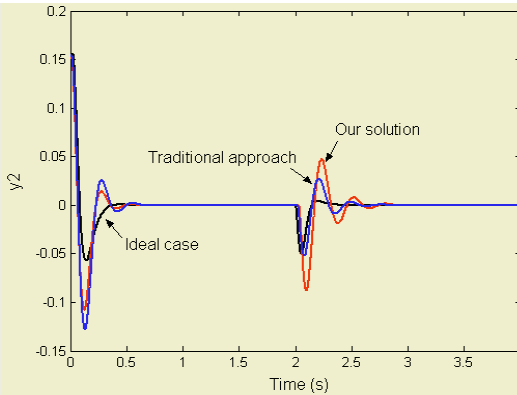
Fig. 3. Transient response of the second inverted pendulum

ance for both pendulums. From t = 2s, the CPU becomes overloaded because the minimum requested CPU utilization $U_{min} = 8/33 + 8/20 = 0.642 > U_{ref} = 0.5$. With traditional approach, the system becomes unstable and the first inverted pendulum falls down, because the control task 1 holds the lower priority. On the contrary, our solution with the NN feedback scheduler provides quite good performance all the time. One can notice that the first pendulum remains stable even when the computing resources become scarce. This is achieved at the expense of slight degradation of the second inverted pendulum's performance. In this way, the overall control perform-ance is maximized. The results demonstrate the practical ability of neural networks to effectively cope with resource limitation and workload uncertainty in multitasking control systems. In the face of overload, graceful degradation of the overall control performance can be achieved online.

5   Conclusions

From the viewpoint of integrated feedback control and real-time scheduling, the real-world resource limitations and workload uncertainties in multitasking control systems are considered. To address these problems, we demonstrate a novel and successful application of neural networks in the newly emerging field of feedback scheduling. The nonlinear approximation capability of feedforward NN is fully exploited. The proposed NN feedback scheduler exhibits several benefits including: 1) as a good approximation to the algorithmic optimizer, it delivers an almost optimal solution; 2) in contrast to the algorithmic approach, it is well-suited for real-time implementation with reduced computational complexity. Therefore, it can be utilized to dynamically maximize the overall control performance of a resource constrained multitasking control system. It is believed that the NN approach will play an increasingly impor-tant role in the convergence of control and scheduling in the near future.

References

1. K. E. Arzen, A. Cervin and J. Eker: An introduction to control and scheduling co-design, Proc. of the 39th IEEE CDC, Vol.5 (2000) 4865 - 4870

2. Feng Xia, Zhi Wang, and Youxian Sun: Integrated Computation, Communication and Control: Towards Next Revolution in Information Technology, Lecture Notes in Computer Science, Vol. 3356, Springer-Verlag, Heidelberg (2004) 117-125
3. A. Cervin, J. Eker, B. Bernhardsson, and K.-E. Årzén: Feedback-Feedforward Scheduling of Control Tasks, Real-Time Systems, 23:1 (2002)
4. K. Shin and C. Meissner: Adaptation of control system performance by task reallocation and period modification, IEEE Proc. 11th ECRTS, York, England (1999) 29-36
5. Johan Eker, Per Hagander, KarlErik Årzén: A feedback scheduler for real-time controller tasks. Control Engineering Practice, 8:12 (2000) 1369-1378
6. Hagan, M. T., H. B. Demuth, and M. H. Beale: Neural Network Design, Boston, MA: PWS Publishing (1996)
7. P. K. Campbell, et al.: Experiments with Simple Neural Networks for Real-Time Control. IEEE J. on Selected Areas in Communications, Vol. 15, No. 2 (1997) 165-178
8. Hagan, M. T., Menhaj, M.: Training Feedforward Networks with the Marquardt Algorithm, IEEE Trans. on Neural Networks, Vol. 5, No. 6 (1994) 989-993
9. Mokhtar S. Bazaraa, Hanif D. Sherali, C. M. Shetty: Nonlinear Programming: Theory and Algorithms, 2nd Edition, Wiley (1993)
10. K.J. Åström, B. Wittenmark: Computer Controlled Systems, Prentice Hall (1997)
11. Dan Henriksson, Anton Cervin and K.-E. Årzén,: TrueTime: Simulation of Control Loops Under Shared Computer Resources. In Proc. 15th IFAC World Congress on Automatic Control, Barcelona, Spain (2002)

An HILS and RCP Based Inter-working Scheme for Computational Evaluation of Manipulators and Trajectory Controller

Yeon-Mo Yang, N.P. Mahalik, Sung-Cheal Byun,
See-Moon Yang, and Byung-Ha Ahn

Department of Mechatronics, Gwangju Institute of Science and Technology
1 Oryong-dong, Buk-ku, Gwangju 500-712, Korea
{yangym,nmahalik,scbyun,smyang,bayhay}@gist.ac.kr

Abstract. We propose a "Hardware-In-the-Loop Simulation" (HILS) and "Rapid Control Prototyping" (RCP) based inter-working scheme for computational evaluation of manipulators and trajectory controller. The objective of the scheme is to minimize the development time in controller design and reduce the efforts in body modeling analysis, especially in the field of robotics. We have derived the analytical model of two manipulators for HILS and implemented an output controller for RCP; then two functional blocks are integrated to verify and validate the usefulness and possibility of HILS and RCP based inter-working (HRI) scheme. Experimental results show that the proposed HRI scheme is an effective means to apply for modeling the two elbow manipulators and implementing the selected controller for the manipulator applications in an early stage of the development cycle. Furthermore, HRI scheme has shown to be particularly beneficial for body dynamics analysis and manipulator control tasks.

1 Introduction

Various computational evaluation methods of observer-based controllers have been considered for the elbow manipulators with respect to modeling and controller design [1, 10]. Most of the models derived till now are for serial-type manipulators, and the associated controllers require a priori knowledge of the reference trajectory that is also uniformly bounded over different physical implementation [2]. In many applications, however, this is not the case considering as follows: first, parallelogram-type is preferred to serial-type manipulator due to the required motor torque variations; second, the reference trajectory is an output of a reference model subjected to numerous input functions rather than a fixed function of time [3]. Typical approaches for manipulator modeling and trajectory controller are the standard real-time trajectory tracking and the use of a first-order linear dynamic equation (reference model) that generates a field of trajectory [4-5, 11].

Ever more, few considerations have been made with respect to the industrial application and simulation of these controllers [9-11]. Some aspects related to restrictions arising in real industrial applications have been considered in detail. To the best of our knowledge, the issue of digital controller and the relevant practical implementation problems of two manipulators such as serial and parallelogram types have not been considered in the available literature [8-11]. The rationale that arises is to intro-

duce the modeling of parallelogram type manipulator and the computational evaluation of a trajectory controller that guarantees the asymptotic convergence of the robot output to the corresponding model output for every reference model input. The objective of the research is to exploit two developments. One is to derive the models of these two types of elbow manipulators and the other is to examine and investigate some of the possible implementation problems that occur in real-time applications using HRI scheme. The manipulator models will be realized as Hardware-In-the-Loop Simulation (HILS) and the controller will be implemented by Rapid Control Prototyping (RCP); then we will integrate the two functional blocks, called HRI scheme.

The rest of the paper is organized as follows. In section 2, the governing model equations of the serial type and parallelogram type elbow manipulators are derived for HILS. In section 3, we present and implement the trajectory controller for RCP using dSPACE® with MATLAB®. The experimental results of two elbow manipulators and the controller by HRI scheme are verified and demonstrated in Section 4. Finally, summary and conclusion are given in the end.

2   Modeling the Elbow Manipulators for HILS

We drive two separate equations for modeling the serial type and parallelogram type elbow manipulators. The modeling is based on the assumptions that the links are rigid and masses of link are uniformly distributed. The equations of motion of the manipulators are obtained in terms of a state-space variable and will be presented in detail as follows.

(a) Serial type   (b) Parallelogram type
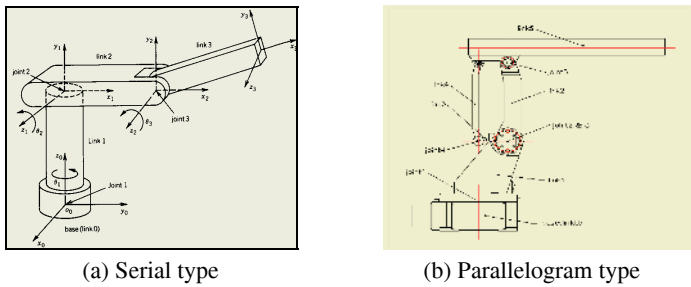
Fig. 2.1. Two elbow manipulators

The serial type elbow manipulator is shown in Fig. 2-1 (a) and its equation of motion is of the form,

$$D(q(t))\ddot{q}(t)+C(q(t),\dot{q}(t))\dot{q}(t)+\phi(q(t))=\tau(t),\ t\geq 0, \tag{2.1}$$

where $q \in R^n$ is the vector containing the joint displacements, $D(q)$ is the inertia matrix with $D(q)=D(q)^T>0$, for all $q \in R^n$, $C(q,\dot{q})\dot{q}$ represents the Coriolis and centrifugal forces, $g(q)$ represents the effect of the gravitational force, and $\tau \in R^n$ is the vector of the applied torques. We assume that the output variable of the robot dynamic system (2.1) is the joint displacements vector q; i.e., we consider the output map $y(t)=q(t)$.

Thus, the state space representation of two type manipulators [7] is given by

$$\dot{x} = f(x) + \hat{D}(x)u, \quad y = Ex \tag{2.2}$$

where

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in IR^{2n}, \quad f(x) = \begin{bmatrix} x_2 \\ D^{-1}(x_1)(-C(x_1,x_2)x_2 - \phi(x_1)) \end{bmatrix}, \quad x_1 = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}, \quad x_2 = \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix}. \tag{2.3}$$

$$\hat{D}(x) = \begin{bmatrix} 0 \\ D(x_1)^{-1} \end{bmatrix}, \quad E = [I_n, 0]. \tag{2.4}$$

3 Controller Design for RCP

The structure of the output controller is presented. To do this, there is a requirement to draw up the basic problem statement with sufficient assumptions. The problem to be solved is: for given the robot dynamic model and the linear reference model (exosystem), let us find an output feedback control law such that the resulting robot output (y) converges asymptotically to the corresponding reference model output y_R, i.e.,

$$\lim_{t \to \infty} e_R(t) = \lim_{t \to \infty} (y(t) - y_R(t)) = 0. \tag{3.1}$$

In addition, the control objective of the output controller is to ensure internal stability of the observer, i.e. the combined system comprising the plant and controller (excluding the reference model) must be stable, so that the corresponding state error vector tends to zero when no tracking is present.

Let x_1^*, x_2^*, and u^* be defined as follows: $x_1^* := C\zeta = y_R$, $x_2^* := CN\zeta = \dot{y}_R$, and $u^* := CN^2\zeta + CNSw = \ddot{y}_R$ where C, N and S the matrices defined in the reference model. It should be seen that, for every admissible w, the state components are

$$\begin{aligned} \dot{x}_1^* &= x_2^*, \\ \dot{x}_2^* &= u^* \end{aligned} \tag{3.2}$$

with $x_1^*(0) = y_R(0) = C\zeta(0)$, $x_2^*(0) = CN\zeta(0) = \dot{y}_R$.

If a controller-observer is synthesized, such that the resulting state trajectory $x(t)$ of the robot dynamic model converges asymptotically to the state

$$x^*(t) = \begin{bmatrix} x_1^*(t) \\ x_2^*(t) \end{bmatrix} \tag{3.3}$$

of the linear reference model, then it guarantees that the output $y(t)$ of the robot converges asymptotically to the output $y_R(t)$. Finally, it should be noticed that the synthesis approach is based on the assumption that an upper bound on the robot initial velocity norm is available beforehand.

Due to stability criterion, the following observer-based controller which is based on [6-7] is utilized:

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2 + F^*(t)y - R(\hat{x}_1 - y), \\ \dot{\hat{x}}_2 &= u^* - L(\hat{x}_1 - x_1^*) - F^*(t)(\hat{x}_2 + F^*(t)y) \\ &\quad - A(\hat{x}_2 + F^*(t)y - x_2^*) - 2(x_2^{*T}u^*)Gy, \end{aligned} \tag{3.4}$$

with the output

$$\hat{y} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_1 + F^*(t)\,y \end{bmatrix} \tag{3.5}$$

where the $n \times n$ time-dependent matrix $F^*(t)$ is defined as $F^*(t) := B + (x_2^{*T}(t)x_2^*(t))G$.
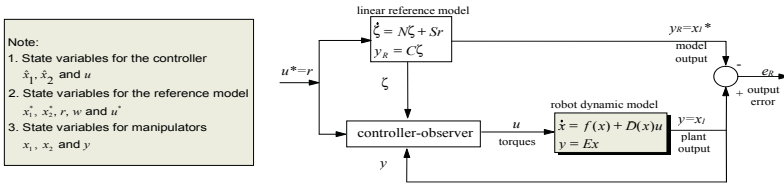
Fig. 3.1. Structure of the controller for robot model matching

Therefore, the input to the robot (the vector of the torques) is given by Eq. (3-6) and the structure of the proposed controller is shown in Fig. 3-1. While the controller is nonlinear, by noting that the functions $x^*(\cdot)$ and $u^*(\cdot)$ are obtained from the linear reference model, it appears that its dynamic part is a linear time-varying system with continuous coefficients and piecewise continuous input function.

$$u = g(y) + C(y, \hat{y}_2)\hat{y}_2 + D(y)[u^* - L(\hat{x}_1 - x_1^*) - K(y - \hat{x}_1) - A(\hat{x}_2 + F^*(t)y - x_2^*)]. \tag{3.6}$$

A constructive proof for the performance of the proposed controller as well as a detailed description of its parameters will not be discussed. Generally, the theoretical proof deals with the concepts of exponential and semi-global stability. In Fig. 3-1, the reference model is to be a double integrator ($u^* = \ddot{y}_R = w$) if we select the matrices C, N and S so that the reference input w is exactly the *acceleration* of the control input. The reference *position* is $x_1^* = y_R$ and the *velocity* is $x_2^* = \dot{y}_R$.

The matrices A, B, G, K, L, R are symmetric constant matrices defined as $K := I_n$, $R := rI_n$, $G := gI_n$, $B := R + \beta G$ and A and L are selected arbitrarily, though the convergence behavior is greatly dependent on this selection. The coefficients r, g, and β are determined by the robot parameters and the upper bound of the reference input r. Finally, it should be outlined again that the initial conditions of the controller are selected arbitrarily, whereas x_1 is a measurable signal (position). Also, the initial conditions for the reference model should be known beforehand.

4   Simulations and Evaluation

Experiments were performed on two elbow manipulators described in Section 3 to verify and validate the applicability of the proposed inter-working scheme. The scheme is implemented on TMS320® based dSPACE system for HILS and the on-board Alpha-processor® based dSPACE system for RCP, respectively. We realized and implemented the output controller by the use of dSPACE integration tool. This is a DSP board hardware platform and is compatible to MATLAB. The advantages using dSPACE in deploying HRI concept are as follows: the flexibility to modify due to change in the algorithm, the supporting real time systems, and the computational power for the complex dynamic systems.

(a) Serial type

(b) Parallelogram type
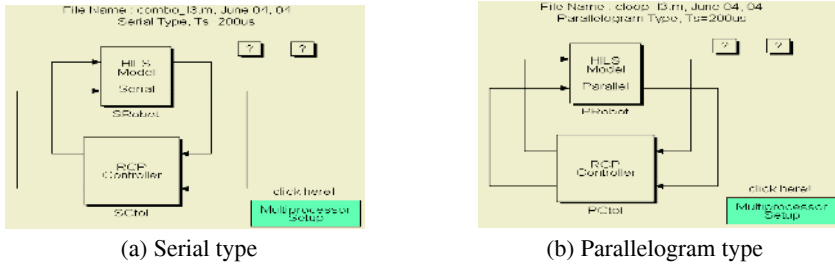
Fig. 4.1. HRI model for Real-time Evaluation

The MATLAB Simulink block diagram of the serial and parallel type elbow manipulators with the controller is shown in Fig. 4-1. The evaluation block consists of two parts: one is the manipulator model implemented on TMS320 dSPACE as in HILS and the other is the output controller realized on Alpha dSPACE as of RCP.

Table 4.1. The physical parameters of two type manipulators

(a) serial

The arms' masses (m)			The arms' lengths (l)		
m_1	m_2	m_3	l_1	l_2	l_3
10kg	5kg	10kg	1m	0.6m	0.8m

(b) parallelogram

The arms' masses (m)					The arms' lengths (l)				
m_1	m_2	m_3	m_4	m_5	l_1	l_2	l_3	l_4	l_5
10kg	10kg	2kg	16kg	5kg	0.1m	0.4m	0.2m	0.4m	0.4m

Table 4.2. The tuned parameters of the controller

Factor	Values	Definition and guideline in choosing the value
η	4.069	Inertia Matrix, $\|D(a)\| \le \alpha$, $\|D(q)^{-1}\| \le \eta$
Kc	3.841	Coriolis and Centrifugal, $K_c - \max \sum \|C_q(q)\|$
γ	31.270	$2\,\eta\,Kc$
g	32.000	$g > \gamma$
δ	10.000	$\delta = v$
v	10.000	$x_2^{*T}(t)\hat{x}_2(t) \le \delta \exp(vt)$
μ	40.900	Positive constant
σ	15.500	$\|\exp(Mt)\| \le \mu \exp(-\sigma t)$
r_e	2.000	Closed ball, $r_e = \sqrt{(\|v_d\| - r_e)^2 + (r_e + r_e)^2}$
φ	2.000	Any $\varphi > 1$
β	182.000	$\beta > 0.25 + 2\sqrt{5}\mu\varphi r_e$
r	254.000	$r \ge (g(\beta + \delta) + (\sigma + v)\varphi)/(\varphi - 1)$

Tables 4-1 (a) and (b) show the evaluation parameters of the serial and parallelogram type elbow manipulators. Considering the variations of total kinetic and potential energy of two types of manipulators, the link lengths and masses has chosen to provide the same condition for the evaluation as shown in the tables. The detailed of tuned parameters of the controller are shown in Table 4-2; and the selected initial conditions of the state equations and the linear model reference input $w(t)$ are as follows:

$$x_1(0) = x_2(0) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \; x_1^*(0) = x_2^*(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \; \hat{x}_1(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \; \hat{x}_2(0) = \begin{bmatrix} -0.7008 \\ -1.4016 \\ -2.1024 \end{bmatrix}, \; w(t) = \begin{bmatrix} 2\sin(5t) \\ 4\sin(10t) \\ 8\sin(15t) \end{bmatrix} u_s(t)$$

where $u_s(t)$ is the unit step function.

As a part of RCP, the three signals of the desired reference and actual output trajectories are shown in Fig. 4-2 (a) and (b), respectively with tuned control parameters. The convergence of the output error function to zero is observed in the manipulator.

(a) reference   (b) actual output
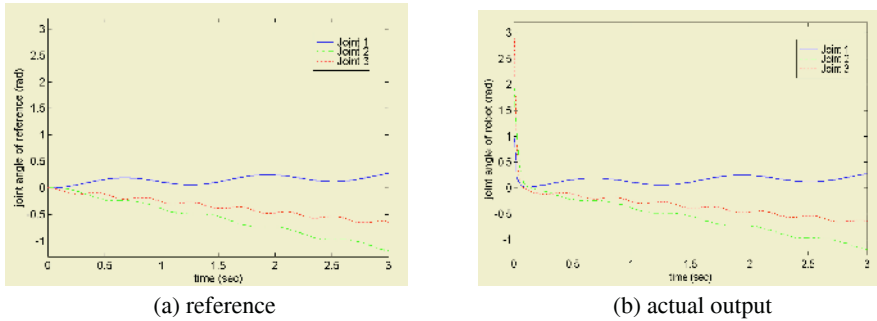
Fig. 4.2. Trajectory curves of each joint of the manipulator

The comparison of the execution time of two manipulators in tracking is illustrated in Fig. 4-3 (a). The upper side of the figure shows that there involves the longer execution time for the serial type elbow manipulator running on dSPACE. In particular, the total executing time for the serial type elbow is 25us on the average; whereas for the parallelogram type is 15us. This unwanted result is caused due to the complexity of the model.

(a) execution time   (b) total required torque
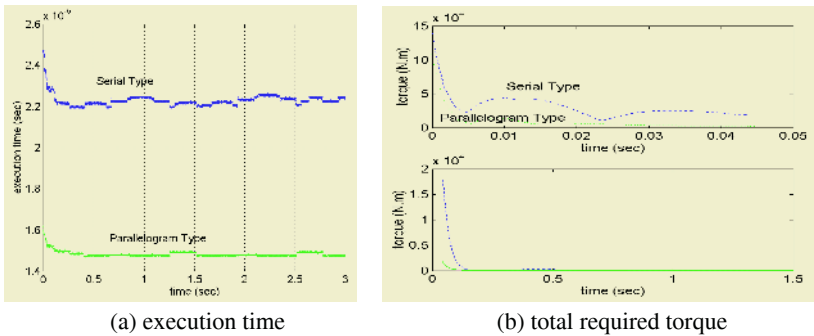
Fig. 4.3. The comparison of two manipulates in execution time and torque variation

By HRI scheme, we have the following results. The overall torques required to get the desired trajectory of two manipulators is shown in Fig. 4-3 (b). It is convincing from the result that the computational time of the parallelogram type manipulator is much less than that of the serial manipulator. Apparently, the parallel manipulator is more effective and efficient in realizing the real robot systems.

The above results suggest that the proposed HRI scheme is very effective and efficient in analyzing the system dynamics and designing the controller at an early stage of the development cycle. This will not only accelerate the design efforts but also improve the quality of design significantly

5   Conclusion

We have presented the HRI scheme for the computational evaluation of manipulators and trajectory controller. Specifically, we show the effectiveness and possibility of the proposed HRI scheme with the development of output tracking control of two elbow manipulators. It has shown how to apply the proposed HRI scheme to two different well-known robot manipulators, namely the serial type and the parallelogram drive type manipulators. The real facts of trajectory controllers and the practical considerations for two different types of manipulators are analyzed and discussed. Experimental results demonstrated that the total execution time of the controller and the torques required by the parallelogram type manipulator to get the desired position and velocity are smaller than that of the serial type manipulator considerably. So, it suggests that the parallelogram body configuration is more suitable when considering the required torques and computational powers.

Through the experiments, it implies that HRI are very useful for the analysis of robot dynamics and especially for the evaluation of the designed model and controller taking into accounts of not only set-point control, but also output tracking control. An experimental consideration to implement the proposed control architecture on a rigid link flexible joint (RLFJ) manipulator is of course a subject for further study. This scheme can also be applied to adaptive and optimal controller strategy.

References

1. Canudas De Wit, B., Siciano, and Bastin G.: Theory of Robot Control, Springer (1996)
2. Nicosia, S. and Tomei, P.: Robot Control by Using Only Joint Position Measurements, IEEE Transactions on Automatic Control. Vol. 35 (1990) 1058-1061
3. Isidori, A.: Nonlinear Control Systems, (2nd Edition), Springer, Berlin, Germany (1989)
4. Slotine, J.-J.E., Li, W.: Applied Nonlinear Control, Prentice-Hall, New Jersey (1991)
5. Spong, M.W. and Vidyasagar, M.: Robot Dynamics and Control, John Wiley and Sons, New York (1989)
6. A. Ailon and R. Segev: Stable Observer-Based Trajectory Controller for Asymptotic Model Matching of a Rigid Robot, Journal of Optimization Theory and Applications, Vol.87 No 3, December (1995) 517-538
7. Ilion A. Bonev, Y.M. Yang, and A. Ailon: Output Controller for Robot Trajectory Tracking System, Mechatronics Project Report, K-JIST, June 14 (1997)
8. H. Hanselmann: Hardware-in-the-Loop Simulation and its Integration into a CACSD Toolset, IEEE International Symposium on Computer-Aided Control System Design, Dearborn, Michign, September 15-18 (1999)
9. J-W Jeon, G-A Woo, K-C Lee, D-H Hwang and Y-J Kim: Developments of ABS controller for aircraft with real-time HILS system, Electrical Machines and Systems, ICEMS 2003. 6th, Vol. 2, Nov. 9-11, (2003) 498-501
10. D. Wang and G. Xu: Full-state tracking and internal dynamics of non-holonomic wheeled mobile robots, Mechatronics, IEEE Transactions on, Vol. 8, June (2003) 203-214
11. Roy, J., Goldberg, R.P., and Whitcomb, L.L.: Structural design, analysis, and performance evaluation of a new semi-direct drive robot arm: theory and experiment, Mechatronics, IEEE/ASME Transactions on, Vol.: 9, Issue: 1, (2004) 10-19

Modeling of Nonlinear Static System
Via Neural Network Based Intelligent Technology

Dong-Won Kim[1], Jang-Hyun Park[2], Sam-Jun Seo[3], and Gwi-Tae Park[1]

[1] Department of Electrical Engineering, Korea University, 1, 5-ka
Anam-dong, Seongbuk-ku, Seoul 136-701, Korea
{upground,gtpark}@korea.ac.kr
[2] Department of Control System Engineering, Mokpo National University, 61
Torim-ri, Chonggye-myon, Muan-gun, Chonnam, Korea
jhpark72@mokpo.ac.kr
[3] Department of Electrical & Electronic Engineering, Anyang University, 708-113
Anyang 5dong, Manan-gu, Anyang-shi, Kyunggi-do, 430-714, Korea
ssj@anyang.ac.kr

Abstract. Modeling of nonlinear static system using neural network based intelligent technology is presented in this paper. The architecture of the intelligent system is combined neural network with polynomial neural network. The composite architecture is designed to get a heuristic approximation method for nonlinear static system modeling. Owing to the approximation capabilities, neural networks have been widely utilized to process modeling, whereas the polynomial neural network is an analysis technique for identifying nonlinear relationships between inputs and outputs of the target system. So the hybrid architecture can harmonize the advantages of the each modeling methodology. Simulation results of the intelligent technology will be shown efficient and good performance.

1 Introduction

Neural networks (NNs) are designed in an attempt to mimic the human brain in order to emulate human performance and thereby function intelligently. NNs have been used in numerous application, many of which have involved either classification of given observations or approximation of the observations' generating function [1-2]. To arrive at a suitable solution for such problems, neural networks use both back propagation of error [3] and the gradient descent technique to optimize an objective function. NNs can be trained by examples and sometimes generalize well for unknown test cases. The worthiness of a network lies in its inferencing or generalization capabilities over such test sets. Connectionist learning procedures are suitable in domains with several graded features that collectively contribute to the solution of a problem. In the process of learning, a network may discover important underlying regularities in the task domain. An advantage of NNs lies in the high computation rate provided by their massive parallelism [4].

Another modeling technique is the polynomial neural network (PNN) [5], based on the GMDH algorithm [6]. The characteristics of the PNN is that it provides an automated selection of essential input variables and builds hierarchical polynomial regressions, partial description (PD), of required complexity. In addition, high-order regression often leads to a severely ill-conditioned system of equations. However, the PNN

avoids this problem by constantly eliminating variables at each layer. Therefore, complex systems can be modeled without specific knowledge of the system or massive amount of data. It is goal of this paper to develop and validate integrating architecture by combining NNs and PNN to find workable synergistic environment and harmonize its own individual advantages. Simulation results shows that the proposed method is much more accurate so it can be considered as efficient modeling architecture.

2   Neural Network Based Intelligent Technology

Neural network based intelligent architecture which combined NNs with PNN is depicted in this section. With reference to the NNs, The NNs [2] consist of multiple layers of simple, two-state, sigmoid processing elements or neurons that interact using weighted connections. After a lowermost input layer there are usually any number of intermediate, or hidden, layers followed by an output layer at the top. There exist no interconnections within a layer while all neurons in a layer are fully connected to neurons in adjacent layers. Weights measure the degree of correlation between the activity levels of neurons that they connect [4].

   One important aspect in a NN development procedure is the learning process. Among the various NN paradigms, the feedforward NN with the back-propagation algorithm is employed. The generalized delta rule is applied for adjusting the weights of the NN to minimize a predetermined cost error function. The rule of adjusting weights is given by the Eqs. (1)-(2)

$$w_{ij}^p(t+1) = w_{ij}^p(t) + \eta \delta_j^p y_j^p + \alpha \Delta w_{ij}^p(t) \tag{1}$$

$$\delta_j = (d_j - y_j) f_j'(\cdot) \tag{2}$$

where w_{ij}^p denotes the connection weight between the i-th node in the layer (p-1) and the j-th node in the layer p. y_j^p is the output of the j-th node in the layer p. η and α are the learning rate and momentum term, respectively. δ is the negative derivative of the total square error with respect to the neuron's output. d_j and $f_j'(\cdot)$ are the desired output and the derivative of the activation function of j-th node, respectively.

   Since the PNN is applied to the hybrid architecture, its fundamentals are briefly explained. The PNN algorithm is based on the GMDH and utilizes a class of polynomials such as linear, quadratic, and modified quadratic types. For an example, specific forms of a PD in the case of two inputs are given as

$$\text{Type 1} = c_0 + c_1 x_1 + c_2 x_2 \tag{3}$$

$$\text{Type 2} = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_1^2 + c_4 x_2^2 + c_5 x_1 x_2 \tag{4}$$

$$\text{Type 3} = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_1 x_2 \tag{5}$$

where c_i is called regression coefficients.

   These polynomials are called as partial descriptions (PDs). By choosing the most significant input variables and polynomial types among various types of forms available, we can obtain the PDs in each layer. The PNN is developed to identify the model of nonlinear complex systems by the use of the input-output data set. This data

set is divided into two parts, that is, training and testing data set. The total number of nodes is given by a combination of a fixed number of inputs among entire input variables. The number of input variables and types of polynomial of each node are therefore determined in advance by designer. By using the chosen input variables and type corresponding to each node, we construct a PD for each node. We determine the coefficients of PD by the least square method by using a given training data set, and finally we obtain the estimated output of each node. Furthermore, we evaluate each PD to check its predictive capability for the output variable using the testing data set. Then we compare these values and choose several PDs which give the best predictive performance. In the sequel, we construct the second layer in the same way, considering the output variable of each chosen node in the first layer as the new input variables to the second layer. Afterwards, we repeat this procedure until the stopping criterion has been satisfied. Once the final layer has been constructed, only the one node characterized by the best performance is selected as the output node. The remaining nodes in that layer are discarded. Furthermore, all the nodes in the previous layers that do not have influence on the selected output node are also removed by tracing the data flow path on each layer. And finally, the PNN model is obtained.
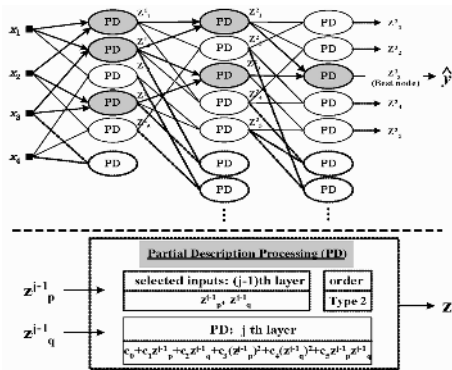
Fig. 1. Typical architecture of the PNN

The overall architecture of the PNN is shown in Fig. 1. In the figure, four input variables $(x_1, \ldots, x_4)$, three layers, and a PD processing example are considered. Z^{j-1}_i indicates the output of the ith node in the j-1th layer, which is employed as a new input of the jth layer. The black nodes have influence on the best node (output node), and these networks represent the ultimate PNN model. Meanwhile, the solid line nodes have no influence over the output node. In addition, owing to poor performance, the dotted line nodes are excluded when choosing the PDs with the best predictive performance in the corresponding layer. Therefore, the solid line nodes and dotted line nodes should not be present in the final PNN model.

The proposed model dwells on the ideas of combining NNs with the PNN. The hybrid architecture of the proposed method is shown in Fig. 2. NNs and PNN merged together with cascade connection for the nonlinear system. In Fig. 2, the z_i $(i=1, \ldots 4)$ is used as input variable to the next PNN. z_i is the product of w_{ij} which connection weight from i-th hidden layer to j-th output layer of NN and node output of the hidden layer.
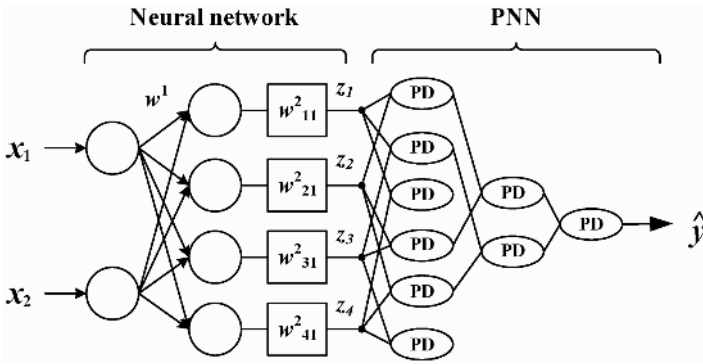
Fig. 2. Architecture of the hybrid neural network

3   Simulation Results

Nonlinear static system with two inputs and a single output which has been widely used to evaluate modeling performance [7-10] is considered. This nonlinear system represents the nonlinear characteristic as shown in Fig. 3. Using equation (6), the inputs are generated randomly and the corresponding output is then computed through the above relationship. Table 1 shows 50 pairs of the input-output data obtained from expression (6). The inputs x_3 and x_4 are dummy variables which has no relation to (6). The performance index is defined as mean squared error by (7)

$$y = (1 + x_1^{-2} + x_2^{-1.5})^2, \ 1 \le \ x_1, \ x_2 \le 5 \tag{6}$$

$$\mathrm{PI} = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2 \tag{7}$$

with y_i being the actual output, $\hat{y}_i$ forming the estimated output, and m is the number of data.
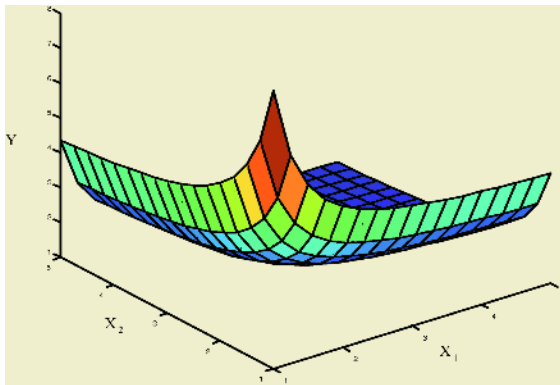
Fig. 3. Input-output relation of the two-input nonlinear system

Table 1. Input-output data of three-input nonlinear function

	Data (1-25)					Data (26-50)				
No.	x_1	x_2	x_3	x_4	y	No.	x_1	x_2	x_3	x_4	y
1	1.40	1.80	3.00	3.80	3.70	26	2.00	2.06	2.25	2.37	2.52
2	4.28	4.96	3.02	4.39	1.31	27	2.71	4.13	4.38	3.21	1.58
3	1.18	4.29	1.60	3.80	3.35	28	1.78	1.11	3.13	1.80	4.71
4	1.96	1.90	1.71	1.59	2.70	29	3.61	2.27	2.27	3.61	1.87
5	1.85	1.43	4.15	3.30	3.52	30	2.24	3.74	4.25	3.26	1.79
6	3.66	1.60	3.44	3.33	2.46	31	1.81	3.18	3.31	2.07	2.20
7	3.64	2.14	1.64	2.64	1.95	32	4.85	4.66	4.11	3.74	1.30
8	4.51	1.52	4.53	2.54	2.51	33	3.41	3.88	1.27	2.21	1.48
9	3.77	1.45	2.50	1.86	2.70	34	1.38	2.55	2.07	4.42	3.14
10	4.84	4.32	2.75	1.70	1.33	35	2.46	2.12	1.11	4.44	2.22
11	1.05	2.55	3.03	2.02	4.63	36	2.66	4.42	1.71	1.23	1.56
12	4.51	1.37	3.97	1.70	2.80	37	4.44	4.71	1.53	2.08	1.32
13	1.84	4.43	4.20	1.38	1.97	38	3.11	1.06	2.91	2.80	4.08
14	1.67	2.81	2.23	4.51	2.47	39	4.47	3.66	1.23	3.62	1.42
15	2.03	1.88	1.41	1.10	2.66	40	1.35	1.76	3.00	3.82	3.91
16	3.62	1.95	4.93	1.58	2.08	41	1.24	1.41	1.92	2.25	5.05
17	1.67	2.23	3.93	1.06	2.75	42	2.81	1.35	4.96	4.04	1.97
18	3.38	3.70	4.65	1.28	1.51	43	1.92	4.25	3.24	3.89	1.92
19	2.83	1.77	2.61	4.50	2.40	44	4.61	2.68	4.89	1.03	1.63
20	1.48	4.44	1.33	3.25	2.44	45	3.04	4.97	2.77	2.63	1.44
21	3.37	2.13	2.42	3.95	1.99	46	4.82	3.80	4.73	2.69	1.39
22	2.84	1.24	4.42	1.21	3.42	47	2.58	1.97	4.16	2.95	2.29
23	1.19	1.53	2.54	3.22	4.99	48	4.14	4.76	2.63	3.88	1.33
24	4.10	1.71	2.54	1.76	2.27	49	4.35	3.90	2.55	1.65	1.40
25	1.65	1.38	4.57	4.03	3.94	50	2.22	1.35	2.75	1.01	3.39

In what follows, experiment of the hybrid neural network architecture is conducted. The results are summarized in the tables. In Table 2, the initial values of the neural network parameters are shown. After 500 epochs with 4 hidden neurons, the final result of the NN was 3.3009 but, PNN is combined with the NN, the performance of the hybrid architecture is remarkably improved. For the PNN, 2 inputs in the 1st layer and 4 inputs in the 2nd layer or higher are employed and Type 1, 2, and 3 in the 1st layer and Type 2 in 2nd layer or higher are used. These results are listed in Table 3.

Table 2. Initial parameter of neural networks

Parameters	Neural networks
Learning rate	0.1
Momentum term	0.0123
Maximum training epochs	500
Number of hidden unit	4

Identification performance is shown in Fig. 4. The model output follows the actual output very well. Where, the values of the performance index of the proposed method are equal to 0.00022.

Table 4 shows the performance of our proposed method and other models studied in the previous literature. The experiment results show that our intelligent technology offers encouraging advantages and has good performance.

Table 3. Results of the neural network based intelligent architecture

Neural networks	PNN inputs		PNN order		PI
PI	1st layer	2-5 layer	1st layer	2-5 layer	PI
			Type 1	Type 2	0.00143
3.3009	2	4	Type 2	Type 2	0.00022
			Type 3	Type 2	0.00056
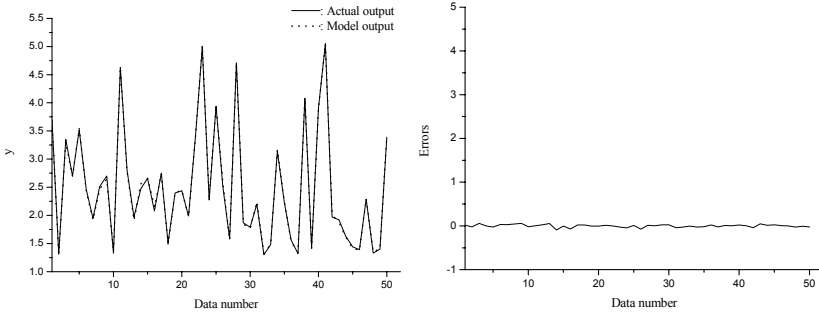
(a) actual output versus model output   (b) error

Fig. 4. Actual output versus model output and its error

Table 4. Comparison of identification performance with some previous models

Model	MSE PI
Sugeno [7]	0.079
	0.010
Kim [8]	0.0089
Gomez-Skarmeta [9]	0.070
Lin [10]	0.0035
Our model	0.0002

4   Conclusions

In this paper, we present neural network based intelligent architecture for modeling of nonlinear static system. The intelligent technology is dwelt on the ideas of combining NNs with the PNN into one methodology. In the simulation studies, hybrid model is applied to the nonlinear static system and it is proved that the proposed modeling technique is a sophisticated and versatile architecture capable of constructing models out of a limited data set and produces superb results.

References

1. Antognetti,P., Milutinovic, V.: Neural Networks: Concepts, applications and implementations. Vols. 1-4. Prentice Hall, New Jersey (1991)
2. Rumelhart, D.E., McClelland, J.L.: Parallel Distributed Processing. Vol. 1. MIT Press, Cambridge MA (1986)

3. Murthy, C. A., Pittman, J.: Multilayer perceptrons and fractals. Inf. Sci. 112 (1998) 137-150

4. Pal, S. K., Mitra, S.: Multilayer Perceptron, Fuzzy Sets, and Classification. IEEE Trans. Neural Networks 3 (1992) 683-697

5. Oh, S.K., Pedrycz, W.: The design of self-organizing Polynomial Neural Networks. Inf. Sci. 141 (2002) 237-258

6. Ivakhnenko, A.G.: Polynomial theory of complex systems. IEEE Trans. Syst. Man Cybern. 1 (1971) 364-378

7. Sugeno,M., Yasukawa, T.: A Fuzzy-Logic-Based Approach to Qualitative Modeling. IEEE Trans. on Fuzzy syst. 1(1993) 7-31

8. Kim, E., Lee, H., Park, M., Park, M.: A simply identified Sugeno-type fuzzy model via double clustering. Inf. Sci. 110 (1998) 25-39

9. Gomez-Skarmeta, A.F., Delgado, M., Vila, M.A.: About the use of fuzzy clustering techniques for fuzzy model identification. Fuzzy Sets Syst. 106 (1999) 179-188

10. Lin, Y., Cunningham III, G. A.: A New Approach to Fuzzy-Neural System Modeling. IEEE Trans. on Fuzzy syst. 3 (1995) 190-198

Bayesian Inference Driven Behavior Network Architecture for Avoiding Moving Obstacles*

Hyeun-Jeong Min and Sung-Bae Cho

Dept. of Computer Science, Yonsei University
134 Shinchon-dong, Sudaemoon-ku, Seoul 120-749, Korea
{solusea,sbcho}@cs.yonsei.ac.kr

Abstract. This paper presents a technique for an intelligent robot to adaptively behave in unforeseen and dynamic circumstances. Since the traditional methods utilized the relatively reliable information about the environment to control intelligent robots, they were robust but could not behave adaptively in complex and dynamic world. On the contrary, behavior-based approach is suitable for generating autonomous behaviors in the environment, but it still lacks of the capabilities to infer dynamic situations for high-level behaviors. This paper proposes a 2-layer control architecture to generate adaptive behaviors, which perceive and avoid dynamic moving obstacles as well as static obstacles. The first level is to generate reflexive and autonomous behaviors with the behavior network, and the second level is to infer dynamic situation of mobile robots with Bayesian network. Experimental results with various situations have shown that the robot reaches the goal points while avoiding static or moving obstacles with the proposed architecture.

1 Introduction

It is difficult to avoid moving obstacles because a mobile robot has to perceive situations by only utilizing its sensors in real-time. The predictability is central to the collision avoidance in non-stationary conditions. If there is no need to predict, we can rely entirely on what is sensed, resulting in a purely reactive approach [1]. However, it is difficult to predict in the real world, even in the best of circumstances. In MIT AI Lab., they proposed a method of learning each circumstance for avoiding moving obstacles [2]. Hashimoto utilized the evolutionary computation and fuzzy system for avoiding moving obstacles [3], and Inoue presented a behavior learning method based on Bayesian networks and experience of interaction between human and robots of avoiding moving obstacles [4]. Nicolescu and Mataric dealt with changing situations by learning and constructing hierarchical structure of previous behaviors for solving the problem of avoiding moving obstacles [5].

The behavior network proposed by Maes [6] can acquire global goals as well as autonomously select behaviors as bestowing goals on the behavior-based system. Although behavior network is proposed for solving problems with goals in uncertain environments, it is insufficient to generate adaptive behaviors in changing and complex situations. To cope with this problem we propose a control architecture in which

---

Bayesian network controls behavior network. This control architecture selects a behavior of the highest weight which is computed by the inference of Bayesian network in behavior network. Bayesian network has advantage that it is independently applied to environment because it is designed by only using the information from sensors.

2   Bayesian Inference Driven Behavior Network

The variables required for the mathematical notations of the proposed method for Bayesian inference driven behavior network are as follows. These variables are also applied to the function for the Bayesian inference [7], the computation of an activation level of each node, and the selection of an active node in the behavior modules.

- α : Activation level
- θ : The initial value of the global threshold which is reduced by 10% if no executable node has an activation greater than it
- ϕ : The weight of environmental sensor inputs and successor links
- γ : The weight of goal inputs and predecessor links
- δ : The weight of protected goal inputs and conflictor links
- t : Current time
- B : A set of behavior nodes
- B_{PS} : A set of preconditions of behavior node b (see table 1)
- B_{AS} : A set of add lists of behavior node b (see table 1)
- B_{DS} : A set of delete list of behavior node b (see table 1)

The function for selecting the activation node b_i in time t is defined as follows.

$$b_i(t) = \begin{cases} 1 & \text{if} \begin{cases} \alpha_{b_i}(t) \ge \theta & (1) \\ executable(b_i,t) = 1 & (2) \\ \alpha_{b_i}(t) \ge \alpha_{b_j}(t), \forall j \cdot \ni \cdot (1) \text{ and } (2) \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

$$executable\ (b_i,t) = \begin{cases} 1 & \text{if } b_i \text{ is executable at time } t \\ 0 & \text{otherwise} \end{cases}$$

The function for the activation level α_{b_i} in behavior node b_i is defined as follows.

$$\alpha_{b_i}(t) = \begin{cases} 0 & \text{if } t = 0 \\ D(b_i,t) + \beta(b_i,t) & \text{otherwise} \end{cases}$$

From this function, the activation value of each behavior $D(b_i, t)$ is defined as follows.

$$\begin{aligned} D(b_i,t) &= S(b_i,t) + G(b_i,t) - P(b_i,t) \\ &+ \sum_{b_j,b_k}(SP_{EW}(b_j,b_k,t) + SP_{FW}(b_j,b_k,t) - SP_{PG}(b_j,b_k,t)) \end{aligned} \quad (3)$$

$$SP_{EW}(b_j,b_k,t) = \begin{cases} 1 & \text{if } \sigma \in B_{PS}, \sigma \in B_{AS} \\ 0 & \text{otherwise} \end{cases}$$

$$SP_{FW}(b_j,b_k,t) = \begin{cases} 1 & \text{if } \sigma \in B_{AS}, \sigma \in B_{PS} \\ 0 & \text{otherwise} \end{cases}$$

$$SP_{PG}(b_j, b_k, t) = \begin{cases} 1 & \text{if } \sigma \in B_{DS}, \sigma \in B_{PS} \\ 0 & \text{otherwise} \end{cases}$$

In the above activation function (b_i, t) is the weight affecting the activation in behavior nodes after the agent infers situation by Bayesian network, and it is defined as follows.

$$\beta(b_i, t) = \begin{cases} \xi & \text{if } b_i \in r_{k,i}, i \in I, \Psi(k,i) = \sigma \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$r_{k,i} = \{s \mid s \in \text{effect\_node}(k), s : i\text{th state}\}$

$I = \{x \mid x \in S_I(t), 1 \le x \le \#(\text{Effect Nodes})\}$ $S_I(t)$: cause at time t

$\Psi(k,i) = P(b_{ki} \mid c_{1j} \cdots c_{mi})$: conditional probability, m: #(cause nodes)

$\sigma = Max\{\Psi(k,i) \mid 1 \le i \le \#(\text{states in } r_i), k : \text{effect node}\}$

In equation (3), $S(b_i, t)$ is the activation value from sensors at time t, $G(b_i, t)$ is the activation value from goals at time t, and $P(b_i, t)$ is the activation value to be deleted from protected goal at time t. In addition to these equations, $SP_{EW}(b_j, b_k, t)$ is the backward spreading from node b_j to b_k and it means both the preconditions of b_j and the add list of b_k and $SP_{FW}(b_j, b_k, t)$ is the forward spreading from node b_j to b_k, and it means both the add list of b_j, and the precondition of b_k. $SP_{PG}(b_j, b_k, t)$ is to take away from node b_j to b_k and it means both the delete list of b_j and the precondition of b_k.

Bayesian network represents all the events with DAG (Directed Acyclic Graph) and each node in DAG models probabilistic independency. It consists of the relation of cause and effect nodes from their probabilities and can infer the results from conditional probability of cause nodes. In Bayesian network, each node corresponds to probabilistic variables C_i, C_j, and E, which are the causes of the effect node and each link is associated with conditional probabilities between linked variables.

In equation (4), $r_{k,i}$ means the result node of Bayesian network at the ith state in the kth effect, and I is the set of indexes of result nodes where preconditions are satisfied at time t. $\Psi(k,i)$ is the conditional probability at the ith state in the kth result node, and σ is the state of the highest probability in the kth result node.

IF $b_i(t) = 1$ THEN execute b_i ELSE $\theta = \theta \times 0.9$ and repeat the procedure $\qquad$ (5)

Table 1. Internal links and external links

Internal Link	
Predecessor Link	(ρ =false) ^ ($\rho \in$ preconditions of node A) ^ ($\rho \in$ add lists of B)
Successor Link	(ρ =false) ^ ($\rho \in$ add lists of node A) ^ (A is executable) ^ ($\rho \in$ preconditions of B)
Conflictor Link	(ρ =true) ^ ($\rho \in$ preconditions of node A) ^ ($\rho \in$ delete lists of B)
External Link	
Sensors	(ρ =true) ^ ($\rho \in$ preconditions of node A)
Goals	($\rho \ge 0$) ^ ($\rho \in$ add lists of node A)

3   Experimental Results

For experiments, we use the YAKS which is the 3D robot simulator, and the experimental environment is shown as Figure 1 (a). The environment of this simulator has

the mobile robot (number 0) which generates behaviors using the proposed method and two robots (number 1 and 2) which act as the moving obstacles. These two robots like obstacles can only detect and avoid pens, and cannot avoid a robot when they collide with other robots. This simulator may change the angle of the view within this simulator. In Figure 1 (a), the white cylinders represent the static obstacles, and they have changeable count in the experiments. The robot randomly starts from any positions for the comparative analysis when using the only behavior network and the proposed method, and also starts from the static position for the various analyses of the proposed method; we compared the success rates to avoid moving obstacles using the only behavior network and the proposed method respectively, and we investigated an avoiding direction in each case when using the proposed method.

The goals in these experiments are to avoid the moving obstacles and to reach the goal position marked with the light area in the corner of pens. We put randomly 5% of noise in the value from sensors because we may not predict the situation accurately in real world.

3.1 Behavior Network and Bayesian Network

Figure 1 (b) shows the behavior network architecture used in this paper, which has primitive behaviors like 'Follow Light', 'Go Straight', 'Turn Left', and 'Turn Right.' In the network, the predecessor links of 'Follow Light' are 'Go Straight' and 'Near Light', and the successor link is 'Reach Goal'. The predecessor links of 'Go Straight' are also 'Nothing', 'Turn Left', and 'Turn Right', and the successor links are 'Reach Goal', 'Avoid Obstacle', and 'Follow Light'. The predecessor links of 'Turn Right' are 'Shade Area', 'Near Obstacle', and 'Go Straight', and the successor links are 'Avoid Obstacle' and 'Go straight'.
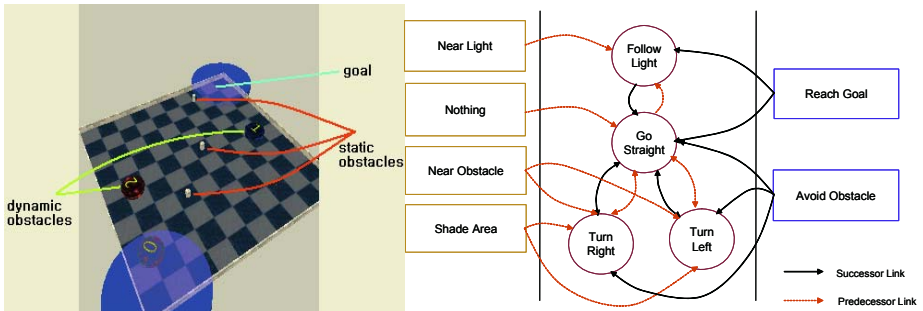
Fig. 1. (a) Experimental environment (left) (b) Behavior network (right)

Bayesian network designed for avoiding moving obstacles as well as static obstacles is as follows. The cause nodes from sensors are 'distance 0', 'distance 1', 'distance 2', 'distance 3', 'distance 4', 'distance 5', 'distance 6', and 'distance 7', and the cause nodes to infer the changing situations from previous behaviors are 'distance of obstacles', 'position of obstacles', 'previous behavior', and the directional change of obstacles like 'rear_object', 'front_object', 'left_object', and 'right_object'. The initial probabilities of each node are normalized as $1/n$ for the n conditions in each node. According to this, the initial probabilities of 'distance 0' are $P(Near0)=0.5$ and

P(*Far*0)=0.5. There are some important factors to the nodes named 'distance of obstacles', 'position of obstacles', 'previous behavior', 'left_object', 'right_object', 'front_object', and 'rear_object' to infer dynamic situations.

For example, if left sensors of robot detect the obstacle in current state, the probabilities of P(*Near*0) and P(*Near*1) in the nodes 'distance 0' and 'distance 1' get higher. The probabilities of P(*Approach*) and P(*GoAway*) in the node 'distance of obstacles' and P(*Front*2*Left*), P(*Left*), and P(*Rear*2*Left*) in the node 'left_object' are decided from the previous information doing by a robot. If P(*Approach*) > P(*GoAway*) in the node 'distance of obstacles', the highest probability in those of P(*NoTurn*), P(*TurnLeft*), and P(*Rear*2*Left*) in the node 'left_turn' is selected. The node 'left_object' consists of P(*Front*2*Left*), P(*Left*), and P(*Rear*2*Left*) which mean the change of obstacle from front to left, on left, and from rear to left, respectively. The conditional probabilities of the effect node 'left_turn' are determined with the above probabilities of cause nodes.

3.2   Comparison

After the experiments to avoid moving obstacles using only behavior network and the proposed method, we have observed that a robot may not avoid moving obstacles when a moving obstacle comes from the opposite direction to a robot. Since a mobile robot avoids collisions as the direction and position of an obstacle from the robot, 'Turn Left' or 'Turn Right' is an important behavior. When starting at the same position and direction in the only behavior network and the proposed method respectively, the robot collides with moving obstacles in the only behavior network, but avoids the obstacle and reaches the goal in the proposed method.

We have analyzed the behavior for avoiding moving obstacles which get nearer from various positions and directions using the behavior network and the proposed method. We have verified the success of 52% and 90% in those of 60 trials using the behavior network and the proposed method, respectively. The cases of failure in the proposed method are when the obstacle gets nearer on the front side or changes its direction abruptly. Figure 2 shows the success rates for avoiding moving obstacles in the only behavior network and the proposed method. In this figure, y-axis is the success rate that ranges from 0 to 1.
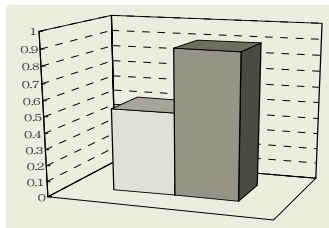
Fig. 2. The success rates for avoiding moving obstacles (Y-axis: Success rate, Dark bar: proposed)

Figure 3 shows the comparison of the behavior network with the proposed method. (a) is the comparison of the behavior sequences and (b) is that of the angles followed by the behaviors. In this figure, the solid line represents the behaviors and angles in

the behavior network and the dashed line represents the behaviors and angles in the proposed method. In the y-axis, 1 is 'Go Straight', 2 is 'Turn Left', and 3 is 'Turn Right'. In the figure, the robot selects 'Turn Left' or 'Turn Right' when using the proposed method, but 'Turn Right' when using the only behavior network.

3.3 Analysis of Results

As mentioned before, we have analyzed the behavior selection of the robot at time t from the moved angles between the robot and obstacles using the proposed method. Figure 4 shows that the robot reaches the goal while avoiding moving obstacles, pens, and static obstacles of 2, 3, and 2 times, respectively. In this figure, (a) shows the trajectories of the robot and moving obstacles, and (b) shows the behavior selection of the robot to reach the goal and has the marked circles representing the collision avoidances with moving obstacles, static obstacles, and pens. In (a), blue line represents the trajectory of the mobile robot and black and red line represent moving obstacles, respectively.

We have subsequently analyzed behaviors of the robot in cases of the same direction and the different direction for comparing behavior selection as the changes of the angles between the robot and obstacles. Figure 5 shows the result of avoiding moving obstacles. In this figure, (a) and (b) are the trajectories in the cases of colliding with an obstacle 2 in the same direction on the left and on the right of the robot, and (c) represents the behavior sequences in (a) with respect to time. Lastly, (d) represents the behavior sequences in (a) with respect to time. In (c) and (d), x-axis is time line and y-axis is the selected behavior with respect to time. The numbers of y-axis represent that 1 is 'Go Straight', 2 is 'Turn Left', and 3 is 'Turn Right'. Figure 6 shows the results to avoid moving obstacle with a real Khepera II mobile robot. In this figure, each of (a, c) and (b, d) shows the situations before and after avoiding a white box in front of the mobile robot.

(a)   (b)
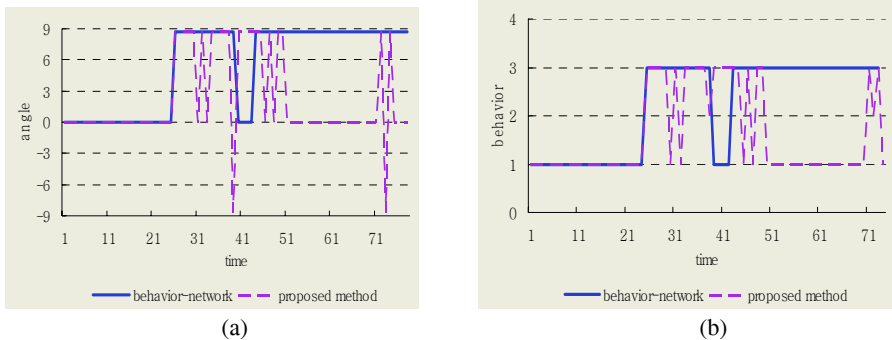
Fig. 3. The behaviors (a) and angles (b) in the behavior network and the proposed method

4 Conclusions

The experiments have validated that mobile robot goes to the goal with avoidance behaviors on moving obstacles using the Bayesian inference driven behavior net-

work. Autonomous behaviors of a mobile robot are generated by the inter-relations of sensors, goals and behaviors in behavior network, and adaptive behaviors of a mobile robot are also generated by expanding the behavior network through the inference of each situation using Bayesian network. The proposed method does not have limitation of re-constructing the system as the changes of environment like the systems with hybrid learning or planning, and overcomes the limitation to re-define many rules such as the systems using fuzzy rules and fuzzy inference [2].

(a)   (b)
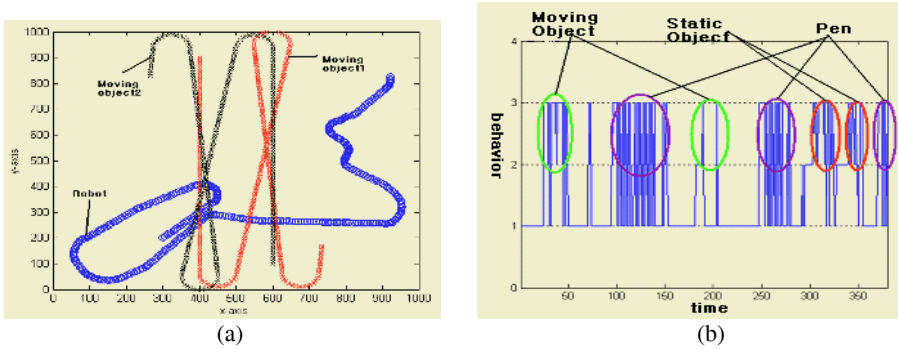
Fig. 4. The trajectory (a) and each behavior (b) by the time of avoiding moving obstacles in the proposed method

(a)   (b)   (c)   (d)
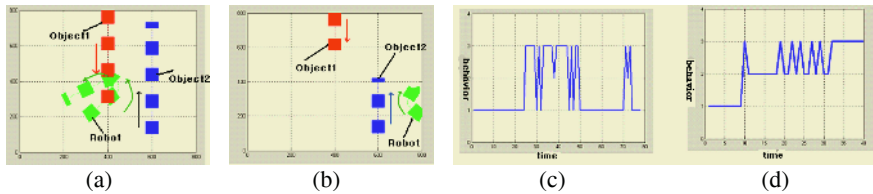
Fig. 5. Avoiding moving obstacles as the same direction with the robot using the proposed method. Each figure shows the avoidance in the right side of robot (a, c) and in the left side of robot (b, d)

(a)   (b)   (c)   (d)
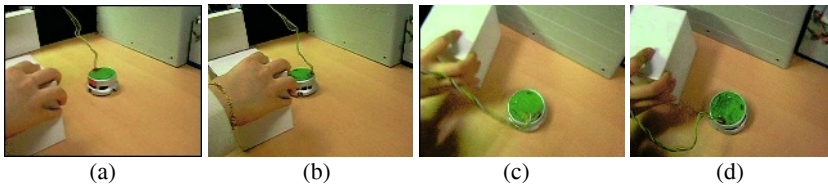
Fig. 6. Avoiding moving obstacle with the real robot. Each shows the avoidance in the middle (a,b) and at the left (c,d) of the robot

References

1. R. C. Arkin, *Behavior-Based Robotics,* MIT Press, 1998.
2. W. D. Smart, *Making Reinforcement Learning Work on Real Robots*, Ph.D.Thesis, 2002.

3. S. Hashimoto, F. Kojima, and N. Kubota, "Perceptual system for a mobile robot under a dynamic environment," *Proc. 2003 IEEE Int. Symposium on Computational Intelligence in Robotics and Automation*, pp. 747-752, 2003.
4. T. Inamura, M. Inaba, and H. Inoue, "User adaptation of human-robot interaction model based on Bayesian network and introspection of interaction experience," *Proc. of the 2000 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 2139-2144, 2000.
5. M. N. Nicolescu and M. J. Mataric, "A hierarchical architecture for behavior-based robots," *Autonomous Agents and Multi-Agent Systems*, pp. 227-233, 2002.
6. P. Maes, "How to do the right thing," *Connection Science J*ournal, vol. 1, no. 3, pp. 291-323, 1989.
7. J. Pearl, Probabilistic *Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.

Loss Minimization Control
of Induction Motor Using GA-PSO

Dong Hwa Kim and Jin Ill Park

Dept. of Instrumentation and Control Eng., Hanbat National University
16-1 San Duckmyong-Dong Yuseong-Gu, Daejon City, Korea, 305-719
Tel: +82-42-821-1170, Fax: +82-821-1164
kimdh@hanbat.ac.kr
www.aialab.net

Abstract. This paper deals with the GA-PSO (Genetic Algorithm-Particle Swarm Optimization) based vector control for loss minimization operation of induction motor. It is estimated that more than around 50% of the world electric energy generated is consumed by electric machines such as induction motor, DC motor. So, improving efficiency in electric drives is important and control strategy for minimum energy loss is needed as one of optimal operation strategy. In this paper, vector control approach is suggested for an optimal operation of induction motor using GA-PSO tuning method through simulation.

1 Introduction

Induction motors have been widely used in various industries as actuators or drivers to produce mechanical motions and forces. Generally, variable-speed drives for induction motors require both wide operating range of speed and fast torque response in the required conditions, regardless of load variations. Since it is estimated that more than around 50% of the world electric energy generated is consumed by electric machines improving efficiency in electric drives is important [1], [2]. That is mainly, for two reasons as follows: economic saving and reduction of environmental pollution [3]. Generally, since induction motors have a high efficiency at rated speed and torque, it has been widely used as driver in the industrial field. However, at light loads, iron losses increase dramatically, reducing considerably the efficiency [2], [3]. To improve the motor efficiency, the flux must be reduced, obtaining a balance between copper and iron losses [3]. Generally, there exist two different approaches to improve the efficiency in induction motor referred to as power measure based method and loss model based method. In this paper, a new minimum-time minimum-loss control algorithm for induction motors using hybrid system GA-PSO is suggested to obtain high performance, as well as high efficiency, under practical constraints on voltage and current. The validity of the suggested scheme, which carries out minimum-time speed control in the transient state and minimum-loss control in the steady state, will be revealed via simulation, including an induction motor model.

2 Mathematical Model for Loss Minimization

The direct torque and flux control for induction machine drives has been developed as direct torque control (DTC) in [1], [3] and as direct self control (DSC) in [2]. The technique was generalized to all ac drives, as torque vector control (TVC) in [3], and

it was recognized as a viable alternative to field-oriented control (FOC) [4]. Industrial drives with DTC are present on the market today [5]. DTC abandons the stator current control philosophy, characteristic of FOC and achieves bang-bang torque and flux control by directly modifying the stator voltage in accordance with the torque and flux errors. DTC is characterized by fast dynamic response, structural simplicity, and strong robustness in the face of parameter uncertainties and perturbations. It does not employ current controllers and pulsewidth modulation (PWM), and it is well suited for sensorless drives. Classic DTC has still several drawbacks. If significant energy saving are to be obtained, is necessary to optimize the efficiency of motor drive systems. Various methods have been investigated to achieve such a goal [1-3]. In [2], an optimal controller scheme has been proposed and implemented on a conventional field oriented control (FOC) structure of an induction motor. Simulation results have shown that it is possible to significantly improve the induction motor drive efficiency when it is slightly loaded. DTC has evolved from a particular control strategy to a wide concept that employs a broad range of control solutions. Many existing schemes depart from the classic bang-bang approach and achieve high-performance torque and flux control by using modern control techniques. All approaches that realize the motor control by direct modification of stator voltage, based on torque and flux errors, can be considered DTC schemes. Using the stator currents i_{ds}, i_{qs}, and the rotor fluxes ϕ_{dr}, ϕ_{qr} as state variables, and by selecting a reference frame in such a way that the rotor flux, $\phi_{qr} = 0$, the equations of the induction motor can be defined as [2]:

$$v_{ds} = R_s i_{ds} + \sigma L_s \frac{di_{ds}}{dt} + \frac{M}{L_r} \frac{d\phi_{dr}}{dt} - \omega\sigma L_s i_{qs} ,$$

$$v_{qs} = R_s i_{qs} + \sigma L_s \frac{di_{qs}}{dt} + \omega \frac{M}{L_r}\phi_{dr} + \omega\sigma L_s i_{ds} , \quad Mi_{ds} = \phi_{dr} + \frac{L_r}{R}\frac{d\phi dr}{dt} ,$$

$$\omega = \omega_m + \frac{MR_r}{L_r}\frac{i_{qs}}{\phi_{dr}} , \quad T_e = \frac{M}{L_r}\phi_{dr}i_{qs} = J\omega_m + KB\omega_m + T_r , \quad \sigma = 1 - \frac{M^2}{L_r L_s}$$

(1)

In these equations, (R_s, L_s) and (R_r, L_r) are the stator and rotor equivalent windings resistance and inductance; M is the mutual inductance between rotor and stator dq windings; σ is the leakage coefficient of the motor; ω_m and its derivative are respectively the rotor angular speed and acceleration with respect to the stator; ω is the stator reference frame speed. The total loss in an induction motor can be defined as the difference between the input and output powers [1]:

$$\eta = \frac{P_{out}}{P_{out}} , \quad P_{in} = v_{ds}i_{ds} + v_{qs}i_{qs} , \quad P_{out} = \omega_r T_e ,$$

(2)

where, ω_r is the rotor speed, T_e, is the electromagnetic torque, $v_{ds}, v_{qs}, i_{ds}, i_{qs}$ are the stator voltages and the stator currents in the dq reference frame, respectively.

3 Loss Minimization Vector Control Using GA-PSO

This paper focuses on the advantage of PSO into the mutation process of GA, for improving the GA learning efficiency. Euclidean distance is used on crossover to

avoid local optimal and obtain fast running time of solution. We illustrate the performance of the method using four test functions [4, 5].Using the conventional GA or PSO approach optimal solutions are obtained mostly with some initial differentiated data and there is a high possibility for obtaining local optimal solutions. The proposed approach uses data points with the longest Euclidean distance for crossover process to avoid such local optimization. The idea is to obtain global solutions by considering the entire search space (all the data points). We used the position and speed vector of PSO as follows:

$$v_{f,g}^{(t+1)} = w \cdot v_j^{(t)} + c_1^* rand()^* \left(pbest_{j,g} - k_{j,g}^{(t)} \right) + c_2^* Rand()^* \left(gbest_g - k_{j,g}^{(t)} \right),$$

$$j=1,2,...,n,\ g=1,2,...,m,\ k_{j,g}^{(t+1)} = k_{j,g}^{(t)} + v_{j,g}^{(t+1)},\ k_g^{min} \le k_{j,g}^{(t+1)} \le k_g^{max}. \tag{3}$$

n : The number of agent in each group, m : The number of member in each group

t : Number of reproduction steps, $v_{j,g}^{(t)}$: The speed vector of agent j in reproduc-

tion step of tth . $V_g^{min} \le v_{j,g}^{(t)} \le V_g^{max}$, $k_{j,g}^{(t)}$: The position vector of agent j in repro-

duction step of tth , w : Weighting factor, $c1, c2$: Acceleration constant,

$rand(), Rand()$: Random value between 0 and 1, $pbest_j$: Optimal position vector of

agent j , $gbest$: Optimal position vector of group.
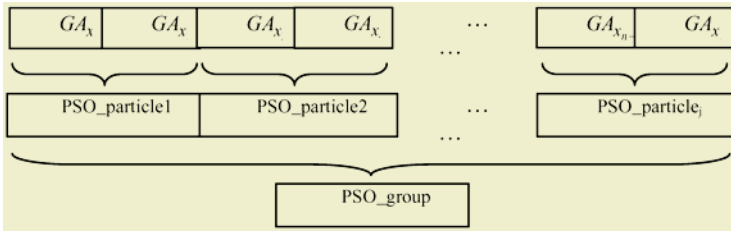
Fig. 1. Individual structure combined by PSO and GA
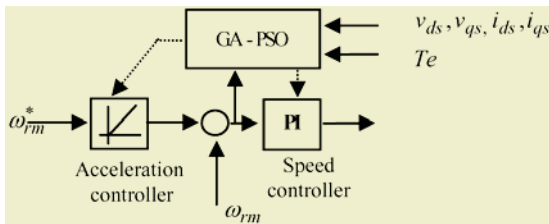
Fig. 2. Control block diagram using GA-PSO

4 Simulation and Discussion

This paper used squirrel induction motor 20HP, 220V, 60Hz with parameters as shown in Table 1 to simulate [5]. For safety, reference speed is defined as half speed of rated speed. Weighting function of fitness function is arranged as ration of torque and speed (1:50). Table 1 is parameters of GA-PSO selected for motor simulation.

Table 1. GA-PSO Parameters

Number of particles in a group	4*3
Maximum generation	30
Number of members in a particle	3 (P,I, γ)
Acceleration constant	2
Mutation rate	0.5
Maximum inertia weight factor	0.9
Minimum inertia weight factor	0.4
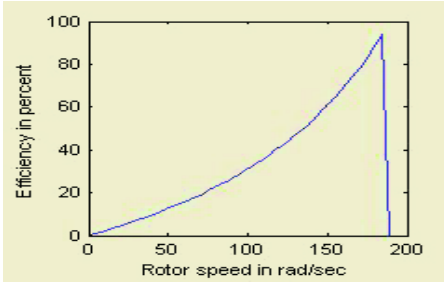
Fig. 3. Efficiency versus speed of the simulation motor
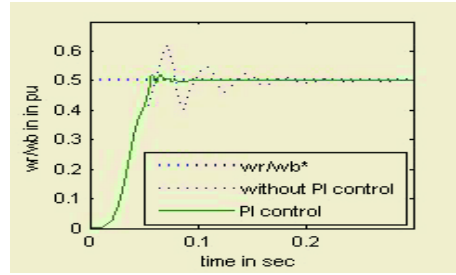
Fig. 4. Speed characteristics of GA-PSO based tuning PI Controller: GEN=30

4.1 V/F Control Without PI Current Controller

In this section, to represent the characteristic of how Volt/Frequency device affect motor control system before illustrate the characteristic of PI speed controller shown in Fig. 4, we control motor using only Volt/Frequency control referred to in reference [3]. Fig. 3 gives these characteristics. Speed following performance is not good as expecting. Speed following function of Fig. 5 is stable after 0.2sec and torque characteristic of Fig. 6 is showing stable torque after 0.2sec. Fig. 6 of torque following characteristic without PI represents oscillation.

4.2 Tuning of PI Parameter Using GA-PSO

This section tunes P, I parameters simultaneously together, using GA-PSO. Figs. 5-7 and Table 2 show the results of this simulation. Fig. 5 illustrates the best characteristic acceleration curve is able to be obtained to the given load, by means of a simultaneously tuning of acceleration constant γ and PI parameter using GA-PSO. Also, Fig. 5 and Figs. 6-7 of torque and efficiency characteristic curves using the tuned acceleration constant γ and PI parameter based on GA-PSO.

5 Conclusions

In this paper, the loss minimization optimum vector control with speed following function, moderate torque to load change, and good efficiency is suggested. In this method, a simultaneous tuning approach of acceleration constant and PI parameters is used on GA-PSO. Optimal solution could be obtained by only small number of indi-

viduals in GA-PSO, using multiobective function of GA-PSO. This strategy for efficiency can be well performed for the load change. Overall, the best strategy is performed. It is needed for real-world application hat the requirements of more application that will eventually determine which strategy is more advantageous.
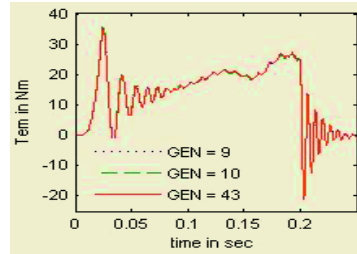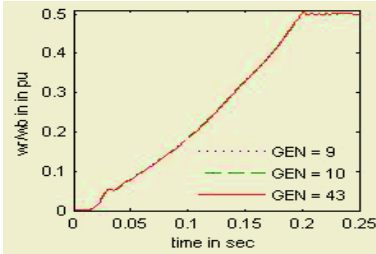
Fig. 5. The acceleration curve using PI parameter based on GA-PSO

Fig. 6. Torque characteristic using PI parameter based on GA-PSO
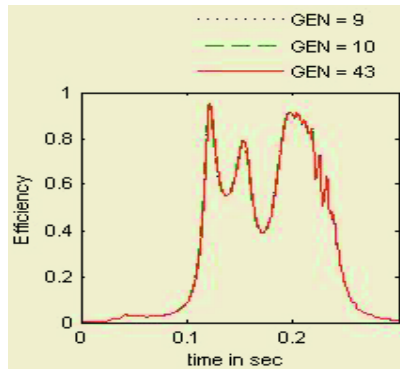
Fig. 7. Efficiency characteristic using the tuned acceleration constant γ and PI parameter based on GA-PSO

Table 2. Parameters in generations

Generation	P	I	D	Fitness_value
5	9.8239	4.2431	2.7900	1.2050
10	9.9557	2.5413	2.7900	1.2074
43	9.9428	4.6268	2.9302	1.2078

Acknowledgement

References

1. I. Takahashi and T. Noguchi, "A new quick response and high efficiency control strategy of an induction motor," in *Conf. Rec. IEEE-IAS Annu. Meeting*, 1985, pp. 495–502.

2. I. Boldea and S. A. Nasar, "Torque vector control (TVC)—A class of fast and robust torque, speed and position digital controllers for electric drives," in *Conf. Rec. EMPS'88*, vol. 15, 1988, pp. 135–148.
3. C. Lascu and A. M. Trzynadlowski, "Combining the principles of sliding mode, direct torque control and space vector modulation in a high performance sensorless AC drive," in *Conf. Rec. IEEE-IAS Annu. Meeting*, vol. 3, 2002, pp. 2073–2078.
4. Dong Hwa Kim, "Robust PID controller tuning using multiobjective optimization based on clonal selection of immune algorithm," Proc. Int. Conf. Knowledge-based intelligent information and engineering systems. Springer-Verlag. pp. 50-56, 2004.
5. Zwe-Lee Gaing, "A Particle Swarm Optimization Approach for Optimum Design of PID Controller in AVR System", IEEE Trans. Energy Con. Vol. 19, pp. 384-391, No. 2, June, 2004.

Context-Restricted, Role-Oriented Emotion Knowledge Acquisition and Representation

Xi Yong[1,2], Cungen Cao[1], and Haitao Wang[1,2]

[1] Key Lab of Intelligent Information Processing
Institute of Computing Technology, Chinese Academy of Sciences
{yongxi,cgcao,htwang}@ict.ac.cn
[2] Graduate School of the Chinese Academy of Sciences, Beijing 100080, China

Abstract. Emotion knowledge is a fundamental part of human commonsense knowledge. For the lack of solid knowledge extraction work in this domain, this paper focuses on acquiring and representing the antecedent situations of various emotions as well as the emotion subjects' interpretation to these situations in real-world emotion-eliciting scenarios. After a comprehensive analysis on the meaning structure of three critical notions in emotion understanding and modeling (i.e. role, context and event), this paper introduces a pragmatic emotion knowledge acquisition method and presents a context-restricted, role-oriented, frame-based representational model of emotion knowledge, with a case study which demonstrates the applicability of the model.

1 Introduction

There is a body of commonsense knowledge about human psychology that we all draw upon in everyday life to interpret our own emotions and those of the people around us[7]. We define this special kind of knowledge as emotion knowledge.

Nowadays, researchers increasingly realize the lack of commonsense knowledge about emotion has highly limited the ability of intelligent systems to interact with users more naturally and socially[4]. The most instructive work which has formally been done in this area includes the Affective Reasoner designed by Elliot[1], the logic for emotions proposed by Sanders[7] and the reasoning formalism for emotional agents presented by Meyer[5]. They laid a good foundation for our further exploration on the construction of an operable emotion knowledge base. Since there are presently no large-scale knowledge bases of this kind, this study will be very beneficial in this direction and can help us a lot to establish more sophisticated emotional systems.

To prepare for this work, we extracted large amounts of text segments describing some emotion-eliciting scenarios from textual Web pages on WWW, in proper length and with proper keywords (e.g. angry, sad or happy), as our initial corpus. On this basis, we first explored one basic emotion (i.e. anger) in detail, summarized all possible influencing factors of anger and proposed a parametric framework to define different categories of anger situations[11]. Then, as we noticed the deep influence of one's role and psychological content on the generation of emotions, we introduced the two notions in our model and presented a role-oriented representational model of the emotion knowledge accounting for some role-player's emotional experiences[12].

In this paper, we attempt to integrate and extend previous work into a systematic architecture of the whole emotion-generating mechanism, and present a pragmatic

emotion knowledge acquisition method and a context-restricted, role-oriented, formal representational model as a result. This model incorporates all the key ideas embodying in our previous work and integrates frameworks of different structure into a comprehensive architecture, which can give explicit explanation to every elicited emotion.

This paper is organized as follows. Section 2 analyzes the meaning structure of role, context and emotion-eliciting scenario and gives the emotion knowledge acquisition method. Section 3 presents the context-restricted, role-oriented, formal representational model of emotion knowledge, and in Section 4, we formalize a concrete case to demonstrate the applicability of our model. Finally, section 5 concludes the paper.

2 Emotion-Eliciting Mechanism

In present psychological literature, the predominant theories of emotion are appraisal theories, which suggest that emotions are generated by appraisal of a subjectively important event by the person experiencing that emotion[6]. The appraisal process has its special nature of subjectivity, depending on the subject's particular psychological content which is tightly connected with his/her role in the current social context. A basic idea of our research is to analyze and understand emotion-eliciting scenarios in terms of the emotional agents' roles in the restrictive contexts, and then going on to equip those roles with sets of beliefs, desires, and other emotion knowledge which contributes to form the interpretation of external events in the appraisal process.

2.1 Role

We've discussed the notion of natural type and role in detail in [12]. While a natural type is related to some intrinsic characteristic of a person such as sex, age, personality trait, and so on, a role indicates a certain position where the individual stands in relation to some other entity in a social context, and a person can enter and leave that context and accordingly acquire and abandon the role without losing his/her identity.

Definition 1 (Role). Given an entity e, R is a role of e if and only if it satisfies:
 1) There exists another entity e' relating to e by R, i.e. RR'(e, e'), where R' is the role of e';
 2) e can still exist and be itself before getting R and after losing R.

As stated in [9], given the set of role types R, there is a role hierarchy (R; $\leq_{RR}$) defined by a partial order $\leq_{RR}$ on R. For every r' $\leq_{RR}$ r, r is a more general role compared to r', and r' is called a subrole of r and r a superrole of r'. Furthermore, a role always comes with its own natural type. We can assign different natural types to the same role, and provide them with their special psychological content besides the common knowledge of this role. Thus, different views to the world held by the same role of different natural types won't cause any inconsistency in the knowledge base.

According to Elliot's twenty-four emotion-type version of the emotion-eliciting condition theory[1], to consider four constituent parts of psychological content will be adequate to explain all the twenty types of rudimentary emotions concerned in our current study. We don't take the four emotion types in Elliot's theory rooted in subject's preferences with respect to objects into account in this study because people usually need no reason for liking or disliking something.

Definition 2 (Psychological content). Psychological content, denoted by PC, is a tuple PC=(Beliefs, Desires, Norms, Emotion-patterns), where

1) Beliefs={bel_1, bel_2, …, bel_j}, which refers to the set of things that the person playing the role believes to be true. It contains various facts and rules the appraisal mechanism relies on to map an external situation into one's internal concerns which may be congruence or incongruence with some desire and/or norm.
2) Desires={des_1, des_2, …, des_k}, which refers to the set of all kinds of desires and needs of the person bearing the role. They are used to judge whether the role-player wants the event or situation to occur or not.
3) Norms={$norm_1$, $norm_2$, …, $norm_l$}, which refers to the set of all the social norms the role-player should abide by in the social context. They are used to determine one's appraisal of right and wrong to the current happening.
4) Emotion-patterns={ep_1, ep_2, …, ep_m}, which refers to the set of the emotion-eliciting conditions of the person playing the role. They are taken as appraisal rules in the appraisal process, which will be matched with the role-player's construal of the external event to discern which emotion will be generated.

In our representation, we introduce four modal operators (i.e. B, D, N and EP) to represent one's psychological content. Assuming x is the variable representing the agent playing the role, and φ is the predicate formula representing the content of the corresponding PC item, then bel_i, des_i, $norm_i$, and ep_i can be respectively represented as "$B(x, \varphi)$", "$D(x, \varphi)$", "$N(x, \varphi)$", and "$EP(x, \varphi)$".

2.2 Context

As most roles are connected with certain contexts and every external event happens in some social context, we employ the context mechanism to organize the emotion knowledge base. A context binds the concerned agents and other restrictive conditions together to form a sharable meaning space of the psychological content of each agent and the interpretation of the events taking place in the context. It indicates the relative network an agent is discussed in as well as the background situation of an event.

Definition 3 (Context). A context, denoted by C, is a tuple C=(Agents, Objects, Roles, Relations, Assumptions), where

1) Agents={a_1, a_2, …, a_j}, which refers to the set of all the participants involved in the context. Each agent is assigned with one or more roles, depending on the relations between this agent and other agent or object.
2) Objects={o_1, o_2, …, o_k}, which refers to the set of all the non-personal things involved in the context.
3) Roles={r_1, r_2, …, r_l}, which is the set of all the roles occurring in the context.
4) Relations = {rel_1, rel_2, …, rel_m}, which is the set of the binary relations specifying all the associations among agents (inter-agent relations) and between agents and objects (agent-object relations) in the context.
5) Assumptions=$assum_1 \wedge assum_2 \wedge … \wedge assum_n$, which refers to the conjunction of other presumed antecedents for entering the context.

Given the set of contexts C, we can define a context hierarchy (C; $\leq_{CC}$) through a partial order $\leq_{CC}$ on C. For every c' $\leq_{CC}$ c, c' is a more special context compared to

c, and c' is called a subcontext of c and c a supercontext of c'. On the one hand, in the role hierarchy, lower roles inherit the psychological content of their superroles in the same context; on the other hand, in the context hierarchy, roles in lower contexts inherit the psychological content of the same role in their supercontexts. Thus, we needn't specify the content in those superroles and supercontexts again in them.

2.3   Emotion Knowledge Acquisition Method

An emotion-eliciting scenario consists of a pivotal stimulating event, the happening context of the event, and an agent's elicited emotion(s) with its behavioral reaction(s) under the emotion(s). Each piece of acquired knowledge corpus describes an emotion-eliciting scenario. During the knowledge acquisition process, we carefully extract the useful information about those components of emotion-eliciting scenario from each scenario description, and manually formalize all the newly occurring knowledge into appropriate category or instance frames.

   On the whole, the practical acquisition process can be arranged into following six steps:

1) Locate the emotion-eliciting event in the scenario description, with which spot the involved participants and objects. Then identify the emotion-producing agent among those participants as well as the resulted emotion and reactive behavior;
2) If there is other participant or object, chase down the relationship between every pair of agents or agent and object and accordingly assign the corresponding role(s) to each agent;
3) Analyze necessary implied assumptions in the scenario description which have some influence on the agent's interpretation to the event. Then with other recognized conditions, determine the context structure of the event, and if it hasn't existed in the context base yet, create a new context for situations of this type;
4) Instantiate the emotion-eliciting event as a frame instance of this event category;
5) Analyze how the emotion is inferred from the external event and the reasoning process involves what psychological content. Then choose a proper role of the emotion-producing agent to contain this content and add it into its corresponding part if it hasn't existed there.
6) Check on each newly created or modified frame to ensure the knowledge is well and truly special to that context and role; otherwise generalize the context or role into an appropriate supercontext or superrole to contain this content.

3   Emotion Knowledge Representation

Based on previous analysis, this section presents a context-restricted, role-oriented, formal representational model of emotion knowledge. Firstly, we give a specification of the unary and binary predicates used in our model as shown in Table 1.

   Moreover, we can attach one or more supplementary argument items to any predicates to provide more information about the states or events. These additional items include *Time, Location, Aspect*, *Content*, *Manner*, *Instrument*, *Cause*, *Result, Degree*, *Frequency* and *Possibility*. They can effectively enhance the expressive capability and flexibility of our representation.

Table 1. Predicate Specification

Predicate Type		General Form	Predicate Implication	Argument Specification
Nominal Predicate (NP)	Attribute Predicate(AP)	P(x, y)	describe the state of an entity in terms of certain attribute	P: the attribute name x: the entity y: the attribute value
	Type Predicate(TP)	P(x)	define the type of an entity	P: the type name x: the entity
	Relation Predicate(RP)	P(x, y)	define a relation between two entities	P: the relation name x, y: the entity
Verbal Predicate (VP)	Single Agent Predicate(SAP)	P(x)	define an event voluntarily executed by an agent	P: the event name x: the agent
	Single Patient Predicate(SPP)	P(x)	define an event a patient involuntarily undergoes	P: the event name x: the patient
	Dual Agent Predicate(DAP)	P(x, y)	define an event executed by two agents together	P: the event name x, y: the agent
	Single Agent Single Recipient Predicate(SSP)	P(x, y)	define an event actively executed by an agent to an passive recipient	P: the event name x: the agent y: the recipient
	Dual Patient Predicate(DPP)	P(x, y)	define an event two patients undergo together	P: the event name x, y: the patient

The representational framework consists of role representation, context representation and event representation, as shown in Fig.1 through to Fig.4.

The role representation consists of two layers: the first layer defines the role according to our understanding (see Fig.1); and the second layer specifies its concrete psychological content (see Fig.2, where the *component* will be replaced by *belief*, *desire*, *norm* or *emotion-pattern* in actual instances). The <context-assignment> in Fig.2 and Fig.4 is represented as "Context: <context-name>". It plays two roles: first, it introduces parameters defined in the context to be used in the frame body specification; second, it is by default part of antecedents of assertions specified in the frame.

4   Case Study

In this section, we give an example showing a family context, a family role with its partially specified psychological content and two successive events taking place in the context as a case study (see Fig.5 through to Fig.12) to demonstrate the applicability of our model. The formalized knowledge in these frames is extracted from an interesting scenario "*The child incautiously broke up the vase at home. He was afraid of being blamed and hided its fragments*" according to our acquisition method.

In Fig.10, "BP" represents the predicate type *Behavior Predicate*, which consists of SAP, SSP and DAP. In actual matching process, P1 and P2 can be matched with any predicates of this type, and the symbol "*" can be matched with none or one necessary argument of the predicates, depending on the specific types of the predicates.

According to the *Child*'s norms in Fig.9, the child thinks it wrong to incautiously break up the vase. According to the *Child*'s beliefs in Fig.7, the child believes if the parent knows this thing, he/she will blame him/her. Finally, according to the *Child*'s desires in Fig.8, the child doesn't want to be blamed by the parent. So, according to the Child's emotion patterns in Fig.10, when the child incautiously breaks up the vase in this context, he/she will fear for being blamed. This inference is according with our common sense, which implies that our emotion pattern is quite reasonable.

```
defrole <role-name>
'{'
        slot: definition
            :type WellFormedFormula
        slot: subrole-of
            :type SeqOf String
        slot: belief
            :comment "taking an instance of
            <belief-name> as its value."
        slot: desire
            :comment "taking an instance of
            <desire-name> as its value."
        slot: norm
            :comment "taking an instance of
            <norm-name> as its value."
        slot: emotion-pattern
            :comment "taking an instance of
            <EP-name> as its value."
'}'
```

Fig. 1. Role Representation (First layer)

```
defframe <component-name>
'['<context-assignment>']'
'{'
        slot: have-subset
            :type SeqOf String
        slot: component
            :type WellFormedFormula
            :comment "indicating the con-
            crete knowledge contained in the
            <component-name>"
'}'
```

Fig. 2. Role Component Representation (Second layer)

```
defcontext <context-name>
'{'
        slot: subcontext-of
            :type SeqOf String
        slot: agent
            :type Char String
            :facet type
            :facet role
        slot: object
            :type Char String
            :facet type
        slot: relation
            :type ConjOf RP
        slot: assumption
            :type ConjOf Literal
'}'
```

Fig. 3. Context Representation

```
defevent <event-name>
'['<context-assignment>']'
'{'
        slot: event
            :type VP
            :comment "taking the
            <event-name> as the
            <predicate-name>"
        slot: emotion
            :type ConjOf AP
            :comment "indicating the
            agent's elicited emotion"
        slot: reaction
            :type DisjOf VP
            :comment "indicating the
            agent's one or more possi-
            ble reactions under the
            emotion"
'}'
```

Fig. 4. Event Representation

```
defcontext ParentChildContext22
{
        subcontext-of: ParentChildContext
        agent: p
            :type StrictAdult
            :role Parent, and Owner
        agent: c
            :type Juvenile
            :role Child
        object: v
            :type Valuables
        relation: ParentChild(p, c) ∧ Own(p, v)
}
```

Fig. 5. *ParentChildContext22* Frame

```
defrole Child
{
        definition:  Forall(x:  Person)
        (Child(x) ↔ Exists(y: Person)
        ParentChild(y, x))
        subrole-of: JuniorGeneration
        belief: ChildBelief
        desire: ChildDesire
        norm: ChildNorm
        emotion-pattern:     ChildEmo-
        tionPattern
}
```

Fig. 6. *Child* Frame

```
defframe ChildBelief
[Context= ParentChildContext22]
{
      B(c, Know(p, BreakUp_Man-
      ner(c,v,incautiously)      Blame
      _Possibility(p, c, very-likely))
      B(c, Hide_Cause(c,v,BreakUp
      (c,v))      ¬Know_Possibility (p,
      BreakUp(c, v), likely))
}
```

Fig. 7. *ChildBelief* Frame

```
defframe ChildDesire
[Context= ParentChildContext22]
{
      BreakUp_Manner(c,  v,  incau-
      tiously)     D(c, ¬Blame(p, c))
      Hide_Cause(c, v, BreakUp(c, v))
         D(c, ¬Know(p, BreakUp(c,v)))
}
```

Fig. 8. *ChildDesire* Frame

```
defframe ChildNorm
[Context  =  ParentChildCon-
text22]
{
      N(c,  ¬BreakUp_Manner
      (c, v, incautiously))
}
```

Fig. 9. *ChildNorm* Frame

```
defframe ChildEmotionPattern
[Context= ParentChildContext22]
{
      EP(c, Forall(P1,P2: SSP)(N(c, ¬P1(c, v)) ∧
      D(c, ¬P2(p, c)) ∧ B(c, Know(p, P1(c, v))
      P2_Possibility(p, c, very-likely)) ∧ P1(c, v)
         Emotion_Content(c, fear, P2(p, c))))
      EP(c, Forall(P1,P2: BP)(D(c, P1(p*)) ∧ B(c,
      P2(c*)     P1_Possibility(p*, likely)) ∧ P2(c*)
         Emotion_Content(c, hope, P1(p*))))
}
```

Fig. 10. *ChildEmotionPatterm* Frame

```
defevent BreakUp3
[Context= ParentChildContext22]
{
      event: BreakUp_Manner(c, v,
      incautiously)
      emotion: Emotion_Content(c,
      fear, Blame(p, c))
      reaction: Hide_Cause(c, v,
      BreakUp(c, v))
}
```

Fig. 11. An Event Frame

```
defevent Hide6
[Context= ParentChildContext22]
{
      event:      Hide_Cause(c,      v,
      BreakUp(c, v))
      emotion:     Emotion_Content(c,
      hope, ¬Know(p, BreakUp(c, v)))
}
```

Fig. 12. The Successive Event Frame

To end this section, we can conclude that this representational architecture is compatible with our knowledge acquisition method, and is suitable to extract and describe the emotion knowledge behind most relatively unsophisticated social situations. We can see this model is effective and rational, and we can use this knowledge to predict or explain the emotion of a person with a certain role in our everyday life.

5   Conclusion

Human emotion knowledge is a special kind of commonsense knowledge, which can contributes a lot to improve the believability of intelligent systems. This paper integrated the results of our previous research on emotion knowledge and extended them into a comprehensive representational architecture underlying the mechanism which accounted for the generation of human emotions in real world. We introduced the notion of context and employed the context mechanism to organize the emotion knowledge. Context provides a sharable meaning space to both psychological content

statement and interpretation of external events. This study led to a pragmatic emotion knowledge acquisition method and a context-restricted, role-oriented representational model of emotion knowledge described by a frame-based formal language. This work would facilitate the juxtaposition of common sense from across various social contexts, ultimately improving our understanding of emotions and the underlying mechanisms. Finally, we must point out that this model isn't meant to be the ultimate outcome of our research, but rather a promising first step in this direction. There are quite a few more complicated issues deserving serious consideration in our future work, such as the treatment of compound emotions, the dynamical update of one's psychological content, the efficient organization of the emotion knowledge base, and so on.

References

1. Elliott, C.D.: The Affective Reasoner: A Process Model of Emotions in a Multi-agent System. Ph. D. Thesis, the Institute for the Learning Sciences, Northwestern University (1992)
2. Gratch, J., Marsella, S.: A Domain-independent Framework for Modeling Emotions. In: Journal of Cognitive Systems Research, vol. 5(4) (2004) 269–306
3. Lenat, D: The Dimensions of Context Space. Available on the Web-site of the Cycorp Corporation, http://www.cyc.com/publications (1998)
4. Liu, L.G., Lieberman, H.: A Model of Textual Affect Sensing Using Real-world Knowledge. In: Proc. 8th International Conference on Intelligent User Interfaces. Florida (2003) 125–132
5. Meyer, J.-J.Ch.: Reasoning about Emotional Agents. In: Proc. 16th European Conference on Artificial Intelligence, IOS Press (2004) 129-133
6. Ortony, A., et al.: The Cognitive Structure of Emotions. Cambridge University Press (1988)
7. Sanders, K.E.: A Logic for Emotions: a Basis for Reasoning about Commonsense Psychological Knowledge. In: Proc. 11th Annual Conference of the Cognitive Science Society, Michigan (1989)
8. Sowa, J.F.: Knowledge Representation–Logical, Philosophical, and Computational Foundations. Pacific Grove, California, Brooks/Cole (2000)
9. Steimann, F.: On the Representation of Roles in Object-oriented and Conceptual Modeling. In: Data & Knowledge Engineering (2000) 83–106
10. Tian, W.: Research on the Human Psychological Commonsense. Ph.D. Thesis, Institute of Computing Technology, Chinese Academy of Sciences (2003)
11. Yong, X., Cao, C.G.: Acquiring Emotion Knowledge of Anger from WWW. In: Poster Proc. 27th Annual German Conference on Artificial Intelligence, Germany (2004) 137–144
12. Yong, X, Cao, C.G.: A Formal Representational Model of Emotion Knowledge. In: Proc. 2nd International Conference on Knowledge Economy and Development of Science and Technology, Beijing, China (2004) 310-316

User Preference Learning for Multimedia Personalization in Pervasive Computing Environment

Zhiwen Yu[1], Daqing Zhang[2], Xingshe Zhou[1], and Changde Li[1]

[1] School of Computer Science, Northwestern Polytechnical University, P.R. China
{zhiwenyu,zhouxs}@nwpu.edu.cn, changde_lee@126.com
[2] Context Aware Systems Department, Institute for Infocomm Research, Singapore
daqing@i2r.a-star.edu.sg

Abstract. Pervasive computing environment and users' demand for multimedia personalization precipitate a need for personalization tools to help people access desired multimedia content at anytime, anywhere, through any devices. User preference learning plays an important role in multimedia personalization. In this paper, we propose a learning approach to acquire and update user preference for multimedia personalization in pervasive computing environment. The approach is based on Master-Slave architecture, of which master device is a device with strong capabilities, such as PC, TV with STB (set-on-box) or PDR (Personal Digital Recorder), etc, and slave devices are pervasive terminals with limited resources. The preference learning and update is done in the master device by utilizing overall user feedback information collected from different devices as opposed to other traditional learning methods that just use partial feedback information in one device. The slave devices are responsible for observing user behavior and uploading feedback information to the master device. The master device is designed to support multiple learning methods: explicit input/modification and implicit learning. The implicit user preference learning algorithm, which applies relevance feedback and Naïve Bayes classifier approach, is described in detail.

1   Introduction

The rapid advances of communication technologies precipitate more and more embedded computing devices, such as PDAs, cell phones, etc., to be used in a wireless environment. This has led to a shift from traditional computer-centered to human-centered information access mode, which is known as *Pervasive Computing*. Multimedia information is widely used in pervasive computing environment in many application fields, such as multimedia digital libraries, home entertainment, live camera remote surveillance, etc. A major trend and requirement in today's multimedia information service is personalization. The capability to model and learn user interests is at the heart of a personalized information filtering system [1]. It is also true for multimedia personalization. Since the interest of a user is changing as time goes by, the quality of personalization mainly depends on whether the user profile really reflects the user preference. So we can conclude that user preference learning plays an important role in multimedia personalization.

Acquiring and updating user preference for multimedia personalization in pervasive computing is challenging due to the following factors:

- The poor human-machine interactivity of pervasive devices causes user explicitly inputting his preference to be nearly impossible.
- The limited computing power and storage make complex learning algorithm impractical.
- The mobility of user and devices causes insufficient time to acquire and update user preference. In pervasive computing environment, multiple devices are attached to a user and used at anytime anywhere. Each of the devices can gather user feedback information, but may be very fractional.

Fortunately, many of today's users own a device with strong capabilities, such as PC (Personal Computer), TV with STB (set-on-box) or PDR (Personal Digital Recorder), etc. In this paper, we address the above challenges by proposing a centralized learning approach with the "strong capability device" been used in conjunction with the mobile user devices for user preference acquisition and update. Our approach collects fragments of user feedback information in different pervasive devices so as to build an abundant feedback repository, and then use a strong capability device to implicitly learn user preference from the feedback repository.

Our early work, TV3P [2] employs an implicit and explicit profiling scheme for personalized TV experiences by integrating explicit input/modification, explicit feedback, and implicit feedback. However, it mainly serves for TV or desktop environment, but not for pervasive computing environment.

The approach proposed here has several advantages. First, it learns user preference by utilizing overall feedback information as opposed to other traditional methods that just use partial feedback information in one device. Second, our approach can relieve pervasive devices with limited resources from computation-consuming and storage-consuming learning tasks. Third, it can make full use of existed user profile residing in other devices. For instance, a user has watched television program through TV set for a long period, so the user profile in the TV set may be comprehensive and correct. When the user wants to watch program through his PDA, he can export his/her preference information from the TV set, and no further learning is required.

2   Overall Learning Approach

The preference leaning is done in a strong capability device, while the pervasive devices are relieved from computation-consuming and storage-consuming learning tasks. The only thing the pervasive devices need to do is to observe user behavior and upload feedback information to the strong capability device. We call the device with strong capability as *master device*, and corresponding pervasive devices as *slave device*.

Fig. 1 shows the schematic architecture of the learning strategy. There are three primary elements: multimedia service provider, master device, and slave device. There are six interfaces between different elements in the whole learning process. They are depicted as follows:

1) When the user wants to access personalized multimedia information through his/her pervasive mobile devices, the slave device first connects with his/her master device asking for the latest user profile.
2) Then the master device sends user preference to the slave device.
3) The slave device sends a request to multimedia service provider, encapsulating user preferences, slave device capabilities, and network characteristics into the request information.
4) Now, the service provider selects and delivers multimedia information to the slave device in an appropriate modality, e.g. video, audio, image, text, or synthetic model, so as to meet the display capability of the device.
5) In the viewing process, the slave device observes user behavior for a specific multimedia content, and sends the feedback information to the master device. The feedback captured by the slave device can be sent at once when it is obtained or sent after a set of feedbacks gathered.
6) When the user profile is updated, the master device sends the latest user profile to the slave device. If the user will, the process goes to (3), otherwise keeps the updated user profile in the slave device.

The master device is designed to support multiple methods for preference learning: explicit input/modification and implicit learning. Explicit input/modification provides Graphic User Interface for the user to input interests when registration or modify preference after login. Implicit learning analyzes user viewing history or feedback information generally applying probability statistics, and then updates user preference. The feedback information is collected through any devices at anytime anywhere. The implicit learning algorithm is described in Section 3 in detail.
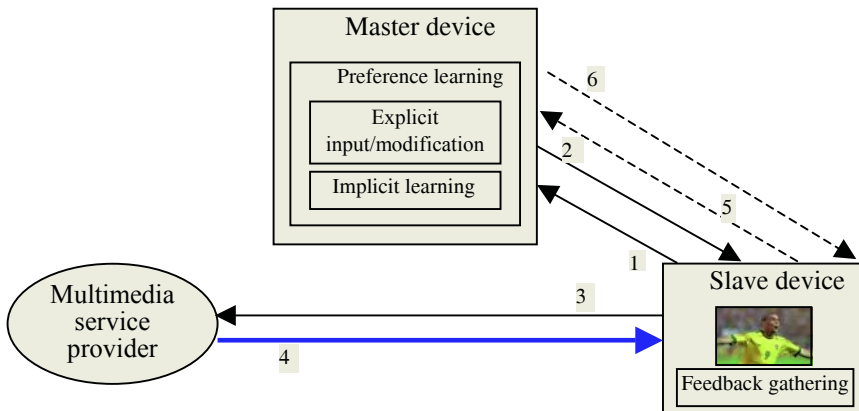
Fig. 1. System schematic architecture

In a pervasive computing environment, the working spot of a device may be at home, in office, on a train, or on road, etc. A diverse set of network technologies are used by different slave devices to communicate with master device. For instance, the master device may be a PC at home. The slave devices may include an interactive TV, PC in office, laptop, PDA, and cellular phone. To communicate with the master device, the interactive TV at home can use IEEE 802.11/Bluetooth, PC in office can use Ethernet, laptop can use ISDN, PDA and cellular phone can use GPRS.

3 Implicit User Preference Learning Algorithm

The general model of the implicit learning algorithm is based on relevance feedback [3], which is an effective and efficient information retrieval technique that can be used to adjust user preference approaching to user's real interests. The algorithm is defined as follows:

$$W_i' = \alpha \times W_i + \beta \times F_{pos} - \gamma \times F_{neg} \tag{1}$$

where W_i' is the updated weight of feature f_i; W_i is the initial weight of feature f_i. α, β and γ are the feedback parameters to be set. The well-known and effective feedback parameters is instantiated by Rocchio [3] as $\alpha = 1$, $\beta = 2$, and $\gamma = 0.5$. F_{pos} and F_{neg} represent the weights of positive feedback and negative feedback to feature f_i respectively.

The positive estimation (F_{pos}) and negative estimation (F_{neg}) were performed by using Naïve Bayes classifier [4] approach. All user actions for multimedia objects are divided into two classes, *Positive* or *Negative*, that is, positive feedback records class *PC*, and negative feedback records class *NC*. The records are about user's viewing attributes and corresponding viewed multimedia objects. The general form is:

```
< class | action time | action type | watching duration | total
duration | title | genre | actor | director | language | keyword >
```

The "class" determines which class (*PC* or *NC*) the action belongs to; "action time" specifies the time that the action takes place; "action type" specifies the type of action performed by the user, e.g., "Record", "View", etc.; "watching duration" specifies the duration of the action; "total duration" specifies the total duration of the multimedia content, such as a TV program; for static modality, such as image, or text, the "watching duration" and "total duration" can be set as the same default value (e.g. 5 seconds); "title", "genre", "actor", "director", "language", and "keyword" are descriptive information about the multimedia content, which can be extracted from the multimedia metadata.

F_{pos} and F_{neg} equals the conditional probabilities of feature f_i in class *PC* and *NC*:

$$F_{pos} = P(f_i \mid PC) \tag{2}$$

$$F_{neg} = P(f_i \mid NC) \tag{3}$$

Suppose $n(PC, f_i)$ means the times of feature *fi* occurring in *PC*, while $n(NC, f_i)$ means the times of feature *fi* occurring in *NC*. $n(PC)$ and $n(NC)$ denote the sum of times of all features occurring in *PC* and *NC* respectively.

For a recommendation item, we consider three action types: Record, View, Delete.

(1) If the action type is "Record", obviously the feedback is *positive*, $n(PC, f_i)$ and $n(PC)$ are calculated as follows:

$$n(PC, f_i)' = n(PC, f_i) + 1 \tag{4}$$

$$n(PC)' = n(PC) + 1 \tag{5}$$

$n(PC, f_i)'$: the updated $n(PC, f_i)$;

$n(PC)'$: the update $n(PC)$;

(2) If the action type is "View" and the recommended item is a video/audio content, whether the feedback is positive or negative depends on the user's real watching duration. We assume that if the real watching duration is longer than a threshold θ (such as 30 seconds), the user really likes the content (the user gives implicit positive feedback); otherwise, the user dislikes it (the user gives implicit negative feedback). $n(PC, f_i)$, $n(PC)$, $n(NC, f_i)$, and $n(NC)$ are calculated as follows:

$$n(PC, f_i)' = n(PC, f_i) + \frac{T_r - \theta}{T_t} \qquad T_r \geq \theta \tag{6}$$

$$n(PC)' = n(PC) + \frac{T_r - \theta}{T_t} \qquad T_r \geq \theta \tag{7}$$

$$n(NC, f_i)' = n(NC, f_i) + \frac{\theta - T_r}{\theta} \qquad T_r < \theta \tag{8}$$

$$n(NC)' = n(NC) + \frac{\theta - T_r}{\theta} \qquad T_r < \theta \tag{9}$$

T_r : user's real watching duration;

T_t : total duration of a specific multimedia object;

θ : the threshold of the time duration;

$n(NC, f_i)'$: the updated $n(NC, f_i)$;

$n(NC)'$: the updated $n(NC)$;

If the recommended item is an image or text, the action "View" is regarded as the same as "Record". $n(PC, f_i)$ and $n(PC)$ are calculated according to Equation (4) and (5).

(3) If the action type is "Delete", obviously the feedback is *negative*, $n(NC, f_i)$ and $n(NC)$ are calculated as follows:

$$n(NC, f_i)' = n(NC, f_i) + 1 \tag{10}$$

$$n(NC)' = n(NC) + 1 \tag{11}$$

Suppose $|V_1|$ and $|V_2|$ denote the total number of features occurring in PC and NC respectively. According to Lidstone's Law of succession [5], $P(f_i | PC)$ and $P(f_i | NC)$ can be estimated as follows:

$$P(f_i | PC) = \frac{n(PC, f_i) + \lambda}{n(PC) + \lambda |V_1|} \tag{12}$$

$$P(f_i | NC) = \frac{n(NC, f_i) + \lambda}{n(NC) + \lambda |V_2|} \tag{13}$$

λ is a positive number, normally between 0 and 1. The case of $\lambda = 0.5$ is called Jeffrey Perks Law or Expected Likelihood Estimation (ELE), which is widely adopted in various applications. If λ is set as 1, the Lidstone's Law is the same as Laplace's Law.

If the updated W_i' is larger than its upper bound, let $W_i' =$ upper_bound (e.g. 1.0).

If the updated W_i' is less than its lower bound, let W_i' =lower_bound (e.g. −1.0);

If the absolute value of updated W_i' is not less than a preset threshold ξ (that is, $|W_i'| \geq \xi$), we will keep it in the user's profile, otherwise discard it, because it is too trivial.

4   Performance Evaluation

We evaluate system performance in terms of learning speed and learning efficacy. Learning speed is crucial for mobile devices, because long-time computation will consume large battery power. Learning efficacy, on the other hand, reflects how much the learned preference approaches to the user's real interests, which directly influences the quality of personalization. The data sets used for experimentation and performance analysis were taken from the Internet Movie Database (IMDb, http://us.imdb.com). The description information for each movie in IMDb is abundant, which includes title, director, genre, actor, keywords, language, country, etc. For each movie, 15 features were extracted from its description information in IMDb to represent its metadata in our system.
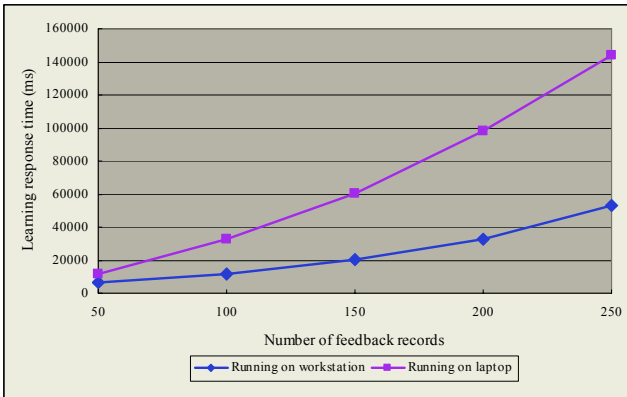
Fig. 2. Experimental result of learning speed

Experiment 1: We evaluate learning speed by measuring learning response time on the master device (workstation) and the slave device (laptop) respectively. We compared the time costs by varying the number of feedback records ranging from 50 to 250. The evaluation result is shown in Fig. 2, from which we can see that the user preference learning is computationally intensive and the learning time is largely dependent on the size of the feedback records. We also can observe that the learning response time on the laptop takes much longer than that on the workstation, and with the number of feedback records increasing from 50 to 250, the time difference is remarkable. When the size of feedback records reaches 250, the learning time on the laptop is very long (144s), which consumes much battery power and causes inconvenience to the user. But the learning time on the workstation is much less and accept-

able. This result proved that the idea of running preference learning on a strong capability device other than pervasive mobile devices is appropriate.

Experiment 2: We use *precision* [6] to evaluate the efficacy of user profile learning. It denotes the ratio of the number of movies that user really interested and recorded to the total number of movies recorded. We took 50 multimedia contents as training set, and 200 contents as testing set. The testing set is divided into 5 sessions with each session having 40 contents. We first let the user browse all of the movies' title and plot beforehand, and classify them into two classes: *interested* and *not-interested*. Then, the system chooses and records movies for the user according to the user profile learned. After every session finished, we calculated the values of precision. The result is encouraging with the average precision value being 0.83. The experimental result proved that our system could keep track of user preference changing over time, and perform good filtering effectiveness, i.e. record most of the movies that the user likes.

5   Conclusion

In this paper, we propose an approach of user preference acquisition and update for multimedia personalization in pervasive computing environment. The major contributions of the paper are: (1) introducing a centralized learning approach to utilize overall feedback information and relieve pervasive devices from learning tasks; (2) providing the implicit user preference learning algorithm which applies relevance feedback and Naïve Bayes classifier approach.

Acknowledgments

References

1. Dwi H. Widyantoro, Thomas R. Ioerger and John Yen: An Adaptive Algorithm for Learning Changes in User Interests. Proc. of the ACM Intl Conf. on Information and Knowledge Management, Kansas City, Missouri, USA (1999) 405-412
2. Zhiwen Yu and Xingshe Zhou: TV3P: An Adaptive Assistant for Personalized TV. IEEE Transactions on Consumer Electronics. Vol. 50, No. 1 (2004) 393-399
3. J.J.Rocchio: Relevance feedback in information retrieval. The Smart System--Experiments in Automatic Document Processing. Prentice Hall (1971)
4. Joachims T.: A probabilistic analysis of the rocchio algorithm with TFIDF for text categorization. Proc. of the 14th Intl Conf. on Machine Learning (1997) 143-151
5. Lidstone, G.J: Note on the general case of the Bayes-Laplace formula for inductive or a posteriori probabilities. Transactions of the Faculty of Actuaries, 8:182-192.
6. C.J. van Rijsbergen: Information Retrieval. Butterworths, second edition (1979)

Emotion-Based Smart Recruitment System

Rajiv Khosla and Chris Lai

Business Systems & Knowledge Modelling Laboratory
School of Business, La Trobe University, Victoria 3086, Australia
{R.Khosla,C.Lai}@latrobe.edu.au

Abstract. Emotions form an important component of human behaviour and decision making. In this paper we employ image processing and soft computing techniques to design and implement emotional state profiling of a sales candidate while they provide psychological inputs related to their selling behaviour to an intelligent sales recruitment and benchmarking system. The work reported has implications in not only recruitment and benchmarking of sales persons but also web personalisation and internet based system in general.

1 Introduction

There are many situations involving the use of computer, internet and mobile communication devices in which emotions of a user can play an important role. Education, selling and customer behaviour, e-tourism (destination selection and planning), and web information personalisation are instances of such situations [1,2]. Facial expressions are an important physiological indicator of human emotions [3]. The existing work on analysing human emotions has primarily centred around embodiment of emotional or affective characteristics like happiness, anger and fear in robots or software agents [4]. Some work has been done in recognition of user emotions in highly controlled environments [5]. In these controlled environments users have to mimic prototypical emotional expressions like happy and angry which are then recognized by the computer or an emotional software agent. However, there is very little research done in terms of processing emotions for providing intelligent e-decision support in natural or real-world environments [6].

Most existing approaches of salespersons' recruitment rely on interview processes and psychometric techniques [7] for evaluating behaviour profiling and behaviour categorization of a sales candidate. These approaches have met limited success. We have developed Salesperson Recruitment and Benchmarking System (SRBS) [1], which is based on integration of selling psychology based behaviour model and intelligent technologies like expert systems and self organizing maps. SRBS employs psychological data to determine a sales candidate's selling behaviour category.

Within the context of personality, temperament and moods, emotions rise and fall in short-term cycles. In this paper we report on the correlation of a candidate's emotional state profile with that candidate's selling behaviour profile. The two profiles and their correlation will provide smart recruitment decision support to a sales or human resource manager.

The paper is organized as follows. Section two outlines the selling behaviour model. Section three outlines an affect space model and section four outlines design of a dynamic facial expression analysis system for capturing the emotional state information of the candidate. Section five presents preliminary results of the system and a discussion on the correlation between the psychological emotional data inputs and their implications for user profiling in general. Section six concludes the paper.

2   Selling Behavioural Model

The behavioural model developed by Buzzotte, Lefton & Sherberg [8] has been used for predicting selling behaviour profiles of salespersons [9]. The two dimensional behavioural model [9, p171] used for predicting four behavioural categories, namely, Dominant-Warm (DW), Dominant-Hostile (DH), Submissive-Warm (SW) and Submissive-Hostile (SH) is shown in Figure 1. This model has been used based upon interactions with senior managers in the sales and human resources arena in the consumer and manufacturing industries in Australia [9].

Dominant

Dominant-Hostile	**Dominant-Warm**
The salesperson must impose their will on the customer by superior determination and strength. Selling is a struggle the salesperson must win.	Sales are made when customers become convinced that they can satisfy a need by buying. The salesperson's job is to demonstrate to the customer that their product would best satisfy the customer's need.

Hostile   **Warm**

Submissive- Hostile	**Submissive- Warm**
Customers buy only when they are ready to buy. Since persuasion does not work, salesperson's job is to take their order when the customer is ready to give it.	People buy from salespersons they like. Once a prospect becomes a friend, it is only reasonable that he should also become a customer.

Submissive

Fig. 1. Salesperson Behaviour Profile [9, p171]

For analysing the selling behaviour profile of a salesperson 17 areas related to the model in Figure 1 have been identified for evaluation of a sales candidate behaviour profile as *selling as a profession, assertiveness, decisiveness, prospecting, product, customers, competition, success and failure, boss, peers, rules and regulations, expenses and reports, training, job satisfaction, view about people, relationship with non-selling departments, general attitudes* [10]. These areas have been identified after several discussions with sales managers and knowledge available in the literature [8, 11, 12]. Weights have been assigned to 17 areas on a scale of 1 to 10 using AHP (Analytical Hierarchy Process) technique [13]. The selling behaviour attributes associated with the 17 areas are designed in the form of a questionnaire. The questionnaire

consists of 76 questions with at least four questions corresponding to each area. The psychological inputs or answers provided by a sales candidate (invited for a sales recruitment interview) are used for determining the selling behavioural profile of the candidate. A sample set of four questions related to the area of *success and failure* is shown in Figure 2. In the rest of the paper we describe the design, implementation and correlation of emotional state profiling of a sales candidate while they are providing the psychological inputs to the 76 questions.

1. I hate to lose or fail. I just don't seem to be able to digest failure. Success matters most in life.	Behavioural Category:	DH
2. You take failures in stride. People are important up to a point. Success and progress is the ultimate goal.	Behavioural Category:	DW
3. You do your job. Failures mostly happen due to product and pricing policies of the company	Behavioural Category:	SH
4. It is more important to be associated with people than thinking of success or failure.	Behavioural Category:	SW

Fig. 2. A sample set of four questions related to the area of *success and failure*

3   Affect Space Model

Facial expressions are an important physiological indicator of human emotions [3]. An affect space model used by psychologists [14] is shown in figure 3. The model involves three dimensions, namely, Valance (measured on a scale of pleasure (+) to displeasure (-)), Arousal (measured on scale of excited/aroused (+) to sleepy (-)) and Stance (measured on a scale of high confidence (+) and low confidence (-)). Facial expressions correspond to affect states like happy, surprise and tired. Figure 3 shows the affect space model with several labelled emotional states. The model can be divided into quadrants, each quadrant being considered to represent positive or negative emotional states.

Like in everyday life, in human-computer interaction people's emotions are characterised more by subtle variations or transient changes in facial/emotional expressions (during the interaction) rather than as prototypical emotional expressions [6].

The proposed system attempts to make use of the candidate's emotional state to determine the correlation or commitment the candidate has to the entered response. Rather than attempt to determine the absolute emotional state of a person or sales candidate i.e., exactly where their emotional state lies in the affect space model, we have modelled a change in emotional state in either a positive direction or a negative direction. It is proposed that a positive change in emotional state of the candidate that coincides with the answering of a question indicates a candidate's higher commitment to the answer given. Conversely a negative emotional state change indicates a reduced commitment of the candidate to the answer given. Note that what we are talking about here is change in emotional state with time and whether this change is in a direction towards a positive or negative quadrant of the affect space model. The psychologists point out that facial expression alone may not be an accurate indicator of the emotional state of a person but changes in facial expression may indicate a change in
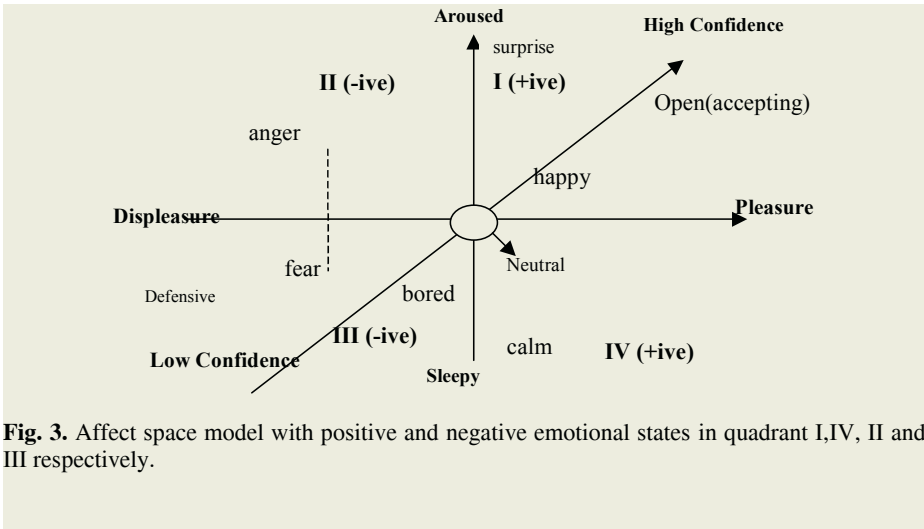
Fig. 3. Affect space model with positive and negative emotional states in quadrant I,IV, II and III respectively.

emotional state [15, 16]. In our case this is exactly what we want to model, that is the transient or temporal changes in facial expressions and emotional states as the sales candidate answers different questions.

4   Dynamic Facial Expression Analysis

The main aim of the work reported here is to track transient facial expression and correlate emotional state indicated by that facial expression with the selling behaviour profile of a sales candidate. To capture the transient facial expressions of a candidate a webcam was used to record video of the candidate while answering SRBS profiling questions. A timestamp was recorded with question answers for temporal correlation between recorded video and answers.

4.1   Pre-processing

Pre-processing consisted of finding the eye coordinates in the video stream, geometric normalisation centred on the eyes and generation of difference images. To expedite the location of the eyes for the purposes of this study a simple method of placing two green dots on the candidate's forehead while recording video was used. Eye coordinates where estimated relative to these green dots.

Define the current image I_t , the image n frames earlier I_{t-n} and the difference image I'_t then we have

$$I'_t = I_t - I_{t-n} \qquad \text{Eq. 1}$$

n was selected to be 3 i.e. the subtracted image was the image three frames earlier. The video was recorded at 10 frames per second. So the time difference the difference image represented was approximately 300 milliseconds. Figure 4 shows three images and their difference images.

4.2   Representation

Two common techniques being used for feature representation in face recognition and analysis are eigenvectors [17] and Gabor wavelets [18]. Gabor wavelet filters at varying scales and orientations have been used to represent facial expression images [19]. We adopt this methodology to represent the difference images in this paper.

Image representation was by a vector of Gabor Jets calculated at points on a grid overlaid on the difference image I'_t. A Gabor jet is the magnitude of the complex valued Gabor wavelet filter output when applied to the image. The filters used were at three scales an octave apart and at six orientations within each scale. The bandwidth of the masks was approximately an octave. For the readers reference, the filters response is given in [19].

4.3   Neural Network Classifier

The neural net was trained on a selection of image sequences from the Cohn-Kanade facial expression data base [20]. The sequences of images in this database are taken from video and so it was assumed the frame rate is approximately 30 frames per second. For the difference images generated to train the neural net n in equation 1 was selected to be 8 representing a time difference of approximately 270 milliseconds. The sequences selected were the ones that represented joy and anger and were classed into positive and negative emotion respectively. Difference images representing no change i.e. relatively flat images where generated artificially and were classed as neutral emotion i.e. no change in emotional state. The neural net architecture was 1296 input nodes, corresponding to the dimensionality of the Gabor jets vector, 10 hidden nodes and 3 output nodes. Backpropagation was the training algorithm used.

For display purposes the output of the network was visualised by an image where different areas of the image represented different classes and/or relative strength of the three different classes namely neutral, positive or negative emotion state change. The output of the neural net, three nodes representing the three classes, where mapped to red, green and blue for negative, positive and neutral change respectively. The top half of the visualisation image was a colour representing all three classes' colour mixed together. The bottom half of the image was divided up into three equal areas, each devoted to each of the three colour/classes. Ideally it would be expected to display either of red, green or blue. However mixing all three classes in the top half if the display image makes it easy to judge where the neural network classification contained an overlap. Examples of the classification display images are shown in figure 4.

5   Results

Presented here are some preliminary results in support of the proposed concepts outlined in section 3. Figure 4 shows an excerpt from the video sequence of a volunteer candidate answering questions from SRBS. The expressions presented are not contrived, that is, the expressions are genuine responses to the questions being presented. The candidate was not asked to make expressions consciously to indicate what he felt was his emotional state. Figure 4 also shows the difference images and visualisations

of the neural network classification. The timing of this video image sequence is approximately 300 ms between frames shown, i.e. n = 3 as described in section 4.1. The sequence shown is taken at a time between a new question being presented to the candidate and the candidate answering that question. Thus it can be assumed the facial expression indicates an emotional response to the presented question. The question was "You take failures in stride. People are important up to a point. Success and progress is the ultimate goal." And the candidate answered: 'To a large extent yes'. This question relates to the area of *success or failure* and the Dominate-Warm behaviour category outlined in section 2.
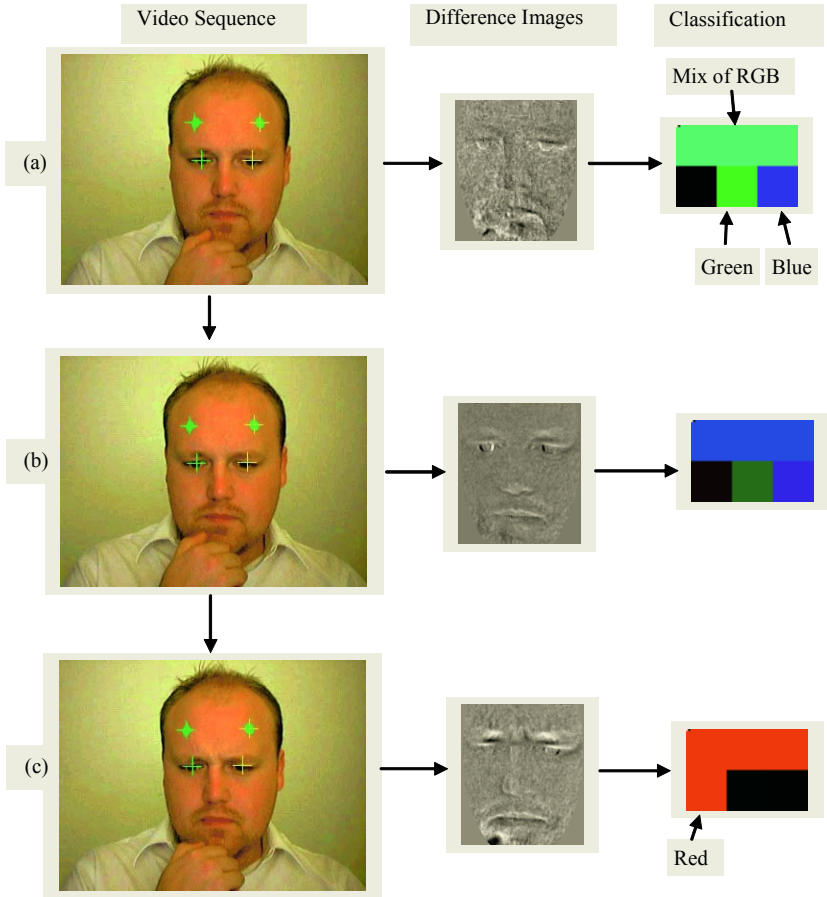
Fig. 4. Sequence of images from video, approximately 300 milliseconds apart, with their difference images and visualisation of their neural net classification. (the image prior to the first image used to generate the first difference image is not shown)

Figure 4 (a) was classified as a mix of roughly equal proportions of neutral and positive as indicated by the blue and green respectively and absence of red i.e. negative emotion change in the bottom half of the classification output and the cyan of the

top half. Figure 4 (b) was classified as primarily neutral indicated by the diminished green in (b) and the shift from cyan to a more blue colour in the top half of the classification image with respect to (a). Figure 4 (c) classification image indicates a dominance of red. It is not asserted that the classification for (a) and (b) is as subtly accurate as indicated. The subtle variations in these two are more likely to be the system, in its present state, not coping very well with noise and other non-linear artefacts introduced in the recording and pre-processing which need to be refined, as well as the classifier, to get more consistent results at a higher level of subtle accuracy. However the complete absence of red or negative emotional change is obvious in (a) and (b). This is contrasted with the dominance of red i.e. negative emotional change and absence of green and blue in (c). The candidate has developed a frown in video image (c) compared to (a) or (b). This intuitively correlates well with the classifier result.

So a negative emotional transition by the candidate in response to a question relating to the area of *success and failure* and dominant-warm behaviour category has been observed. Figure 5 is a graph of the candidate's answers to questions in the area of success and failure, affirmative and negative for each of the behaviour categories. The graph also shows his emotional state as detected by this system on the same axis. The continuous line represents answers and the dashed line is emotional state. There is a divergence with the question relating to the dominant-warm, DW, behaviour category. This correlation indicates a conflict with the answer given by the candidates in the indicated area/behaviour category. This should translate, in practice, to a follow up in an interview if this system were to be used in a recruitment situation.
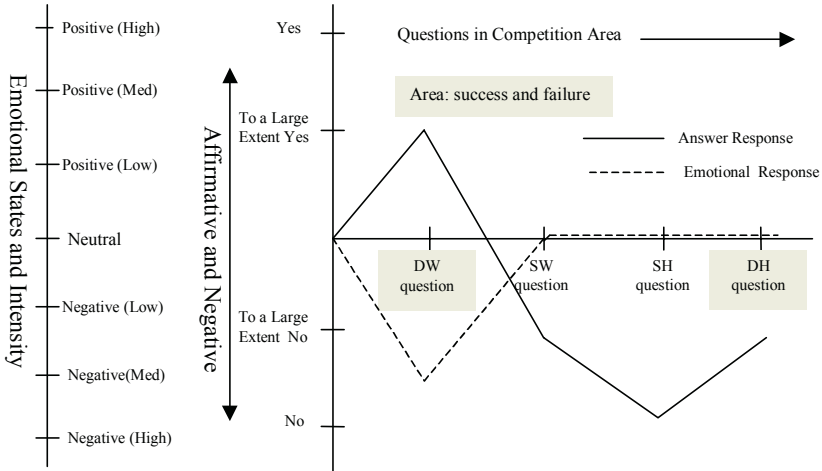
Fig. 5. Correlation of emotional state profile and selling behaviour profile.

6 Conclusion

In this paper we have reported a unique comparison of psychological data based selling behaviour profile with emotional state profile of sales candidate for smart recruitment decision support. The work reported in this paper can also be applied in web personalisation and other decision support systems.

References

1. R. Khosla, C. Lai and T. Goonesekera, "Behaviour Profiling Based on Psychological Data and Emotional States", in LNCS/LNAI proceedings, 8th International conf. on Knowledge-based Intelligent Information and Eng. Systems, N.Zealand, Sept. 2004.
2. S. Gountas. & J. Gountas , The Influence of Consumers' Emotions on their Service Product Evaluation,Tourism Analysis, vol. 8,No. 2-4, p125, 2004.
3. Facial expression analysis and Human emotions - http://www.ri.cmu.edu/projects/project_10.html
4. Bartneck, C., Reichenbach, J., & Breemen, A. v. In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions. Proceedings of the Design and Emotion Conference , Ankara 2004.
5. I. Essa et. al.. 'Coding, analysis, interpretation and recognition of facial expressions. Pattern Analysis and Machine Intelligence,, 7:757-763, July 1997.
6. A. Kapoor, et. al., Fully Automatic Upper Facial Action Recognition, in IEEE Workshop on Analysis and Modeling of Faces and Gestures , Oct 2003.
7. Murphy, K. A. and Shon, R. De :Progress in Psychometrics: Can Industrial and Organizational Psychology Catch Up?, Personnel Psychology vol. 53 913-924 2000.
8. Buzzotte, V.R., Lefton, R.E. and Sherberg, M.: Effective Selling Through Psychology: Psychological Associates New York 1981.
9. Khosla, R. Goonesekera, T.and Mitsukura, T.:Knowledge Engineering of Intelligent Sales-Recruitment System Using Multi-Layered Agents Methodologies, presented at 14th International Symposium on Methodologies for Intelligent Systems (ISMIS), 28-31 October Maebashi Japan 2003.
10. Khosla, R. and Goonesekera, T.: An Online Multi-Agent e-Sales Recruitment Systems, presented at IEEE/WIC International Conference on Web Intelligence (WI) Halifax Canada 2003.
11. Szymanski, D.M.: Determinants of Selling Effectiveness: The Importance of Declarative Knowledge to the Personal Selling Concept, J.of Mkting vol. 52 64-77 1988.
12. Weitz, B.A., Sujan, H. and Sujan, M.:Knowledge, Motivation and Adaptive Behavior: A Framework for Improving Selling Effectiveness, J.of Mkting vol. 50 174-191 1986.
13. Saaty, T.L.: The Analytic Hierarchy Process, NY McGraw Hill, 1980.
14. P. J. Lang. The emotion probe: Studies of motivation and attention. American Psychologist, 50(5):372-385 1995.
15. A. Irfan et. al., A , 'Coding, Analysis, Interpretation, and Recognition of Facial Expressions', IEEE Trans on Pattern analysis and Machine Intelligence, vol 19, no. 7, July 1997.
16. C.E. Izard, "Facial Expressions and the Regulation of Emotions,"J. Personality and Social Psychology, vol. 58, no. 3, pp. 487-498, 1990.
17. M.A. Turk, A.P. Pentland, "Face Recognition Using Eigenfaces", Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on , Pages:586 – 591, 3-6 June 1991.
18. T. Lee, "Image Representation Using 2D Gabor Wavelets.", IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(10):9590971, Oct. 1996.
19. S. Akamatsu, J. Gyoba, M. Kamachi, M. Lyons, "Coding facial expressions with Gabor wavelets", Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on , Pages:200 – 205, 4-16 April 1998.
20. Kanade, T., Cohn, J.F., and Tian, Y., "Comprehensive Database for Facial Expression Analysis", The 4th IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), France. March 2000.

Evolvable Recommendation System in the Portable Device Based on the Emotion Awareness

Seong-Joo Kim[1], Jong-Soo Kim[1], Sung-Hyun Kim[2], and Yong-Min Kim[3]

[1] R&D Center, Multichannel Labs. Co., Ltd.
#401-1302 Bucheon Techno Park, 193 Yakdae-dong Wonmi-gu, Bucheon, 420-734, Korea
[2] Dept. of Electronics, Tongwon College
Sinchon-ri, Silchon-eup, Gwangju-si, Kyounggi-do, Korea
[3] School of Computer Engineering, Chungcheong College
330 Weolgok-ri, Kangnae-myon, Cheongwon-gun, Chungbuk-do, 363-792, Korea
bluebird@xclef.com

Abstract. Recently, the portable devices became very popular and many people have a portable device. The functionality of the portable device is demanded to be advanced more highly. Moreover, the portable device has intelligence as a result of technical advance. In fact, people hope that the portable device can work and operate by itself for user's satisfaction. This paper introduces the portable device that can be an intelligent and evolvable system and also can recognize the emotional awareness. The emotional awareness means the feeling of user who uses the portable device. It is very intelligent function for the portable device to provide suitable operation for the user being based on the status of user's emotion. In this experiment, the suitable operation generated by portable device will be intelligent recommendation function such as genre selection or music recommendation. In order to show the performance of the proposed intelligent portable device, the genre selection will be introduced as a result. The intelligent portable device having an evolvable recommendation function may have various functions in the future according to the type of intelligent application and the accuracy of emotional awareness.

I Introduction

This paper proposes the evolvable recommendation system in portable device based on the emotion awareness with fuzzy inference engine and neural network. The evolvable recommendation in the portable device is generated with an intelligent module, which includes intelligent functions. As a result, the portable devices will be more friendly and efficient to user. Recently, the mobile and portable technologies are rapidly changed and there also are too many portable devices for multimedia. Finally, the functions generated by each device will be similar to each others. It may be called as a digital convergence. A digital convergence means that the function in a device is applied for another device keeping its original function unchanged.

As such a reason, the intelligent module which can operate much intelligence can be introduced instead of the separated and various modules which may be applied for only a specific device. Among many intelligent items, in this paper, the proposed portable device has the intelligent module generating the emotional decision.

The proposed portable device can perform the basic function as a portable multi-media device and also intelligent functions such as menu selection and media recommendation by considering the emotional status of the user. It is very useful for the portable device to provide intelligent functions during operation. On behalf of providing intelligent functions, the portable device recognizes the emotional status of user using several parameters those can be used for decision of emotion. The intelligent functions based on emotional recognition are in friendly relations with user. The device having intelligent method on several interface and algorithm has a meaning of humanoid system. The proposed intelligent portable device will be an example of the model in the future. In order to introduce the intelligent portable device, in this paper, the intelligent module will be proposed as an independent module.

The rest of this paper will be organized as follows.

In the chapter II, the fundamental intelligent method will be introduced and the intelligent module for portable device will be introduced at the chapter III. The implementation of portable device will be introduced in the chapter IV, the result of experiment will be shown in the chapter V. Finally, the conclusion and future works will be made in the chapter V.

II Fundamental Intelligent Method

A. Basic Concept for Implementation

This paper introduces the implementation of intelligent recommendation for portable device by generating emotional recognition module. First of all, the basic concept for implementation is as follows.

In case of using MP3 player, the user used to select function, 'random play'. The user may complain on the function, 'random play', because of simple repeat for the selected music. On the other hand, in case of 'random play', the user may complain on the selected music according to his (her) emotional status. Therefore, the function that can recommend and play songs considering the emotional status of the user even in the random play mode may be regarded as an advanced and human-kindly function. Otherwise, the emotional status of the user may be various according to the users and the feeling about the recommended song may be different. Therefore, in this paper, the intelligent recommendation in the portable device recognizing the emotional status according to the user will be implemented by using MP3 player. Furthermore, the MP3 player is just an example of portable devices which can be represented by digital convergence and the application of the proposed intelligent recommendation module can be applied to the other portable devices. The contents which can be implemented by the proposed intelligent module may include the game types or multimedia contents.

For the proper implementation of the intelligent module, the learning process should be contained and the learning process can be generated by the neural network structure and theory.

B. Fuzzy Inference Engine (FIE)

The fuzzifier has the effect of transforming crisp measured data (e.g., blood pressure is 120) into suitable linguistic values (i.e., fuzzy sets, for example, blood pressure is

normal or medium). The fuzzy rule base stores the empirical knowledge of operation processing the domain experts. The inference engine is the kernel of a FIE, and it has the capability of simulating human decision making by performing approximate reasoning to achieve a desired decision strategy.

As mentioned above, FIE has fuzzy rule base, decision rule and is composed with inference engine. The decision rule of fuzzy system is the form of "IF-THEN". The example of "IF-THEN" rule is as follows:

- IF BP is PB and BT is PB, THEN Emotion is BAD.

The notation in the above rule is follows. BP means Blood Pressure, BT means the Body Temperature, PB means Positive Big, BAD means not good and Emotion means emotional status.

In this paper, we use triangular method as a fuzzifier and Mamdani's min-max method as a fuzzy inference method and center of gravity method as a defuzzifier respectively.

C. Neural Network

In general, neural networks consist of a number of simple node elements, which are connected together to form either a single layer or multiple layers. The relative strengths of the input connections and also of the connections between layers are then decided as the network learns its specific task. This learning procedure can make use of an individual data set, after which the strengths are fixed, or learning can continue throughout the network's lifetime, possibly in terms of a faster rate in the initial period with a limited amount of 'open mind' allowance in the steady-state. In certain networks a supervisor is used to direct the learning procedure, and this method is referred to as 'supervised learning', whereas it is also possible to realize self-organising networks which learn in an unsupervised fashion [1].

In this paper, neural networks will learn the relations between user's emotional status and genre for recommendation. As a learning method, back propagation algorithm is used.

III   Intelligent Module for Portable Devices

The flow for implementation of intelligent and human friendly portable devices consists of emotional estimation, inference of suitable genre and learning process.

The flow chart is given as the following diagram.

Fig. 1. Modules for providing intelligent contents

In the figure 1, the input can be obtained from bio sensors, which can measure the biological information such as blood pressure, volume of sweat and blood bit rate.

To estimate emotional status, the hardware interface should be prepared to get information. Fuzzy logic will be used to estimate emotional status organized as a rule base, which explains the relations between emotion and biological information. Generally, fuzzy logic is one of soft computing methods and can be prepared by experts who know relations of system input and output.

In this chapter, sensor module architecture, creation of rule base that can explain emotional status and finally learning module for genre by learning relations and inference result of fuzzy rule base will be introduced.

A. Sensors

In order to estimate the emotional status with bio-information, in this paper, the volume of sweat, pulse, blood pressure and body temperature may be considered. For the experiment, skin resistor sensor to measure the volume of sweat, pulse measure sensor to measure pulse, blood pressure sensor to measure blood pressure and temperature sensor to measure body temperature are used.

The sensors should be stuck to body and the sensor fusion system that can integrate information determined by sensors should be portable type. The sensor fusion system can acquire information at the same time and it is implemented by micro computer. The serial communication may also be used to transfer information and protocol for communication should be pre-defined.

B. Fuzzy Rule Base

The fuzzy rule base can be constructed by bio information from several bio sensors that is explained in the above. The rule base may be different according to personality about the similar information. For implementation, the rule base can be constructed in general case. Bio information can be classified by three kinds such as weak, medium and strong according to normalized sensor value. The following is the figure of fuzzy membership functions.
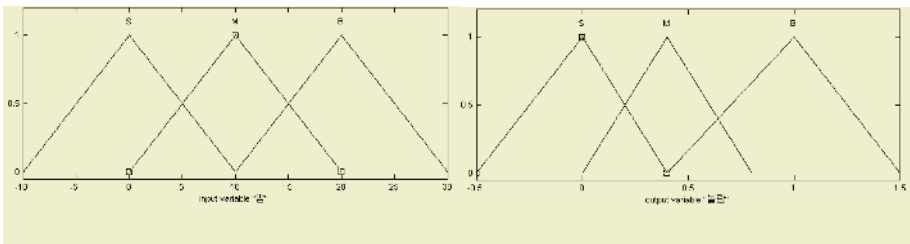
Fig. 2. Membership functions

The above left side figure in figure 2 shows the pre-conditional membership function and also right side one shows the post-conditional membership function.

The premise membership functions are ready for 4 sensory inputs and the membership functions in consequent are ready for 4 emotional states, which include surprising, comfort, pleasure, melancholy.

The below figures show the relations between sensory inputs and emotional states.
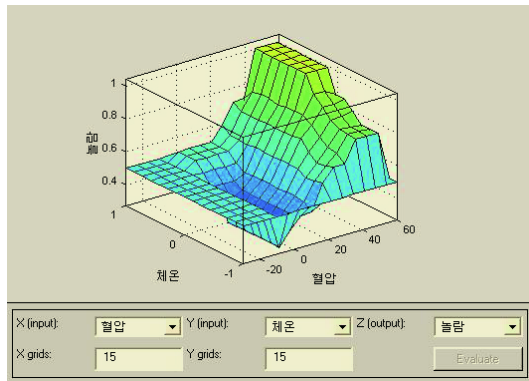
Fig. 3. Relations between blood pressure, body temperature and emotional state, surprising.
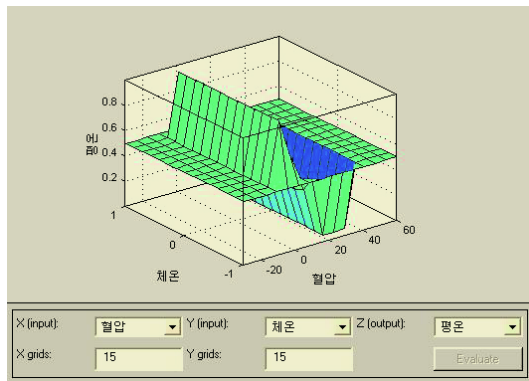
Fig. 4. Relations between blood pressure, body temperature and emotional state, comfort.

In figure 3 and 4, it is known that the relations may be different by inputs and emotional states. In the both figure 3 and 4, the input X is blood pressure, input Y is body temperature and the output Z is calmness.

IV   Experiments

The experiments are generated by simulation in the PC and the results are applied to the portable device, MP3 player.

Using the fuzzy membership functions and pre-defined virtual sensory inputs, the PC can calculate the fuzzy inference engine with its fuzzy membership functions, which is shown in the figure 2. The following figure shows the result of fuzzy inference engine.

In the figure 5, the emotional states can be found. From left side, the characters mean blood pressure, pulse, sweat, body temperature, calmness, surprise, pleasure and melancholy, respectively, by turn.

The following figure shows real implementation on the portable device.
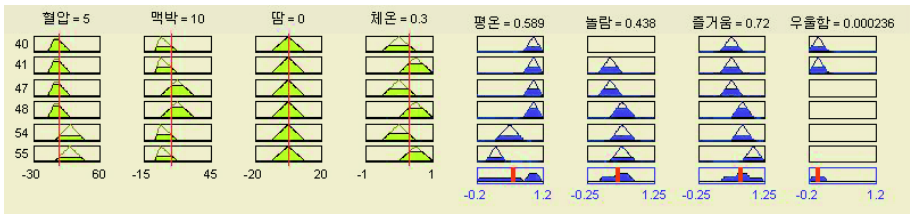
Fig. 5. The emotional states inferred from fuzzy inference engine about 4 sensory inputs.

Fig. 6. The implementation on the portable device using fuzzy inference engine.

Fig. 7. The implementation on the portable device using fuzzy inference engine.

In the figure 6 and 7, the screen on the portable device shows different genre that is selected by fuzzy inference engine according to the sensory input and fuzzy membership functions. The intelligent portable device can support automatic genre selection function using fuzzy inference engine. Automatic selection can be applied to various functions and can be defined intelligent recommendation. The intelligent recommendation may provide end-user who uses portable device friendly and intelligent functional support.

V   Conclusions

In this paper, the intelligent recommendation for portable devices is introduced. The intelligent functional implementation is possible by applying fuzzy inference engine, which uses various fuzzy membership functions and rules. The relations between sensory inputs and emotional states are also analyzed and implemented by fuzzy rules. The relations make portable device support intelligent functional approaches. The methodology introduced in this paper may be applied many functional implementations on the systems as well as portable devices.

In the future, the graphical display according to the emotional state will be applied to the portable device, MP3 player that has already intelligent recommendation function.

Acknowledgement

References

1. G. W. Irwin, K. Warwick, and K. J. Hunt, *Neural Network Applications in Control*, The Institute of Electrical Engineers, 1995.
2. Schmitz, P., "Multimedia goes corporate," *IEEE Transaction on Multimedia*, Vol. 9, Issue 3, pp. 18-21, 2002.
3. Hartwig S., Luck M., Aaltonen J., Serafat R., and Theimer W., "Mobile multimedia-challenges and opportunities," *IEEE Transactions on Consumer Electronics*, Vol. 46, Issue 4, pp. 1167-1178, 2000.
4. B. Zheng, and Atiquzzaman M., "A novel scheme for streaming multimedia to personal wireless handheld devices," *IEEE Transactions on Consumer Electronics*, Vol. 49, Issue 1, pp. 32-40, 2003.

Emotional Extraction System
by Using the Color Combination

Keiko Sato[1], Yasue Mitsukura[1], and Minoru Fukumi[2]

[1] Okayama University 3-1-1, Tsushimanaka, Okayama, 700-8530 Japan
{ged17066,mitsue}@cc.okayama-u.ac.jp
[2] University of Tokushima 2-1, Minami-Josanjima, Tokushima, 770-8506 Japan
fukumi@is.tokushima-u.ac.jp

Abstract. Recently, many researches using the human interface have been done. In particular, the KANSEI information processing is attracted as the multimedia information processing on the human interface. The color coordination system which connects colors with feelings is expected as the system supporting the color design. Therefore, to analyze the relation between colors and feelings is one of problems in the field of the KANSEI engineering. In this research, the method for judging the impression caused by the color automatically is proposed. In this paper, the correlation with the impression caused by the color and the color feature is analyzed as the first stage of this research. Concretely, by using the principal component analysis, the correlation with an amount of the sensibility and the color feature is found.

1 Introduction

In recent years, the consumer get to be concerned about colors, and the color coordination of the product in the multidisciplinary is requested. Usually, the color coordination of the product depend on designer's sensibility, therefore it is difficult to match the color arrangement of the product to the impression that the consumer requests. Particularly, the color design is one of the important elements that influence the impression of the product, therefore the technology which understands and reflects the consumer's sensibility is needed in the color arrangement. Especially, the color coordination system that treats the relation between the color and the sensibility is enumerated as a tool that supports the color design. This system is expected as a system that achieves the advice of the color arrangement and the support of the product design in the dress and the interior and so on. Along with this, the development of the feeling classification system is being done actively. Therefore, requesting the relation between the color and the impression becomes one of problems in the field of the KANSEI engineering. This research pays attention to the influence that the color gives to our feelings, and proposes the method for judging the impression caused by the color automatically. In this paper, the correlation with the color and the impression received by the color is analyzed as the first stage. Concretely, by analyzing the color feature using of the principal component analysis, the correlation with an amount of the sensibility and an amount of the color feature is found. Moreover, in order to show the effectiveness of the proposed method, we invent the application by the color based on the KANSEI information processing.

2   Proposed Method

In this paper, to collect data of the color impression as the initial stage of this re-search, the investigation concerning the color impression is conducted. Next, the color feature is added to the data collected by this investigation, and afterwards the data matrix is created. Finally, feature vectors of colors at each impression are extracted by the principal component analysis. This procedure is done, divided into single color and two colors. Figure 1 shows the flowchart of the proposed method.

Investigation of the color

↓

Addition of the color feature

↓

Creation of the data matrix

↓

Extraction of the feature vector
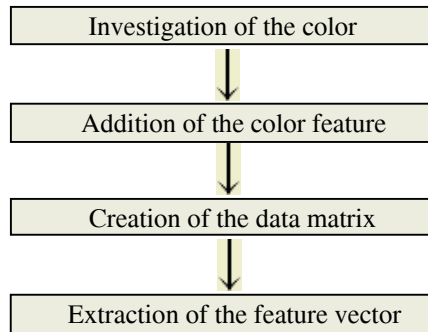
Fig. 1. The flowchart of the proposed method

3   The Investigation of the Color Impression

In this paper, to collect the data of impressions received from the color, the investiga-tion concerning the color impression is conducted. First of all, the application to use of this investigation is made. The data of the color impression can be easily collected by this application.

3.1   The Investigation Procedure

The subjects selects an adjective (impression word) near the impression received the color displayed on the screen from the list of the adjective shown in Table 1. This procedure of each subject is done 60 times about single color and done 30 times about two colors. The data is preserved by the text-base as follows:

Single Color = R, G, B and Impression number.

$2Colors = R_1, G_1, B_1, R_2, G_2, B_2$ *and   Impression number*

In this paper, this investigation was done to 43 subjects. Figure 2 shows the ex-periment screen of this questionnaire.

3.2   Impression Words

The impression received from the color is various. In this paper, an adjective (impres-sion word) that expresses the color impression is greatly classified into seven impres-sions referring to reference1. and 2.
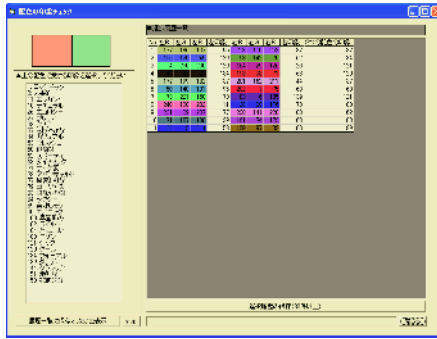
Fig. 2. The experiment screen

Table 1. The impression word to classify

Impression number	Representative's impression words
1	———— Clear
2	———— Natual
3	———— Casual
4	———— Dynamic
5	———— Elegant
6	———— Wild
7	———— Classic

4   Addition of the Color Feature

In this paper, the color feature was added as the former stage of the feature extraction by the principal component analysis. The collected data is preserved by RGB-value. However, it is difficult to extract the feature only by the RGB-value because the correlation among three attributes of R, G and B is strong. Therefore, the HSI transform is adopted in addition to a general color specification system of RGB. This is expressed the color by three elements of hue (H), saturation(S) and intensity (I). In this paper, we convert from RGB to HSI by using the HSI 6 pyramid color model shown in Figure 3, and add it to the color feature.

5   Creation of the Data Matrix

After the color feature is added to the collected data, the data matrix is created. In case of single color, the color feature is expressed by RGB-value, hue, saturation and intensity *(1)*. When thinking about the arrangement of colors, it has an important meaning to relate between two colors. Then, in this paper, the difference of R-value, G-value, B-value, hue, saturation and intensity are added *(2)*. The color feature is as follows:
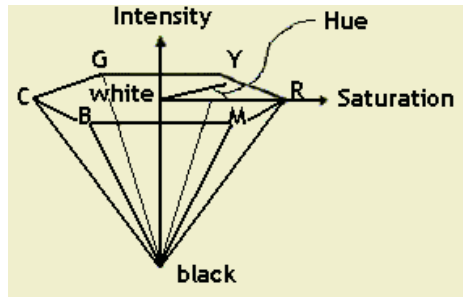
Fig. 3. The HSI 6 pyramid color model

- *Single Color (R, G, B, H, S, I).*   … (1)
- *Two Colors =*

$$(\quad R_1, G_1, B_1, H_1, S_1, I_1,$$
$$R_2, G_2, B_2, H_2, S_2, I_2,$$
$$R_1 - R_2, G_1 - G_2, B_1 - B_2,$$
$$H_1 - H_2, S_1 - S_2, I_1 - I_2 \quad) \quad ...(2)$$

We create the data matrix based on the color feature.

6   Principal Component Analysis

The principal component analysis is a method for mutually consolidating it as small number of overall indices of piece to be mutually independent without the loss of information as much as possible when the measurement variable is many or equivocal. This method is used for the feature extraction from colors. In this paper, the principal component analysis is done by using the data matrix, and the feature vector is extracted. The number of the major ingredients used in consideration of the contribution rate is decided; the major ingredients that the rate of accumulated contribution is more than 70% are used in this paper.

7   Results

Figure 4 show the distribution of feature vectors created by the principal component analysis in single color. According to this figure, it is thought that the feature extraction was performed by the principal component analysis. From these results, it is confirmed that the proposed method work well.

8   Conclusions

In this paper, the color features were extracted by principal component analysis. As the result, it was suggested that this proposal method is valid, and the feature extraction from the color by the principal component analysis was effective. Henceforth, on the basis of this result, we will research about the method of classifying impressions automatically.
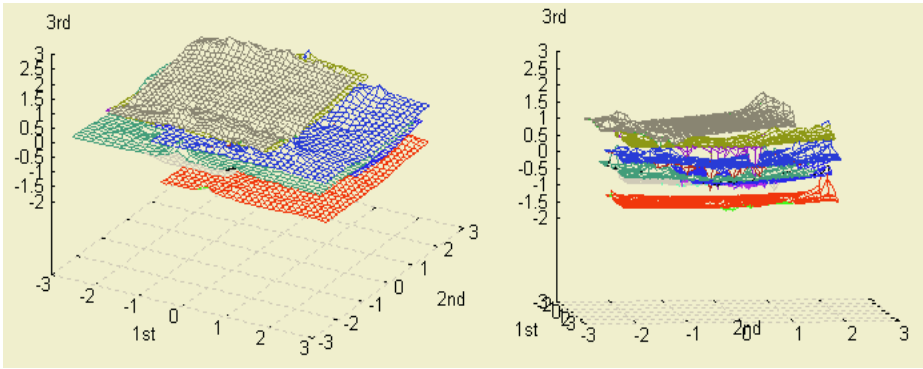
Fig. 4. the distribution of the 1st to 3rd component score in single color

References

1. S.Tsuji: "Science of the sensibility – Approach to the KANSEI information processing", Science Publishers Ltd., 1997 (in Japanese)
2. S.Kobayashi: "Color Image Scale", Kodansha Publishers Ltd., 1995 (in Japanese)
3. H.Nagumo: "Color Image Chart", Graphic Publishers Ltd., 1999 (in Japanese)
4. Japanese color society: "Handbook of color science- second edition-", The University of Tokyo publication association, 1998 (in Japanese)
5. M.Takagi, H.Shimoda: "Handbook of image analysis", The University of Tokyo publication association, 1991 (in Japanese)
6. M.Tokumaru and K, Yamashita: "Color Coodinate Evaluationg System Using Fuzzy Reasoning", IEICE, Vol.J83-D-ll, No.2, pp.680-689, 2000 (in Japanese)
7. T.Yagyu, Y.Hisamori, Y.Yagi, and M.Yachida: "Color Coodinate Supporting System with Navigating State of User's Mind", IEICE, Vol.J79-A, No.2, pp.261-270, 1996 (in Japanese)

Research on Individual Recognition System with Writing Pressure Based on Customized Neuro-template with Gaussian Function

Lina Mi[1] and Fumiaki Takeda[2]

[1] Department of Intelligent Mechanical System, Kochi University of Technology
185 Miyanokuchi, Tosayamada-cyo, Kami-gun, Kochi, 782-8502, Japan
086408w@gs.kochi-tech.ac.jp
[2] Department of Intelligent Mechanical System, Kochi University of Technology
185 Miyanokuchi, Tosayamada-cyo, Kami-gun, Kochi, Japan
Takeda.fumiaki@kochi-tech.ac.jp

Abstract. In our previous research, neuro-template matching method[1] was proposed for currency recognition. In this paper, neuro-template with sigmoid as activation function is applied in the individual recognition system with writing pressure, and the experiment shows that this method is effective on the known pattern recognition, however it suffers from poor rejection capability for counterfeit signatures. To solve previous problem, Gaussian function is proposed as activation function of neuro-template and optimal parameters are customized for neuro-template of each registrant. The experiment shows that the customized neuro-template with Gaussian activation function is seemed to be very effective on improving the rejection capability of the system for counterfeit signatures with ensuring the recognition capability satisfied.

1 Introduction

Recently, the application of biometrics information such as fingerprint, iris has been increasingly applied into individual recognition system.

Being one of biometrics information, writing pressure, detected dynamically from signature procedure, is employed as personal information in our individual recognition system taking advantage of unique individual character in terms of signature force rhythm. Besides being friendly and has little spirit resistance, writing pressure has higher security than handwriting because of being dynamic signal and invisible to users. The application of our system has shown that taking writing pressure as personal information for individual recognition is feasible and reliable.

In our system, neuro-template with sigmoid as activation function is used to recognize the user logging on the system. Though this system is effective on recognizing the known registrants, the rejection capability for counterfeit writing made by imitating the legitimate registrant's signature by intruders is not satisfying. To solve this problem, Gaussian function is proposed as activation function of the neuro-template and furthermore optimal parameters of Gaussian function are customized for each registrant's neuro-template in this paper. The experiment results indicate that the

Gaussian function with customized parameter proposed in this paper is seemed to have high effectiveness on improving the counterfeit rejection performance of the system with ensuring recognition capability for known pattern satisfied.

2   Basic Structure of the Individual Recognition System

The system described in this paper is composed of hardware and software sections. The hardware section which construction is illustrated as Figure 1 includes the electronic pen, data collection box and personal computer. Writing pressure data are detected by the electronic pen, and then loaded to PC via the data collection box.

Registration based on Neuro-template

Recognition based on Neuro-template

Data Collection Box

Electronic Pen
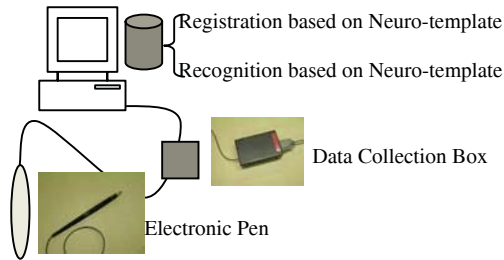
Fig. 1. Construction of the individual recognition system with writing pressure

Software section includes registration part and recognition one which are independent each other and both of them are based on neuro-template which will be interpreted in section four. Registration subsystem is in charge of recruiting new member and recognition subsystem recognizing the identity of the user entering the system.

3   Preprocessing

The writing pressure data from hardware section must be preprocessed because of too much redundant and noise data. Three stages are included in data preprocessing. First, the data from electronic pen are transformed into normal data, in this stage the individual feature is extracted and data scale is also greatly compressed. Second, validity check is conducted on normal data obtained previously. Last, normal data are transformed into slab data for neuro-template by filtering normal data with the mask data.

4   Algorithm

For traditional of multiplayer feedforward neural network (MFNN), recruitment of new pattern inevitably leads to restructuring and retraining on the whole network, that would contribute to expensive computation and time cost. Further more, with increase of patterns in the output layer, the network will become too complicated to work well. To address previous problem, neural template matching method, the strategy of

which is shown as Figure 2, is introduced into our system. In this method, one neuro-template corresponds to one registrant and each neuro-template is constructed by three layers feedforward neural network with fixed structure of 50×35×2 as shown in Figure 3. In new registration new one neuro-template is constructed and trained for the new registrant, other existing neuro-templates will remain untouched unless mis-outputs are caused by new registration. Namely instead of training all templates, only the training for new template and the retraining for the existing templates in which the mis-recognition occurred are involved in the new registration procedure. Therefore registration procedure is significantly simplified and the limitation on number of registrants is also removed because of relative independence of each neuro-template.
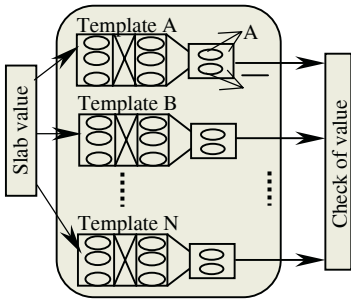
Fig. 2. Neuro-template matching method
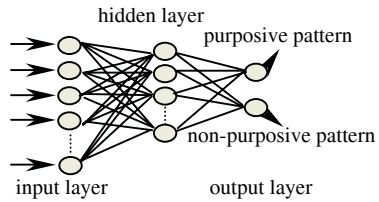
Fig. 3. Structure of neuro-template

Although the original system with sigmoid activation function is effective on classifying the known signatures, it suffers from poor rejection capability for counterfeit signature. That mostly due to the limitation of sigmoid function. Taking two-dimensional space as example for convenience, the recognition space of neuro-template with sigmoid function is divided into two open sub-domains shown as figure 4, purposive pattern and non-purposive one. The network is effective on recognizing registrants whose pattern has been learned, but for counterfeit writing which is neither purposive pattern nor non-purposive one (namely unknown pattern), it will also be classified as one of them, that contributes to poor rejection capability for counterfeit writing. Unlike the one with sigmoid function, neuro-template with Gaussian function, the recognition space of which is demonstrated as Figure 5, learns the pattern probability density instead of dividing up the recognition space.
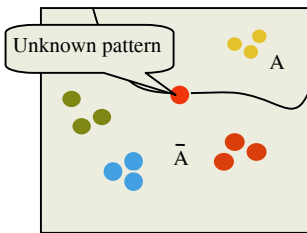
Fig. 4. Recognition space of neuro-template with sigmoid function
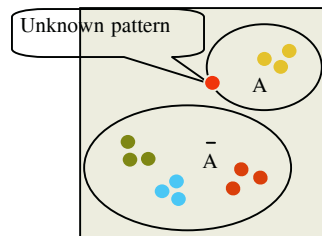
Fig. 5. Recognition space of neuro-template with Gaussian function

In the recognition space as figure 5, when a pattern which is out of known categories is evaluated, it is likely to be recognized as unknown pattern. Therefore for counterfeit writing, there is high probability of being rejected as unknown pattern. Based on the previous analysis, Gaussian function is proposed as activation function of neuro-template to improve the rejection capability of the system for counterfeit writing in this paper. Its expression is described as following equation:

$$f(x) = \exp\left(-\frac{1}{2\sigma^2}\|x\|^2\right) \tag{1}$$

Where x is input vector, σ is the width parameter of Gaussian function.

In the neuro-template training procedure, improved Back Propagation (BP) is employed to modify the weights between neighbor layer neurons and corresponding expression is described as following:

$$\Delta W(t) = -\eta \cdot \frac{\partial E}{\partial W(t)} + \alpha \cdot \Delta W(t-1) + \beta \cdot \Delta W(t-2) \tag{2}$$

Where $W(t)$ is weight vector of hidden layer or output layer at t th iteration, and $\Delta W(t)$, $\Delta W(t-1)$ and $\Delta W(t-2)$ are variations of weight vector at t th, $t-1$ th and $t-2$ th iteration respectively. η, α and β are the learning rate, momentum coefficient and oscillation coefficient respectively. The momentum item is used to accelerate convergence and the oscillation item to escape local minimum.

In the earlier research, width parameter σ is selected manually and kept constant once proper value is decided. In this paper, further study is conducted and optimal value of σ is automatically customized for each neural units in middle layer and output layer of neuro-template. With proper initial value, parameter σ is modified basing on same improved Back Propagation algorithm as described immediately before:

$$\Delta\sigma(t) = -\eta \cdot \frac{\partial E}{\partial\sigma(t)} + \alpha \cdot \Delta\sigma(t-1) + \beta \cdot \Delta\sigma(t-2) \tag{3}$$

Where η, α, β have same meaning as that in equation (2), $\Delta\sigma(t)$, $\Delta\sigma(t-1)$ and $\Delta\sigma(t-2)$ are the variations of width at iteration t, $t-1$ and $t-2$ respectively. Though the convergence speed of $\Delta\sigma(t)$ varies with different neuro-template, the convergence performance is excellent in general, and two hundred iterations are needed at most.

5   Experiments

In this section, simulation experiments are made to verify the effectiveness of proposed method. Three registrants (labeled as A, B, C) are selected randomly as purposive registrants , 53 purposive signature samples are extracted from each registrant, any three of these samples are used for neuro-template learning and the left for evaluation. 50 samples are extracted as counterfeit signatures for each purposive registrant from any other five people (except A, B, C) with 10 samples per person.

The performance of the system involves recognition capability and rejection capability for counterfeit signatures. Assuming registrants A, B and C included in the system, taking registrant A as purposive registrant for example, signature (purposive signatures of A or counterfeit signatures of A) is fed to each template and evaluated by each template. All possible cases of the system are illustrated in table 1. Note that output of the system will be one of these three registrants or none of them (briefly N).

Table 1. All possible situation of experiments

Signature	Source of signature	Ideal results	Success	Failure	
Purposive	Register A	A	A	B or C	N
Counterfeit	Anyone except A, B and C	N	N	A	B or C

In the previous table, the failure cases in which purposive signature of A is mis-recognized as B or C and the counterfeit signature for A is mis-recognized as B or C are not in consideration because they can easily be overcome by our system for significant difference among the signatures of the respective registrant.

Basing on the previous condition and the sampled data (purposive and counterfeit), neuro-templates with Gaussian Function and that with sigmoid function are trained with the sequence of A, B, C and evaluated respectively, and the experiment results are presented in Table 2 and Table 3.

Table 2. Recognition ratio for purposive signatures (ideal value: 100%)

Activation function	A	B	C	Average
Sigmoid	96%	98%	100%	98%
Gaussian	96%	96%	100%	97.33%
Difference	-0%	-2%	-0%	-0.67%

Table 3. Rejection ratio for counterfeit signatures (ideal value: 100%)

Activation function	A	B	C	Average
Sigmoid	64%	86%	74%	74.67%
Gaussian	82%	90%	84 %	85.33%
Difference	+18%	+4%	+10%	+10.67%

The minus sign (-) here indicates deterioration of the performance of system and the positive one (+) means improvement. According to table 2 and table 3, though the recognition capability of system with Gaussian function decreased slightly (average -0.67%), its rejection capability for counterfeit signatures was improved greatly with average value of 10.67%. This result suggests that the neuro-template with Gaussian function proposed in this paper tends to be effective on improving the rejection capability of system for the counterfeit signatures and at the same time keeping the recognition capability for purposive signatures satisfied.

As mentioned in section four, the recruitment of new registrant may lead to retraining on the exiting templates. That indicates neuro-templates are likely to be influenced by other templates registered later. To investigate the mutual influence of different neuro-templates, experiments are conducted on the systems with different number and sequence of templates. Registrant A and B in previous experiments are selected for convenience. Note that any other template except A and B is not included in the system. Table 4 and 5 show the performances of template A and B under different situation respectively, As can be seen from table 4 and 5, the recognition performance of template A and B are not influenced by register sequence, it indicates that the recognition performance of the system is not affected by increase of templates. According to table 4, with different activation function, the rejection capabili-

ties of template A are enhanced by new register of B and this improvement is especially remarkable with Gaussian function employed. From table 5, the rejection performance of template B under different situation keeps untouched. These experiment results show that recruitment of new templates is sometime helpful to improve the rejection capability of neuro-template for counterfeit signatures though it doesn't always do so.

Table 4. Performance of template A with different function in different situation

Function	Performance (ideal value: 100%)	A only	A, B	B, A
Sigmoid	Recognition	96%	96%	96%
	Rejection	58%	64%	58 %
Gaussian	Recognition	96%	96%	96%
	Rejection	66%	82%	66 %

Table 5. Performance of template B with different function in different situation

Function	Performance (ideal value: 100%)	B only	B, A	A, B
Sigmoid	Recognition	98%	98%	98%
	Rejection	86%	86%	86 %
Gaussian	Recognition	96%	96%	96%
	Rejection	90%	90%	90%

As mentioned before, the optimization of σ for each registrant's neuro-template is conducted for further improvement on our system. The performance of the system with customized σ for each neuro-template are shown in table 6 and 7 respectively. Note that to demonstrate the effect of optimization of σ clearly, only one registrant is involved in the system to remove the influence among templates in each case,

Table 6. Recognition performance of template with different activation function

Function	A only	B only	C only	Average
Sigmoid	96%	98%	100%	98 %
Gaussian with fixed σ	96%	96%	100 %	97.33%
Gaussian with optimized σ	96%	96%	100%	97.33%
Difference	0%	-2%	0%	-0.67%

Table 7. Rejection performance of template with different activation function

Function	A only	B only	C only	Average
Sigmoid	58%	86%	74%	72.67%
Gaussian with fixed σ	66%	90%	84%	80%
Gaussian with optimized σ	82%	92%	90%	88%
Difference	+24%	+6%	+16%	+15.33%

The performance of system with sigmoid function and system with Gaussian function with fixed σ are also included in table 6 and 7 for the purpose of comparison.

The difference item in table 6 and table 7 describes the performance change of the system based on Gaussian with optimized σ comparing with the system with sigmoid activation function. As shown in table 6 and 7, even without the favourable influence of templates, Gaussian function with optimized σ for each neuro-template provides significant improvement on the counterfeit rejection capability of the system without sacrificing the recognition capability compared with Gaussian function with fixed σ.

6   Conclusion

In this paper, the construction and algorithm of the individual recognition system with writing pressure are firstly described. Then, Gaussian function, which is one of RBF, is proposed as activation function of neuro-template to improve the rejection capability of the system for counterfeit signatures. Furthermore the influence among neuro-templates is investigated in this paper. At last, the width parameter σ of Gaussian function is optimized based on improved BP algorithm for further improvement on the rejection capability of the system. The experiments results suggest that the influence among templates is favorable for rejection capability of the system, and more importantly the experiment results show that neuro-template based on Gaussian function with customized width parameter σ is seemed to be very effective on improving rejection performance of the system for counterfeit signatures at same time ensuring the recognition performance satisfied.

References

1. Takeda, F. and Omatu, S., 1995. High speed paper currency recognition by neural networks. IEEE Transaction on Neural Networks, Vol.6, No.1,pp.73-77.
2. Takeda, F. and Omatu, S., 1996. A neuro-system technology for bank note recognition, Proceedings of the Japan/USA Symposium on Flexible Automation, Boston,USA, Vol.2, pp.1511-1516.
3. Takeda, F. and Omatu S., 1998, Neural network systems Technique and applications in paper currency recognition neural networks systems, Technique and Applications, Vo1.5, Chapter 4, pp.133-160, ACADEMIC Press.
4. Takeda, F. Omatu, S., and Nishikege, 1996, A neuro-recognition technology for paper currency using optimized masks by GA and its hardware, Proceedings of the International Conference on Information System Analysis and Synthesis, Orlando, USA, pp147-152.
5. Takeda, F., Nishikage, T. and Matsumoto, 1998, Y., Characteristic extraction of paper currency using symmetrical masks optimized by GA and neuro-recognition of multi-national paper currency, Proceedings of IEEE World Congress on Computational Intelligence, Alaska, USA, Vol.1, pp.634-639.
6. Fnakubo, M., 1991, Pattern Recognition, Japan: Kyouritsu Press,pp15-25.
7. Abhijit S. Pandya, and Robert B. Macy, 1996, Pattern Recognition with Neural Networks in C++, Chapter 5, pp147-151, IEEE Press.

Context-Aware Evolvable System Framework for Environment Identifying Systems

Phill Kyu Rhee, Mi Young Nam, and In Ja Jeon

Dept. of Computer Science & Engineering
Inha University, Yong-Hyun Dong
Incheon, South Korea
pkrhee@inha.ac.kr, {rera,juninja}@im.inha.ac.kr

Abstract. This paper proposes a novel framework for adaptive and intelligent systems that can be used under dynamic and uneven environments by taking advantage of environment context identification. Adaptation to dynamically changing environments is very important since advanced applications become pervasive and ubiquitous. The proposed framework, callesd CAES (Context-Aware Evolvable System), adopts the concept of context-aware and the evolutionary computing, and the system working environments are learned (clustered) and identified as environmental contexts. The context-awareness has been carried out by unsupervised learning, Fuzzy ART. Genetic algorithm (GA) is used to explore the most effective action configuration for each identified context. The knowledge of the individual context and its associated chromosomes representing optimal action configurations is accumulated and stored in the context knowledge base. Once the context knowledge is constructed, the system can adapt to varying environment in real-time. The framework of CAES has been tested in the area of intelligent vision application where most approaches show vulnerability under dynamically changing environments. The superiority of the proposed scheme is shown using three face image data sets: Inha, FERET, and Yale.

1 Introduction

Recently, adaptation capability under dynamically changing environments becomes more important since advanced applications need service-oriented, pervasive, and ubiquitous computing [1, 2] In this paper, we discuss about the framework of CAES (Context-Aware Evolvable System) that can behave in an adaptive and robust manner under dynamic variations of application environments. CAES can be thought as an augmented evolvable system framework which provides not only the capability of evolution/adaptation but also that of context-awareness and context knowledge accumulation. The context knowledge of individual environmental contexts and their associated chromosomes is stored in the context knowledge base. The most effective action configuration of the system is determined for a current environment by employing the popular evolutionary computing method, genetic algorithm (GA). GA explores a most effective chromosome that describes the action configuration for each identified environmental context. The context-awareness consists of context modeling

and identification. Context modeling can be performed by an unsupervised learning method[17] such as Fuzzy Art, SOM, etc. Context identification can be implemented by a normal classification method such as NN, k-nn, etc.

Evolutionary computing is an efficient search and adaptation method by simulating the natural evolution mechanism [1]. GA guides the system adaptive to varying environment. Adaptation to changing environments is an important application of GAs [3, 4, 5]. GAs have been developed and used in many application areas such as optimization and search problems. GA is inspired by biological evolution process [3, 4]. GA is a well known searching algorithm in a huge search space via population. However, GAs can hardly be used under dynamically changing environments alone since they usually consume too much time to evaluate the population of chromosomes in the genospace until finding an optimal solution.

The proposed framework fuses the context-awareness and the evolutionary strategy. The novel strategy make the system adapt itself dynamically under changing and uneven application environments. Similar research can be found in [6,7,8,9,10].

The main difference of the proposed method from other evolutionary computing methods is that it can optimize action configuration in accordance with an identified context as well as stores its knowledge in the context knowledge. Hence, the proposed method can provide self-growing adaptation capability to the system. That is, once the context knowledge is constructed, the system can react to changing environments in real-time. Ubiquitous sensing and recognition under dynamic environment enforce visual sensor based information processing more adaptive to application environments. We show the feasibility of CAES framework in the area of vision applications suitable for dynamically changing applications such as ubiquitous, pervasive, sand service oriented systems. The proposed method has been tested for the popular vision application, face recognition using three data sets: Inha, FERET, and Yale where face images are exposed to dynamically changing and uneven lighting conditions. We achieve encouraging experimental results showing that the performance of the proposed method is superior to those of most popular methods under uneven environment. This paper is organized as follows. In the section II, we present the model of CAES. In the section III, we discuss about design issues such as context learning, context identification, run-time adaptation using genetic algorithm. Finally, we give experimental results and concluding remarks in the section IV and V, respectively.

2   Model of Context-Aware System

In this session, we discuss about the model of CAES with the capability of real-time adaptation and context knowledge accumulation. The goal of CAES can be described as the provision of optimal services (performances) using given system resources by identifying its working environments (contexts), and the major functionalities of CAES can be formalized as follows (see Fig. Fig. 1):

– Identify the system working environment (Context-awareness),
– Configure its structure using autonomous computing method (Evolvable action configuration), and Accumulate knowledge from its experience in a manner of self-growing (Context knowledge accumulation).
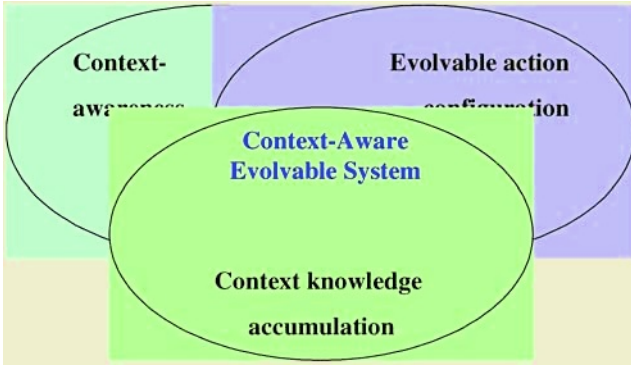
Fig. 1. Major functionality of CAES

Two types of data inputs are used: context data and action data inputs in CAES. The action data, denoted by $\mathbf{x}$, is a normal data being processed. The context data, denoted by $\mathbf{y}$, is used to identify an environmental context of CAES and to construct a proper action (action configuration) based on the identified context. In many cases, the action data itself can be used as the context data. We assume that the context data can be modeled in association with the input action data.

The system action or response can be described as the sequence of actions of action stages. Each action stage has its own action primitives. The action of each action stage is configured by selected action primitives or combined action primitives. The model of action configuration can be formalized as follows. Assume that the system has three action stages. Let $STG_1 = \{ap_{11}, ap_{12}, \ldots, ap_{1s}\}$, $STG_1 = \{ap_{21}, ap_{22}, \ldots, ap_{2t}\}$, $STG_1 = \{ap_{31}, ap_{32}, \ldots, ap_{3u}\}$ be a set of action primitives in the action stage 1, 2, and 3, respectively. The set of all possible action configurations can be expressed as follows:

$$AC = \{STG_1 \times STG_2 \times STG_3\} \tag{1}$$

Total k (= s x t x u) action configurations can be generated. An example of action configuration can be denoted as follows.

$$AC_i = [ap_{1a}, ap_{2b}, ap_{3c}] \tag{2}$$

where $i \leq k$, a = 1, 2, . . ., s, b = 1, 2, . . ., t, and c = 1, 2, ,, , , u. Then, all possible action configurations can be described as,

$$AC = \{AC_1, AC_2, \ldots, AC_k\} \tag{3}$$

Let $\mathbf{R}^n$ be an input attribute space, and $RES = \{res_1, res_2, \ldots, res_d\}$ be a set of response labels. The input attribute space is assumed to be represented by feature vector $\mathbf{x}$ in $\mathbf{R}^n$, i.e. $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$. Individual action configuration assigns response label of RES to an input attribute vector $\mathbf{x}$. That is, $AC_i : \mathbf{R}^n \rightarrow RES$. The output of action configuration AC_i can be represented as a vector in the following.

$$AC_i(\mathbf{x}) = [O_{i,1}(\mathbf{x}), O_{i,2}(\mathbf{x}), \ldots, O_{i,d}(\mathbf{x})] \tag{4}$$

where $O_{i,j}(x)$ is the probability of *jth* output of action configuration AC_i using the input attribute vector **x**. The context-aware action configuration needs identify an environmental context first, and only a small subset of total k action configurations are selected based on the identified environmental context and aggregation is performed as follows.

$$-AC\ (\mathbf{x}, \mathbf{y}) = AAO\ \{CAO(\mathbf{y})\ \{AC_1, AC_2, \ldots, AC_k\}\} \tag{5}$$

where CA-AC represents a context-aware action configuration, *AAO* denotes action aggregation operator, and *CAO* is a context-aware operator. AC_i is an i-th action configuration. If *CAO* decides more than one action configurations, *AAO* aggregates action configuration, e.g., majority voting. The context knowledge for *AAO* is accumulated by some learning method and stored in the CKB (Context Knowledge Base) (see Fig.2). In general, the context knowledge can be expressed by the ontology reflecting application environments. For the simplicity of explanation, we adopt the most simple form, the lookup table that consists of environment contexts and corresponding action configurations in this formulation. However, it can be readily extended more complex ontology description. The details of implementation issues will be discussed in the next session.

The output of context-aware operator can be represented using a context action profile (CAP) as shown in the following equation where context ontology is represented by the lookup table.

$$CAP(\mathbf{x},\mathbf{y}) = CP^t(\mathbf{y})\ x\ AC(\mathbf{x})$$
$$= [CP_1(\mathbf{y}), CP_2(\mathbf{y}), \ldots, CP_k(\mathbf{y})]^t\ x\ [AC_1(\mathbf{x}), AC_2(\mathbf{x}), \ldots, AC_k(\mathbf{x})] \tag{6}$$

The element of $CP_i(\mathbf{y})$ indicates whether the corresponding classifier CLS_i is activated or not under a given context y. That is,

> If $CP_i(\mathbf{y}) = 1$, then AC_i is activated,
> Else if $CP_i(\mathbf{y}) = 0$, AC_i is deactivated.

The intermediate action space considering an identified context is organized by the activated elements in AC(x). The activated elements in AC(x) are decided by applying context-aware operator *CAO*. *CAO* is denoted as the activating vector **CP**, called context profile. The system output is represented as a context decision profile as shown in the following equation.

$$|AC_1(\mathbf{x})| = |O_{1,1}(\mathbf{x}), O_{1,2}(\mathbf{x}), \ldots, O_{1,c}(\mathbf{x})| \tag{7}$$

$$CAP(\mathbf{x},\mathbf{y}) = CP^t(\mathbf{y})\ x\ |AC_i(\mathbf{x})|\ = |O_{i,1}(\mathbf{x}), O_{i,2}(\mathbf{x}), \ldots, O_{i,c}(\mathbf{x})| \tag{8}$$

$$|AC_k(\mathbf{x})| = |O_{k,1}(\mathbf{x}), O_{k,2}(\mathbf{x}), \ldots, O_{k,c}(\mathbf{x})| \tag{9}$$

where if $O_{i,i}(\mathbf{x})$ is zero, the output will be deactivated, i.e., filtered by *CAO*. The output of CA-AC can be described as follows.

$$\text{CA-AC}(\mathbf{x},\mathbf{y}) = AAO\,(\text{CDP}(\mathbf{x},\mathbf{y})) \tag{10}$$

where CAES($\mathbf{x},\mathbf{y}$) denotes the output of CAES using an action input $\mathbf{x}$ under a context input $\mathbf{y}$.

The CAES is the augmentation of CA-AC by employing the capability of context knowledge accumulation. The main difference of CAES from normal GA based adaptable systems is that it can select action configurations in accordance with an identified context or situation. Furthermore, CAES can store its accumulated knowledge or experience in the CKB Context Knowledge Base). Hence, CAES can describe incremental adaptation. In addition, once CKB is constructed, the scheme can react in real-time. It can be used in dynamically varying situation or environment in realtime while normal GA based approaches can hardly be used under such dynamic situations or a dynamic environments. In this way, the proposed classifier combination method can also include the model of dynamically classifier selection discussed in [11,12,13]. Furthermore, emerging computation paradigm such as ubiquitous, invisible, pervasive, etc. computing very often requires open environments and operational situation changes dynamically[11].

3   Framework of Context-Aware Evolvable System

In this section, we will discuss about the framework of CAES discussed in the section II. Most important design issues are as follows: i) how to learn and identify a context (category) in association with application working environments, and ii) how to decide (produce) action configuration for an identified context from all possible combinable action configurations. The selected r most effective action configurations for an identified context are combined to produce a final output of the scheme using some aggregation method. We need devise a method of calculating a context profile CP in order to derive context action profile CAP**.** Classifiers are encoded as artificial chromosomes, and the genetic algorithm is employed to discover a most effective subset of classifiers for the identified context. We assume that environmental data can be modeled as (clustered into) several discretized environmental contexts in association with distinguishable application actions. The proposed scheme operates in two modes: the learning and action modes. The context learning is performed by an unsupervised learning method in the learning mode, and the context identification is implemented by a normal classification method in the action mode. The knowledge of most effective subset of action configurations for identified contexts is accumulated and stored in the context knowledge base (CKB) with associated artificial chromosomes in the learning mode. The knowledge is used to select a most effective subset of classifiers for an identified environmental context, and the produced effective action configurations are aggregated to produce an optimal output of the scheme in the action mode.

Architecture of the CAES

In the proposed scheme, two types of inputs are used: context data and action data inputs as discussed in the session II. The action data is a normal input to be processed.

The context data is used to identify an environmental context by which the scheme adapts to a given working environment. In many applications, the same input data can be used as context and action data. The proposed scheme consists of the context-aware module (CAM), the context knowledge base (CKB), the evolutionary module (EM), the control module (CM), and the action module (AM). (see Fig. 2). The AM consists of action stages each of which has action primitives. The CAM performs the functions of learning and identifying context. The EM accumulates the context knowledge of optimal action configurations for each identified context. It is usually implemented by the genetic algorithm. The accumulated knowledge of action configurations and the corresponding identified context is stored in the CKB. During operation time, the CM searches for a most effective set of action configurations using each identified context and the previously accumulated knowledge in the CKB.
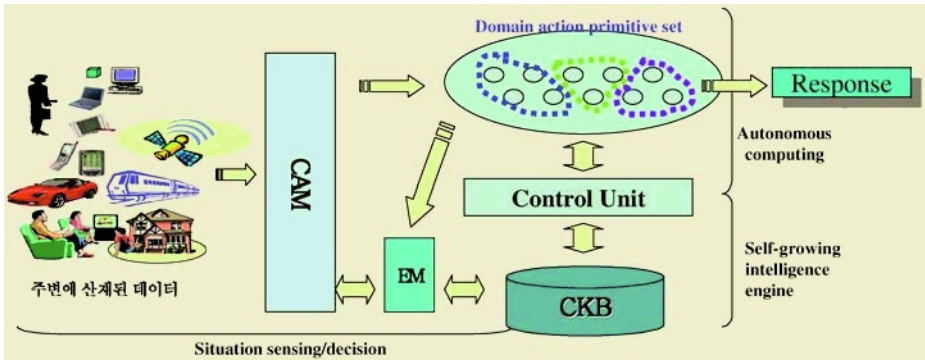
Fig. 2. An example of CAES design

Context can be various configurations, computing resource availability, dynamic task requirement, application condition, environmental condition, etc.[14]. Identified context describes a trigger an associated system action using a context stored in the context knowledge base over a period of time and/or the variation of a set of context data of the system over a period of time[15]. Context data is defined as any observable and relevant attributes, and its interaction with other entities and/or surrounding environment at an instance of time [16]. Context data can be denoted by a context tuple as follows.

$$CD = [t, cd_1, cd_2, . . ., CDs] \tag{11}$$

where t is the time stamp and cd_1, cd_2, . . ., cd_s are a set of context data attributes. Derived context features, which is preprocessed context data, is generated from the context tuple, and represented as follows.

$$DCF = [t, dcf_1, dcf_2, . . ., dcf_t] \tag{12}$$

where dcf_i is a preprocessed attribute from a subset of context data. Finally, identified context is expressed by attributes as follows.

$$IC = [t, ic_1, ic_2, \ldots, ic_u] \tag{13}$$

where t is the time when the context is identified and ic_1, ic_2, . . ., and ic_u are a set of identified context attributes representing an context. Identified contexts can be module configuration, parameter values, thresholds, parameter types, threshold types, etc. Context learning can be performed by an unsupervised learning method such as SOM, FuzzyArt, etc. Learned context can be identified by employing a normal classification method such as NN, k-nn, etc.

Action Module and Action Configuration

Action configuration is carried out using action primitive set and accumulated knowledge in the CKB. Action configuration can be described using the artificial chromosome of GA. The configured action produces a response or output of the scheme.
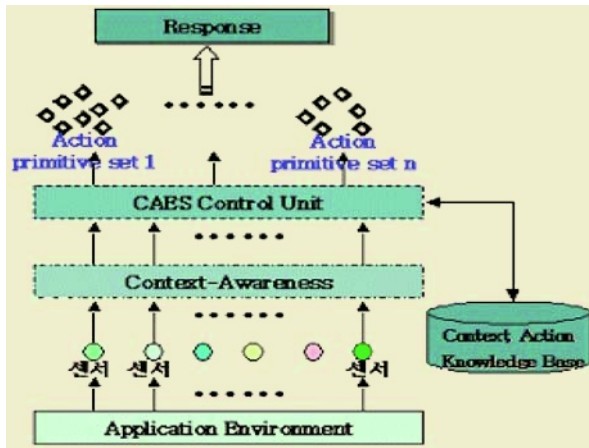
Fig. 3. Action module and action configuration

Initially, the scheme learns application's environmental contexts. It accumulates the knowledge of context-action association, and stores them in the CKB. The knowledge of context-action association denotes that of most effective action configurations for the context.

The details of context knowledge accumulation process is given in the followings.

Step1. The context data associated system working environments is clustered (learned) as environmental contexts (categories) by the CAM.

Step2. The action configuration structure of the AM is encoded as the chromosome, and the fitness of the GA is decided. (Each chromosome instance represents corresponding action configuration of the AM.)

Step3. The most effective subset of action configurations of AM for each identified environmental context is decided by the GA using the associated training action data as follows.
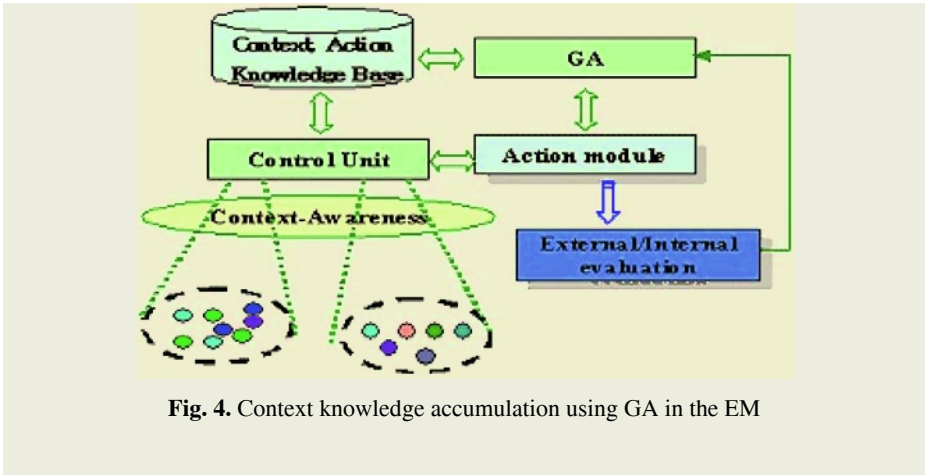
Fig. 4. Context knowledge accumulation using GA in the EM

     3.1 Generate a random population.

     3.2 Evaluate the fitness of each feature vector of the population.

     3.3 Select only a portion of the population as the population of next generation.

     3.4 Repeat 3.2 and 3.3 until a most effective action configuration is found, or a generation exceeds a desired criterion value..

     3.5 Repeat the above for all identified contexts.

Step4. The chromosomes with their associated contexts (the knowledge of context-action relations) are stored in the CKB.

A simple action scenario of the proposed scheme is outlined in the followings.

Step1. The environmental context is identified using inputted context data by the CAM.

Step2. Most effective subset of action configurations for the identified context is searched using the CKB.

Step3. The task of the scheme is performed by aggregating the output of the selected action configurations.

Step4. Whenever either the system identifies the non-trivial change of environment context or the system performance is measured to fall down below the predefined criterion, go to the step 1.

Step5. The CKB can be updated periodically. or whenever it is needed by activating the context-action knowledge accumulation process.

4 Design Examples

The proposed framework is tested in the area of face recognition and location. Extensive experiments have been carried out using three data sets: Inha, FERET, and Yale. We achieved very encouraging results.

Face Recognition Scheme Based on CAES Framework

The proposed CAES method has been tested in the area of face recognition. The changes in image data under changing illumination is modeled based on the illumina-

tion context. Illumination context may include lighting direction, brightness, contrast, and spectral composition. We adopt lighting direction and illumination as the model of illumination context here. The architecture of face recognition based on the CAES is given in Fig. 5.
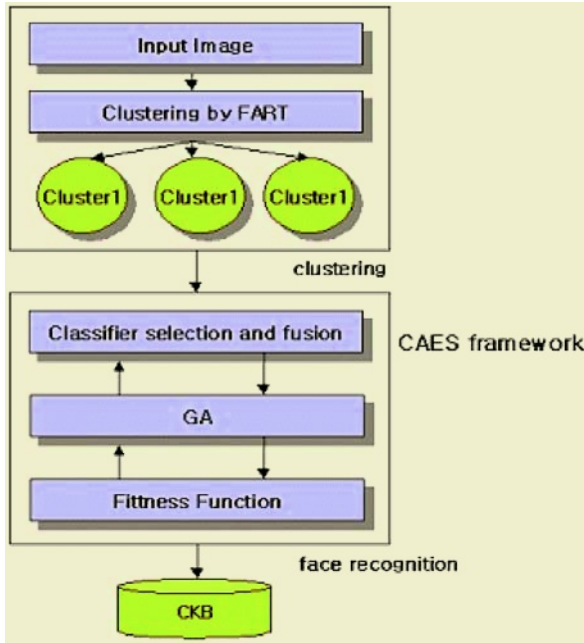
Fig. 5. Face recognition using CAES framework

In the experimental system for the face recognition, the CAM clusters face data images into five contexts, and constructs the context model of face images for each environmental context using the unsupervised learning algorithm, Fuzzy ART [17]. Fuzzy ART is a variant of ART system derived from the first generation of ART, namely ART1. It is a synthesis of ART algorithm and Fuzzy operator. ART1 can only accept binary input pattern, but Fuzzy ART allows both` binary and continuous input patterns [18]. The feature space of object instance with uneven illumination must be clustered properly so that the location error can be minimized. However, the clustering of illumination variations is very subjective and ambiguous. Thus, we adopt Fuzzy ART for achieving an optimal illumination clustering architecture. In this paper, the clustering performance is improved by iterative learning method. Fig. 6shows the clustering result of face images by Fuzzy ART.

The AM of the face recognition consists of three stages: preprocessing, feature representation, and class decision stages. The action primitives of the preprocessing stage are histogram equalization, Retnix, and end-in contrast stretching. The plausible combinations of action primitives in the preprocessing stage considered are Retnix + histogram equalization and histogram equalization. The action primitives in the feature representation stage are PCA and Gabor representations. Gabor representations are divided into four types. The proposed approach employs Gabor feature vector, which

is generated from the Gabor wavelet transform. We adopt four Gabor based feature representation: Gabor3, Gabor13, Gabor28, Gabor30. Thus, we use total five action primitives in the feature representation stage , PCA, Gabor3, Gabor13, Gabor 28, and Gabor30. We adopt two action primitives in the class decision stage: Euclidean distance based and cosine distance based actin primitives.

Fig. 6. Examples of face image clustered by the CAM

The evolutionary module EM explores optimal action configurations of the AM. In the recognition mode, the system searches for a most effective action configuration based on the identified category. The knowledge of effective action configuration and its environmental context is stored in the CKB. GA is employed to search among the different action configuration in each stages. Fig. 3 shows a possible encoding of chromosome description.

Fig. 7. The chromosome description of the CAES aiming at face recognition application, where HE, RX, and EC are histogram equalization, Retnix, and End-In Contrast, respectively. FR1, FR2, FR3, FR4, and FR5 denote PCA, Gabor3, Gabor18, Gabor28, and Gabor30 action primitives in the feature representation stage. CD1 and CD2 indicate Euclidean distance based and cosine distance based action primitives in the class decision stage.

Evolution or adaption will be guided by a fitness function defined in terms of the system accuracy and the class scattering criterion. The evolutionary module derives the classifier being balanced between successful recognition and generalization capability. The fitness function adopted here is defined as follows:

$$\eta(V) = \lambda_1 \eta_s(V) + \lambda_2 \eta_g(V) \tag{14}$$

where $\eta_s(V)$ is the term for the system correctness, i.e., successful recognition rate and $\eta_g(V)$ is the term for class generalization [19]. λ_1 and λ_2 are positive parameters that indicate the weight of each term, respectively.

The feasibility of the proposed method has been tested using Inha, FERET[20], and Yale [19] data set [18]. Experiments have been carried out to compare the recognition performance of the CAES based face recognition scheme and that of other methods. We used 1000 images of 100 persons from Inha data set, 60 images of 15 persons from Yale Face DB [19], and 2418 images of 1209 persons from FERET data set. The above data sets are merged for training and testing the CAM (see session 3). The first experiment is performed using Inha data set accumulated by our lab. We used 5 image for registration for each people. The remaining 500 images are used as the test images. For the Yale data set, we used 15 registration face image and 45 test images, The FERET images of 1209 people is used for registration and 1209 probe_fafb_expression images are used for test.As shown in Table 6, the face recognition scheme based on CAES (90.4%) outperforms the performance of the non-evolvable methods for the integrated data set. Even though Retnix based method (RX) shows the highest performance using FERET fafc data set, it can not be used under normal illumination. Histogram equalization based method (HE) shows the highest performance under normal illumination, i.e. FERET fafb data set, but it shows poor performance under bad illumination, i.e. FERET fafc data set. Thus, if the working environment of the system can be controlled well, non-evovable method may perform better than the proposed adaptive method. However, in general we can not predict or control the system working environment. Thus, we can say that the proposed method is better choice than the non-evolvable methods. Fig. 8 shows the CMC analysis of integrated data set comparing CAES based method to non-evolvable methods.

Table 1. Comparison of face recognition rates of CAES based method to those of non-evolvable methods focused on the preprocessing stage in the AM

Dataset & Methods	RX	HE	ES	RX + HE	HE + RX	CAES
FERET fafc dataset (bad illumination)	82.98 %	79.58%	78%	79.5%	76.48%	83%
Inha dataset	90.7%	96.98%	97.06%	97.30%	93.01%	98.2%
FERET fafb (normal illumination)	75%	94%	89%	79.5%	89.65%	95%
Yale DB	85%	91%	76%	82.3%	80.5%	93.5%
Integrated data set (fafc+fafb+our lab)	64.33%	85.38%	84.35%	85.43%	81.99%	90.4%

HE: the histogram equalization based method; EC: the end-Ii contrast based method; RX: Retnix based method; CAES: Context-Aware Evolvable System framework based method.

Table 2 shows the effect of feature representation stage of th AM between the CAES based method and other non-evolvable methods, Fig. 9 shows the CMC analysis of the integrated data set.

Table 3 shows the comparative test performance by Eigenface using Bayesian theory, linear discriminant, elastic graph matching, and evolutionary pursuit. The recognition rate of Eigenface is 83.4 % and Evolutionary Pursuit is 92.14 %. In own experimental results, the proposed method shows recognition rate of over 95.04 % for

FERET dataset, which exceeds the performance of the other popular methods. The context-based classifier fusion method performs better than sigle classifier method because the image feature is differnt according context information.
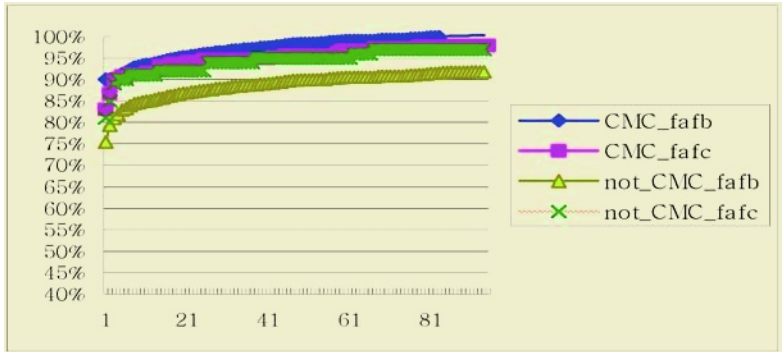
Fig. 8. CMC analysis of integrated data set for individual methods focusing on the effect of the preprocessing stage in the AM

Table 2. Comparison of face recognition rates of CAES based method to those of non-evolvable methods focusing on the feature representation stages in the AM

Method	CAES	PCA	Gabor3	Gabor13	Gaor28	Gabor30
Inha	98.5%	92%	90%	65%	93%	97%
FERRET fafb	95.04%	90%	25%	50.25%	90.25%	94%
FERRET fafc	81.94%	65.5%	20%	26%	65%	68%
Yale	97.5%	85%	10%	30%	85.25%	96.5%
Integrated	93.25%	83.13%	36.25%	42.81%	83.38%	88.88%

Table 3. Comparative Testing Performance : FERET database

Method	Rank1 correct acceptance
Eigenface[20]	83.4%
Eigenface by Bayesian[21]	94.8%
Evolutionary Pursuit[22]	92.14%
Proposed method	95.04%

5   Conluding Remarks

This paper proposes a novel adaptive framework called CAES that can be used for adaptive systems under uneven and dynamic environments by employing the concept of context-awareness, GA, and context knowledge. The system working environments are learned (clustered) and identified as environmental contexts, and GA is used to explore the most effective action configuration(s) for each identified context. The knowledge of the individual context and its associated chromosomes representing

optimal action configurations is accumulated and stored in the context knowledge base. After the context knowledge is constructed, the system can adapt to varying environment in real-time. The CAES framework separates the learning mode from the action mode to solve the time-consuming problem which is the intrinsic weak points of the GA. The learning mode accumulates the context knowledge from varying illumination environments, and the accumulated knowledge is stored in terms of artificial chromosomes. The action mode executes the task of identification using the chromosome knowledge accumulated in the learning mode. The major contribution of the proposed method is to achieve real-time evolvable system under varying environments. Furthermore, it solves the time-consuming problem of the GA by taking advantage of the concept of context-aware architecture. The superiority of the CAES based vision system is shown using three face image data sets: Inha, FERET, and Yale.

References

1. Ong KG, Dreschel WR, Grimes CA, "Detection of human respiration using square-wave modulated electromagnetic impulses," MICROWAVE AND OPTICAL TECHNOLOGY LETTERS, Vol.35 MAR 5 (2003) pp.339-343
2. Slay H, Thomas B, Vernik R, et al. "A rapidly adaptive collaborative ubiquitous computing environment to allow passive detection of marked objects," LECTURE NOTES IN COMPUTER SCIENCE ISSN:3101, (2004) pp. 420-430
3. J. H. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, (1975)
4. D. Goldberg, "Genetic Algorithm in Search, Optimization, and Machine Learning, Addison-Wesley, (1989)
5. N. Mori, et. al., "Adaptation to a Dynamic Environment by Means of the ENvironment Identifying Genetic Algorithm," Industrial Electronics Society, 2000. IECON 2000. 26th Annual Conference of the IEEE, Vol. 4, (2000) pp.2953 - 2958
6. N.Mori, H.Kita and Y. Nishikawa: Adapatation to a Changing Environment by Means of the Feedback Thermodynamical Genetic Algorithm, Proc. of 5th PPSN, Vol5, pp.149-158(1998)
7. K.Ohkura and K.Ueda: A Genetic Algorithm for Nonstationary Function Optimization Problems, Trans. SICE, Vol.8, No.6, (1995) pp.269-276
8. J.J.Grefenstette : Genetic algorithms for changing environments, Problem Solving from Nature 2, Vol.2, (1992) pp.137-144.
9. Kaenampornpan M, O'Neill E, "Modelling context: An activity theory approach," LECTURE NOTES IN COMPUTER SCIENCE ISSN:3295, (2004) pp.367-374.
10. Feng L, Apers PMG, Jonker W, "Towards context-aware data management for ambient intelligence," LECTURE NOTES IN COMPUTER SCIENCE ISSN:3180, (2004) pp.422-431.
11. L. I. Kuncheva, "Change-glasses approach in pattern recognition," *Pattern Recognit. Lett.*, vol. 14, (1993) pp. 619-623
12. L. A. Rastrigin and R. H. Erenstein, *Method of Collective Recognition* (in Russian). Moscow, Russia: Energoizdat, (1981)
13. K.Woods,W. P. Kegelmeyer, and K. Bowyer, "Combination of multiple classifiers using local accuracy estimates," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, (1997) pp. 405-410.

14. S. S. Yau, Y. Wang, D. Huang, and H. In, "A Middleware Situation-Aware Contract Specification Language for Ubiquitous Computing," *Proc. 9th International Workshop on Future Trends of Distributed Computing Systems (FTDCS2003)*, May, Puerto Rico, USA, (2003) pp. 93-99

15. S. Yau, Y. Wang, and F. Karim "Developing Situation-Awareness in Middleware for Ubicomp Environments," *Proc. 26th Int'l Computer Software and Applications Conference (COMPSAC 2002), (2002)* pp. 233-238

16. G.D. Abowd, "Classroom 2000: An Experiment with the Instrumentation of a Living Educational Environment," *IBM Systems. J.*, vol. 38, no. 4, (1999) pp. 508-530

17. Ramuhalli, P., Polikar, R., Udpa L., Udpa S. : Fuzzy ARTMAP network with evolutionary learning. Proc. of IEEE 25th Int. Conf. On Acoustics, Speech and Signal Processing (ICASSP 2000), Vol. 6. Istanbul, Turkey, (2000) pp.3466-3469

18. Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Addison Wesley, (1993) p161~218

19. M.YNam and P.K.Rhee, "An Efficient Face Recognition for Variant Illumination Condition," ISPACS2005, (2005)

20. M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cong. Neurosci. vol. 13, no.1,(1991) pp.71-86

21. B.Moghaddam, C.Nastar, and A,Pentland, "A Bayesian similarity Measure for direct Image Matching, " Proc. of Int. Conf. on Pattern Recogntion, (1996)

22. Chengjun Liu and Harry Wechsler, "Evolutionary Pursuit and Its Application to Face Recognition," IEEE Trans. on Pattern Analysis and Machin Intelligent, Vol.22, No.6, (2000) pp.570–582

Context-Aware Computing
Based Adaptable Heart Diseases Diagnosis Algorithm

Tae Seon Kim and Hyun-Dong Kim

School of Information, Communications and Electronics Engineering
Catholic University of Korea, Bucheon, Korea
{tkim,hdkim81}@catholic.ac.kr
http://isl.catholic.ac.kr

Abstract. It is well known that the electrocardiogram (ECG) signal has crucial information to detect heart disease. However, development of automated heart disease diagnosis system is known as nontrivial problem since the ECG signal is different from patient to patient, measured time and environmental conditions. To overcome this problem, context-aware computing based adaptable heart disease diagnosis algorithm is proposed. Before diagnosis step, patient signal type recognition module groups various types of signal characteristics for patients and genetic algorithm (GA) finds optimal set of preprocessing, feature extraction and classifier for each group of signal types. Evaluation results using MIT-BIH database showed that various types of signal type require different preprocessing, feature extraction method and classifier and the test results showed 98.36% of classification accuracy for best optimization case.

1 Introduction

The electrocardiogram (ECG) signal is a tracing of electrical activity signal generated by rhythmic contractions of the heart and it can be measured by electrodes placed on the body surface. The clinical experience data showed that ECG signal is very effective to detect heart disease and there are steady efforts of researchers to develop automated ECG signal classification algorithms. ECG signal analysis problem for heart disease diagnosis is mainly divided into two parts, feature extraction and classification. Figure 1 shows the typical shape of normal ECG signal. This signal shows only the tracing of electrical potential, therefore feature extraction procedure is required to detect abnormality of ECG signal and classification of disease types. As shown in Figure 1, ECG signal is generally divided into three different waves, P-wave, QRS complex, and T-wave which is the sequence of depolarization and repolarization of the atria and ventricles. Generally, the position information of P-wave, QRS complex, and T-wave, QRS duration, and the P-R, R-R, and Q-T intervals are used as the features of ECG signal. However, it is very difficult to extract features from ECG signal since the appearance of heartbeats varies considerably, not only among patients, but also movement, respiration and modifications in the electrical characteristics of the body [1]. Even, several electrical and mechanical noise components are also added to signal and it made difficult to extract key features. To overcome these difficulties, various wave detection algorithms have been developed through years. Ramli used cross-correlation analysis to extract features from abnor-

mal signal [2]. And also, maximum slope detection based detection method which called as "So & Chan's method" and digital analysis of the slope, amplitude, and the width of the data was developed by Pan and Tompkins showed improved signal detection performance [3]. However, still they can't extract some features for abnormal cases.
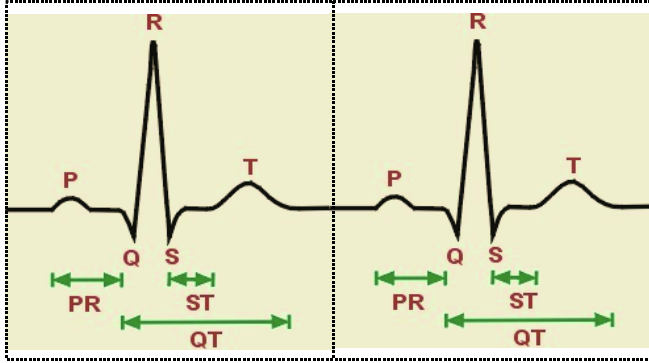
Fig. 1. Typical shape of normal ECG signal

Based on selected features, various pattern classification methods have been developed for heart disease diagnosis including cardiac arrhythmia classification using autoregressive modeling [4] and heart rate variability (HRV) analysis for cardiovascular disease [5]. However, most of classification algorithms suffer difficulty to find general rule base for automated diagnosis system because of wide variations of signal. Practically, it is nearly impossible to construct standardized diagnosis rule for heart disease since ECG signal is different from patient to patient. And also, it is varying on measured time and patient's condition even for same patient. For perfect classification, classifier needs to learn every case of ECG signals for every people, but it's impossible or not cost effective. This problem can be solved if we can group various types of signal characteristics (similarity of signal, gender, age, and etc.) for all patients before diagnosis step, we can develop more accurate patient adaptable ECG diagnosis algorithm. For this, patient type recognition method was developed [6] and genetic algorithm (GA) finds optimal set of preprocessing, feature extraction and classifier for each group of signal types.

2   Patient Adaptable Diagnosis Algorithm

For adaptable patient diagnosis, three channel ECG signals with environmental noise are used as context data. Proposed system is consist of mainly five modules including signal input module, output module for diagnosis results, patient signal type recognition module, diagnosis module and GA based diagnosis step optimization module as shown in Figure 2. The diagnosis module comprises preprocessing block for noise cancellation, feature extraction block and classifier block. Details on each block are described in below bur the detail algorithm and experimental result of patient type recognition module cannot be presented since it is currently patent pending.
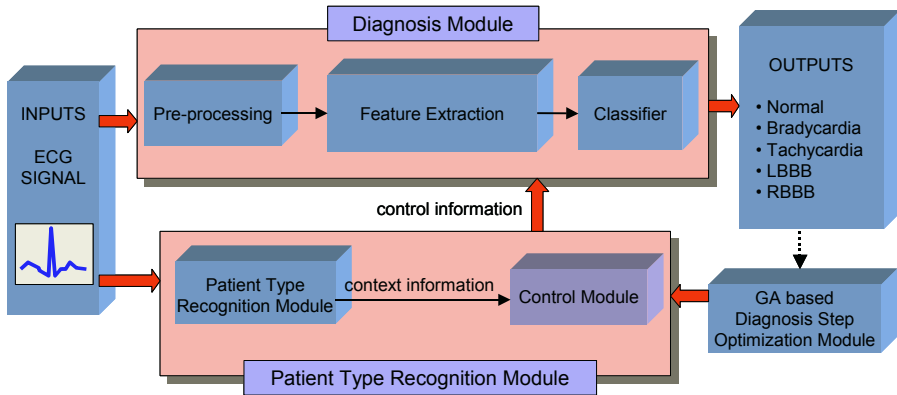
Fig. 2. Conceptual diagram of proposed adaptable diagnosis algorithm

2.1   Preprocessing Block for Noise Removal

Prior to feature extraction and classification steps, noise removal process is required. Generally, measured ECG signal data contains muscle noise, baseline noise, incoupling noise, and 60Hz power line noise. Generally, 1~50 Hz IIR band pass filter was widely used to remove noise factors. Also, high pass and low pass filters are used separately in many other algorithms [7]. In this algorithm, low pass filter with 25 Hz cut-off frequency (LPF), high pass filter with 0.5 Hz cut-off frequency (HPF), 0.5~25 Hz band pass filter (BPF), and 125 Hz, 30 Hz, 60 Hz notch filters (NF) are considered for optimal preprocessing methods. No filter block for preprocessing case is also considered as one of selection method since sometimes filters can make distortion or shape changes to original signal and it leads changes of important features including signal intervals and vertices.

2.2   Feature Extraction Block

As shown in Figure 1, normal ECG signal contains several characteristic waves, intervals and durations which is essential to diagnosis. Therefore, accurate feature extraction is essential to accurate classification. Among them, as shown in Figure 3, QRS complex duration, P-R segment, Q-T interval, R-R interval, and the amplitude of S wave from baseline were selected for classification. Although it is nontrivial problem to diagnosis heart disease, typical shapes of waves, QRS complex duration, and the typical values of interval between waves are generally known by clinical data analysis. In this work, two candidates are considered for feature extraction method, the So & Chan's method and Kim's feature extraction method. So & Chan's method is based on the maximum slope detection with the QRS onset selected when two successive values of the slope exceed the threshold. So & Chan's method showed superior QRS detection capability compare to others [3]. Kim's feature extraction method detects starting points of characteristic waves by finding peak point and specific cross point between signal and baseline [8]. Contrast to conventional feature extraction methods which use complex mathematical derivation, Kim's method requires relatively short computation time.
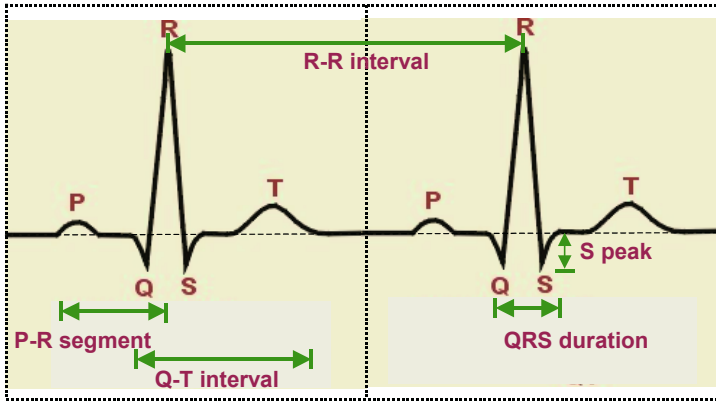
Fig. 3. Selected five features for classification

2.3 Classifier Block

For classification, three types of neural networks are selected for candidates, feed-forward error back-propagation neural network (BP), competitive neural network (CNN), probabilistic neural networks (PNN). In CNN, the net input of competitive layer is computed by finding the negative distance between input vector and the weight vectors. The competitive transfer function accepts a net input vector for a layer and returns neuron outputs of 0 for all neurons except for the winner neuron. In PNN, the first layer calculates Euclidian distances from the input vector to the training vector, and then calculates class probability based on radial basis function (RBF). Finally, the maximum of these probabilities are selected for output [9].

3   Experimental Results

For performance evaluation, clinical records of the MIT-BIH arrhythmia database are used. It was the first generally available set for evaluation and it is widely used in many research literatures [10]. All the data are sample at 360Hz. For training and test, total 14 data records (record # 100, #101, #103, #109, #111, #112, #113, #118, #200, #203, #205, #212, #214,and #232) are selected and seven signal types including four simple disease types, two complex disease types and normal signal are recorded. Considered disease types for classification are sinus bradycardia (SB), sinus tachycardia (ST), right bundle branch block (RBBB), left bundle branch block (LBBB), complex of RBBB and ST (ST+RBBB), and complex of LBBB and ST (ST+LBBB). For network inputs, selected five feature values are used consistently for every neural network classifiers, but outputs have different representation method because of nature difference of network architecture. For PNN and CNN, seven neurons are used for network outputs to represents each class. In contrast to winner-take-all structure, BP needs only four output neurons since it can represents complex disease types by firing multiple neurons simultaneously. For each signal type, sixty vectors are selected and totally 420 vectors are used for training. For test, 35 test vectors are selected at each signal types and therefore, totally 245 data are used for evaluation. For

patient type recognition, lead III based type classification method is applied [6]. To find optimized diagnosis module, simple GA is applied. In this case, searching step is not critical since the search space is only 30 cases (five preprocessing methods, two feature extraction methods and three classifier types) as a preliminary research level. However, with increase of modules, types and steps, GA based search method would be crucial to find optimal combination of diagnosis module.

To evaluate diagnosis performance, classification accuracy, uncertainty and four parameters (sensitivity (SN), specificity (SP), and positive predictive accuracy (PPA), negative predictive accuracy (NPA)) are calculated using equations in below,

$$SN = TP / (TP+FN) \times 100 \tag{1}$$
$$SP = TN / (TN+FP) \times 100 \tag{2}$$
$$PPA = TP / (TP+FP) \times 100 \tag{3}$$
$$NPA = TN / (TN+FN) \times 100 \tag{4}$$

Where TP, TN, FP, and FN represents true positive, true negative, false positive, and false negative, respectively and they are defined on Table 1.

Table 1. Definition of TP, TN, FP, and FN

	Assigned to class i	Not assigned to class i
Belongs to class i	TP	FN
Not belongs to class i	FP	TN

The sensitivity indicates the rate of the true positive events (TP) for each class and the specificity indicates the rate of the true negative events (TN) for each class [11]. Table 2 shows the performance differences among combination of different preprocessing, feature extraction method and classifier for evaluation data. In this example, combination of 'no filtering - Kim's feature extraction - BP' set is selected as the best combination. To show the effect of preprocessing and feature extraction method, after classifier was fixed to BP, diagnosis performance results are evaluated for various combination of preprocessing methods and feature extraction methods. In this experiment, some signals can't be classified for several reasons. For example, So & Chan's feature extraction method can't find QRS complex wave from original ECG signal, and it leads fail to extract some feature information from signal. To consider this factors for performance representation, uncertainty rate are calculated in table 2. Owing to lower feature extraction capability, So & Chan's method showed higher uncertainty compare to Kim's feature extraction method. As shown in this table, Kim's feature extraction method showed insignificant performance difference on preprocessing methods. For So & Chan's method, preprocessing using notch filter module showed about 12% of performance improvement compare to worst case (no preprocessing case).

In addition to dependency of patient types, experimental results also showed that diagnosis performances are very depends on disease types. For example, 'no filtering-Kim's feature extraction-CNN' set showed 94.28% of classification rate for SB and ST+RBBB, but this combination of diagnosis set showed only 37.14% of classification rate for ST. Therefore, it would be desirable to consider patient's medical history to develop more accurate and adaptable diagnosis system.

Table 2. Comparison of diagnosis performance results for various diagnosis modules

Preprocessing and feature extraction method	Classification rate (%)	Uncertainty (%)	SN (%)	SP (%)	PPA (%)	NPA (%)
N/A - Kim's	98.36	0	100	89.74	98	100
LPF - Kim's	87.34	0	98.9	53.23	86.19	94.29
HPF - Kim's	77.55	0	100	38.88	73.8	100
BPF - Kim's	96.32	0	100	79.5	95.71	100
NF - Kim's	96.73	0	100	81.4	96.2	100
N/A - So & Chan's	56.73	59.18	100	97.22	97.14	100
LPF - So & Chan's	57.14	59.18	100	100	100	100
HPF - So & Chan's	62.04	32.65	97.56	76.19	92.3	91.42
BPF - So & Chan's	66.53	32.66	100	94.59	98.46	100
NF - So & Chan's	68.97	28.57	100	85.36	95.71	100

4   Conclusion

In this paper, intelligent patient adaptable heart diseases diagnosis system is developed. To overcome nontrivial characteristics of ECG rule base, patient signal type recognition module groups various types of signal characteristics and genetic algorithm (GA) finds optimal set of preprocessing, feature extraction and classifier for each group of signal types. MIT-BIH database based evaluation results showed that various types of signal types require different preprocessing, feature extraction method and classifier and the test results showed 98.36% of classification accuracy for best optimization case. With successful development of proposed intelligent diagnosis algorithm, it can be applied to evolvable hardware platform to provide real-time reconfigurable diagnosis module for patient adaptable diagnosis.

References

1. S. Barro, M. Fernandez-Delgado, J. Vila-Sobrino, C. Regueiro, and E. Sanchez: Classifying Multichannel ECG Patterns with an Adaptive Neural Network. IEEE Engineering in Medicine and Biology (1998) 45–55
2. A. Ramli and P. Ahmad: Correlation Analysis for Abnormal ECG Signal Features Extraction. Proc. IEEE 4th National Conference on Telecommunication Technology (2003) 232-237
3. K. Tan, K. Chan, and K. Choi: Detection of the QRS Complex, P wave and T wave in Electrocardiogram. Proc. IEEE Advances in Medical Signal and Information Processing (2000) 41-47
4. Z. Zhang, H. Jiang, D. Ge, X. Xiang: Pattern Recognition of Cardiac Arrhythmias Using Scalar Autoregressive Modeling. Proc. World Congress on Intelligent Control and Automation, Vol. 6. (2004) 15-19

5. N. Ahuja, V. Raghavan, V. Lath, and S. Pillai: Heart rate variability and its clinical application for biofeedback. Proc. Computer-Based Medical Systems (2004) 263-266
6. T. Kim and C. Min: ECG Based Patient Recognition Model for Smart Healthcare Systems, Systems Modeling and Simulation: Theory and Applications. Lecture Notes in Artificial Intelligence, Vol. 3398. Springer, Berlin (2005) 159-166
7. B. Kohlerm, C. Hennig, R. Orglmeister: The Principles of Software QRS Detection. IEEE Engineering in Medicine and Biology (2002) 42–57
8. H. Kim, C. Min, and T. Kim: Feature Extraction of ECG Signal for Heart Disease Diagnosis, Proc. IEEK & KIEE Conf. Information and Control Systems (2004) 325-327
9. MATLAB 6, Users Manual, The Math Works (2001)
10. G. Moody and R. Mark: The Impact of the MIT-BIH Arrhythmia Database, IEEE Engineering in Medicine and Biology (2001) 45-50
11. R Silipo and C. Marchesi: Artificial Neural Networks for Automatic ECG Analysis, IEEE Trans. Signal Processing, Vol. 46, No. 5 (1998) 1417-1425

Multiple Sensor Fusion and Motion Control of Snake Robot Based on Soft-Computing

Woo-Kyung Choi[1], Seong-Joo Kim[2], and Hong-Tae Jeon[1]

[1] School of Electronic and Electrical Engineering, Chung-Ang University
Donjak-Gu, Seoul, 156-756, Korea
chwk001@wm.cau.ac.kr
[2] Intelligent Control and Multimedia, Multi Channel Labs, Bucheon Technopark
Bucheon, Gyeonggi-Do, 420-734, Korea

Abstract. The recent development in robot filed shows that practical application of robot has transferred from industry to human's daily life. That is, robots which are modeled on human being as well as various animals have shown up. If a robot just moves around certain place as it controls its links, it is not more than a toy for children. A robot has to mount with various sensors to get information from environment, infer environment from sensor information and act properly as human being does with the five senses. In this paper, we made a snake shaped robot mounted with various sensors such as image, gas, temperature and luminosity sensor. The data from sensors is fused by soft-computing method. The snake robot recognizes environment with the fused sensor information and acts according to the result of expert system which is able to infer what proper action is.

1 Introduction

There are many circumstance limits to human like extreme radioactivity, temperature, chemical toxicity, pressure and so on. But, we need to detect and research region for making our sure safe and state. During long time, human have made a robot that have multi-joint biped robot and mobile robot using wheels. And now, that is developing but it still has some problem that is occurred by condition of the ground. Wheel structure is more adaptive to make high speed drive on flat ground and more efficiency to control the drive than other robot. But if it is on the non-flat ground or sandy road, wheel-based drive is not efficient way[2]. Also biped robot has wonderful adaptation to ground better than wheel-based mobile robot. Considering the condition of the rough ground, multi-joint biped robot that have no running gear like the snake robot have more adaptive than wheel-based robot[3][4]. If human make a moving robot operate like sneak, this sneak robot can solve the big problem that is the saving survivor, examination of harmfulness material in the dangerous situation like disaster(a earthquake, explosion and fire)[3]. Also it has a good adaptability and moving ability to use widely in difficult situation like researching of dangerous surrounding and medical part[3][5][6][7]. It is just like toy if robot merely operates a joint for driving. Human can recognize their surround using by the five senses. For it can be intelligence robot, robot must have independent sensors. It is possible to judge by comparing the information get through the sensor with the own information. Surely robot have to do that judge[1][8]. In this research, we realize the intelligence sneak robot

that is copied from biological sneak and that robot can judge intelligently about their circumstance. Surely, this robot can execute their mission by intelligence judging.

2   System Constitution of Snake Robot

We design a frame of sneak robot from biological structure feature of sneak. Left of figure 1 shows us a module that constitutes vertically connection of two servo-motor. The snake robot consists of 7 modules. And the robot has 12 DOF, because each module has 2 DOF by using 12 linear connecting of servo-motor. A power source is in the tail of robot for freely activity and Main module is in the head. Length between the head and the tail is 970mm.
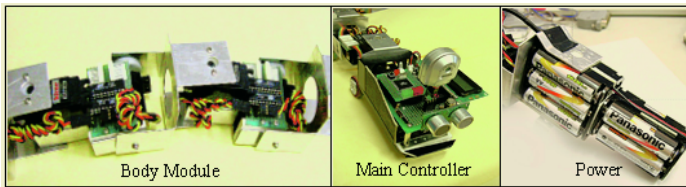
Fig. 1. Body, Controller and Power of Snake Robot

Sneak robot has ultrasonic sensor, CMOS image sensor, gas sensor, temperature, illumination for sensing circumstance. Table 1 means the usage.

Table 1. Sensors of Snake Robot

Sensor	Use	Character
Ultrasonic Sensor	Distance measurement, Avoidance	Range 3cm~3m
Image Sensor	Color recognition	120×90 Pixel
Gas Sensor	Gas detection	LNG, LPG
Temperature Sensor	Temperature measurement	-55~125℃
Illumination Sensor	Illumination measurement	CdS

Micro controller in head executes sensor fusion algorism by using sensor input data and this can reason proper action by expert system. Consequence of reasoning is sent to servo motor controller, and then servo motor working.

3   Sensor Fusion of Snake Robot

The snake robot has ultrasonic, gas, illumination, temperature and CMOS image sensor. Figure 2 is sensor fusion structure of the snake robot.

Preprocessing of sensor input is included in sensor modeling step. After sensor data is modeled, it is transferred to fusion or separate operation step. There are many kinds of fusion method, generally, neural network is used in lower level and rule based fusion system is used in little upper level. In this paper, complementary neural network is used in ultrasonic sensor fusion. In environment sensor fusion, Radial Basis Function Network(RBFN) is used. And CMOS image sensor, that can affect other sensor processing, is processed individually.
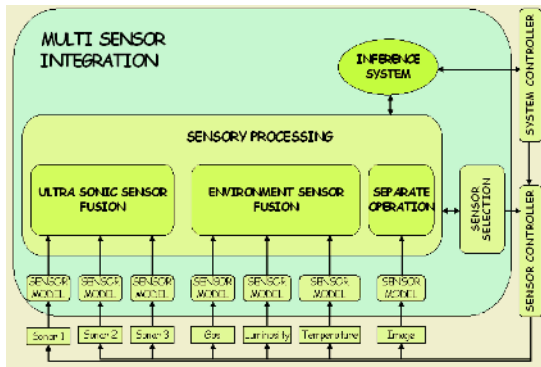
Fig. 2. Block Diagram of Sensor Fusion

3.1   Fusion of Ultrasonic Sensor

Ultrasonic sensor fusion used competitive neural network. Fusion module of ultrasonic sensor calculates distance between front side and the right and left sides, and reason traveling path of the snake robot. At this time, Relative difference of each direction-distance is important element, therefore competitive neural network is profitable than general multi-layer neural network. Competitive neural network uses unsupervised learning and classify input pattern without given information about target value.

3.2   Fusion of Environment Sensors

The snake robot has gas sensor and temperature sensor to grasp the dangerous degree of surrounding situation. Fusion module of environment sensors estimates dangerous degree of place that robot is situated currently with use value of gas and temperature sensor. Fusion module of environment sensor uses Radial Basis Function Network (RBFN).

3.3   Separation Management of Image Sensor

The snake robot can recognizes color and coordinates calculation about target by using CMOS image sensor. Image sensor handles distinction image of 120X90 pixels by YUV format. Due to limit of a memory, method of image process directly operate image received from camera by pixel instead of frame in a memory. If we use this method, memory space and operation speed can increase because image information do not stored in memory.

4   The Snake Robot's Inference Algorithm

The snake robot must judge the state of thing automatically and behave. Autonomous judgment of state and process of behavior can solve by soft computing method. For behavior of the snake robot is realized, it demands learning about environment of

various kinds and reasoning ability about behavior. Generally, neural network and fuzzy rule base are very useful method for learning and reasoning system. But, the snake robot's main process which consists of 8 bit has weak point that go down calculation ability than the PC's process. So, algorithm for reasoning of the snake robot used possible simple structure's neural networks and recognition and reasoning about various environment used rule base based on knowledge of expert.

4.1 Inference System

Sensor fusion information used input of inference system. Inference system use expert system. Knowledge-based method utilizes rule-based expression method. Rule-based expression consists of next structure.

$$\textbf{IF} \ \textit{Antecedent} \ \text{and} \ \textit{Antecedent}, \ \textbf{then} \ \textit{Consequent}. \tag{1}$$

Antecedent input of inference system used result which disposed of department of sensor fusion. So Antecedent of fuzzy rule was consisted of direction of ultrasonic sensor, result of environment sensor fusion, illumination value that handled separately and target value. Consequent of fuzzy rule determine movement of the snake robot which consists of 7 actions, and its movement is forward motion, backward motion, left-turn, right-turn, rest, attack and precaution. Each motion appeared to figure 3.
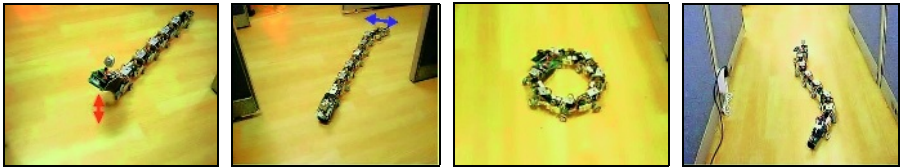
Fig. 3. Attack, Warning, Sleep and Locomotion Mode

Linguistic variables of direction element are made up of 4 kinds that consist of forward, backward, left and right and environment elements are 2 kinds(Danger, Safe) and target elements are 3 kinds(Nothing, Enemy, Prey). Finally, Linguistic variables of illumination elements is 2 kinds(Bright, Dark). We got 48 rules from all antecedent inputs as figure 4.

Rule 1: If *Dr is F*, *En is D*, *T is E* and *L is B*,   then *A is RETREAT*.
Rule 2: If *Dr is F*, *En is D*, *T is N* and *L is B*,   then *A is LOCOMOTION*.
Rule 3: If *Dr is F*, *En is S*, *T is N* and *L is Dk*, then *A is SLEEP*.
        ⋮
Rule 48: If *Dr is Bk*, *E is S*, *T is P* and *L is B*,   then *A is ATTACK*.

- *Dr*: Direction, *F*: Forward, *En*: Environment, *T*: Target, *S*: Safe,
  D: Danger, *Bk*: Backward, *L*: Luminorsity, *B*: Bright, *Dk*: Dark,
  A: Action, *B*: Bright, *E*: Enemy, *N*: Nothing, *P*: Prey

Fig. 4. Rule Base of Snake Robot

5   Experiment

5.1   Computer Simulation

We watched movement of the snake robot through computer simulation before real-time experiment of the snake robot hardware. Input for the simulation used distance values(left, front, right) measured by ultrasonic sensor, temperature, gas, illumination and target color measured by CMOS sensor. we easily ascertain fusion result of ultrasonic sensor and environment sensor fusion by computer simulation and get reasoning result of the snake robot using expert system. Simulation condition of figure 5 are not target and gas in the surrounding of the snake robot and temperature is 27 and illumination is brightness. Also enemy or prey is not appearance. Direction result from fusion module of ultrasonic sensor inferred 'FORWARD' and fusion department of environment sensor inferred that present circumstance is 'SAFETY'. Final inference result obtained form image sensor value is 'LOCOMOTION' by rule 11. Also we perform simulation that final inferences are 'WARNING', 'RETREAT', 'RATHER SAFETY', 'TURN RIGHT' and 'ATTACK'.
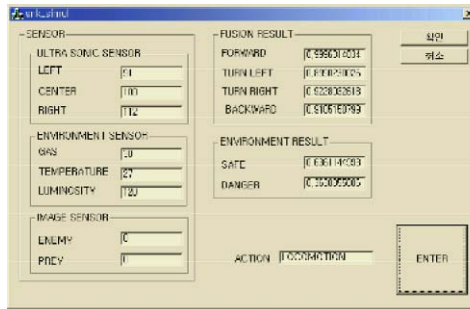
Fig. 5. Result of final inference is 'LOCOMOTION'

5.2   Implementing a Real-Time System of Snake Robot

The result of computer simulation showed that the snake robot determined proper judgment in every situation. We implemented the snake robot which consists of several frames to assure that the algorithm operate well in real-time system. We defined the blue object as a prey and the red object as an enemy in the experiment. For example, the snake robot meets across a prey in figure 6. We defined the blue object as a prey and the red object as an enemy in the experiment. For example, the snake robot meets across a prey in figure 6. The snake robot estimates the environment whether it is in safety situation or not and then, it attack the prey when it is in safety. Figure 7 shows that environment around the snake robot traveling is getting dark. The snake robot takes a rest when preys and enemies are not detected by image sensor and gas and temperature sensor display its safety. Figure 8 show that the snake robot detected a forward obstacle in traveling. It avoids the obstacle and turns left and goes straight after its the ultrasonic fusion module computes the distance between the snake robot and the obstacle.
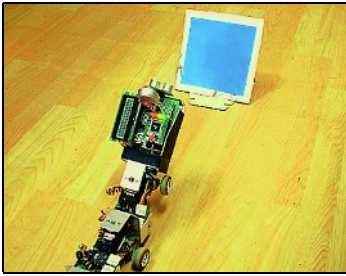
Fig. 6. Attack Motion
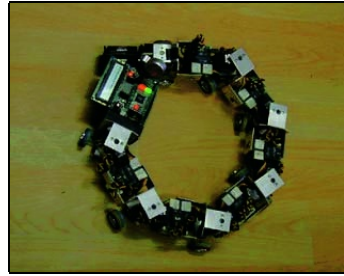
Fig. 7. Rest Motion

Fig. 8. Obstacle Avoidance

6 Conclusions

The goal of this paper in a snake robot and sensor fusion is that the snake robot which imitates a real snake's an activity and being adapted to topography and has multiple sensors operates well with considering environment around it. To avoid overloads of a processor and process a huge data of multiple sensors in distribute methods, fusion module of sensor is constructed in a module. In a low level of sensor processes, we worked sensor fusion processes using neural networks to have adaptability. In a high level of sensor processes, we make the snake robot operate intelligently using expert system in fused sensor data to infer activity of the snake robot. The snake robot is long and elliptical and has sensors on a head. If sensors are located on the snake robot's body and in addition magnetic and voice recognition sensors are used, abilities of the snake robot will be improved. To have the ability of learning in the snake robot, a processor to have a high ability must be used instead of using the processor used to operate the snake robot in this paper.

References

1. Wako Tojima., Robot kyoshitu, Kobunsha, 2001.
2. Masashi Saito, Masakazu Fukaya and Tetsuya Iwasaki, Serpentine Locomotion with Robotic Snakes, in IEEE Control System Magazine. February 2002.
3. K. Dowling, Limbless locomotion: Learning to crawl with a Snake Robot, in proc. IEEE Int. Conf. Robot. Auto., Detroit, MI, 1999, pp. 3001-3006.
4. Honda Motor Corporation, Honda Biped Robot, News Release. Several news articles including Reuters from December 20, 1996.
5. S. Hirose, Biologically Inspired Robots: Snake-Like Locomotors and Manipulators, New York: Oxford Univ. Press, 1993.

6. Hirose, Morishima, A., Disign and Control of a Mobile Robotic with an Articulated Body, Int. J. of Robot. Res, Vol.9, No.2, pp. 99-114, April 1990.
7. M. Nilsson, Snake robot free climbing, IEEE Control Syst. Mag., Vol. 18, pp. 21-26, 1998.
8. Ren C. Luo, Multisensor Fusion and Integration: Approaches, IEEE SENSORS JOURNAL, Vol. 2, No. 2, APRIL 2002
9. J. Gray and H. Lissmann, The Kinemetics of locomotion of the grass-snake, J. Exp. Biol., Vol. 26, pp. 354-367, 1950.
10. B.C. Jayne, Kinemetics of terrestrial snake locomotion, Copeia, Vol. 4, pp. 915-927, 1986.
11. H. Lissmann, Rectilinear locomotion in a snake(Boa occidentalis), J. Exp. Biol., Vol. 26, pp. 368-379, 1950.

Human Face Detection Using Skin Color Context Awareness and Context-Based Bayesian Classifiers

Mi Young Nam and Phill Kyu Rhee

Dept. of Computer Science & Engineering Inha University
Yong-Hyun Dong, Incheon, Korea
rera@im.inha.ac.kr, pkrhee@inha.ac.kr

Abstract. We propose a cascade detection scheme by combining the color feature-based method and appearance-based method. In addition, the scheme employs illumination context-awareness so that the detection scheme can react in a robust way against dynamically changing illumination. Skin color provides rich information for extracting rough area of the face. Difficulties in detecting face skin color come from the variations in ambient light, image capturing devices, etc,. Appearance-based object detection, multiple Bayesian classifiers here, is attractive since it could accumulate object models by autonomous learning process. This approach can be easily adopted in searching for multiple scale faces by scaling up/down the input image with some factor. The appearance-based method shows more stability under changing illumination than other detection methods, but it is still bordered from the variations in illumination. We employ Fuzzy ART and RBFN for the illumination context- awareness. The proposed face detection achieves the capacity of the high level attentive process by taking advantage of the illumination context-awareness in both color feature-based detection and multiple Bayesian classifiers. We achieve very encouraging experimental results, especially when illumination condition varies dynamically.

1 Introduction

Much research is carried out in automatic face detection area [1,2]. Even though human beings can detect face easily, it is not easy to make computer to do so. The difficulties in face detection are caused by the variations in illumination, pose, complex background, etc. Robust object detection needs high invariance with regard to those variations. Much research has been done to solve this problem [3,4,5].

The face detection systems can be divided into three major categories [2]: the model-based method, feature-based method, and appearance-based learning method. In the model-based method, face is detected by a geometrical deformable model by grouping facial attributes such as eyes, nose, and mouth. In the feature-based approach, some obvious and most invariant facial features are used for face detection. Color is the most obvious feature among those of the feature-based approach. Color-based face detection is very much intuitive and has long history. However, few papers are published. Cahi et. al [6] have compared five color chrominance planes and two skin modeling methods: lookup table and Bayes skin probability map. Jones et. al. [7] have provided the comparison of nine chrominance spaces and two parametric techniques. Skin–color provides rich information for extracting rough area of the face. Difficulties in detecting face skin color comes from the fact that ambient light, image capturing devices, etc. can affect skin color too much.

In the appearance-based approach, the pixel representation of human face itself is considered as a pattern. This approach can search for multiple scale faces by scaling up or down the input image with some factor. The appearance-based method is more stable than two other methods, but it is still bordered from the variations appeared in input images. Recently, the concept of context-awareness has been applied for robust face detection under dynamically changing environments [8].

The goal of this paper is to explore the possibility of illumination-insensitive face recognition by adopting the concept of context-awareness and combining color feature-based and appearance-based methods in cascade manner. The context-awareness is employed to improve robustness of the proposed method. Context, in this paper, is modeled and applied as the change of illumination. As illumination context changes, skin color model is restructured or reselected. Restructuring the skin color model enables the scheme to adapt to changing illumination and to perform an action that is appropriate over a range of the change.

In this paper, we propose a cascade detection scheme using adaptive color model and multiple Bayesian classifiers. The scheme also employs illumination context-awareness so that the detection scheme can react in a robust way against dynamically changing illumination. We employ Fuzzy ART and RBFN for the context modeling and awareness, respectively. They provide the capacity of the high level attentive process by the illumination context-awareness in face detection. We achieve very encouraging results from extensive experiments. The outline of this paper is as follows. In section 2, we will present the architecture of the proposed face detection scheme. Adaptive color modeling using illumination context awareness will be discussed in section 3. In section 4, we will present multiple Bayesian classifies based on illumination context-awareness. We will give experimental results in section 5. Finally, we will give concluding remarks.

2   The Proposed Face Detection Architecture

The specific task we are discussing in this paper is a real time face detector which is adaptive against changing illumination environments. The system consists of illumination context–aware module, search area reduction module using illumination context based color model, multi-resolution pyramid, face/non-face determination using multiple Bayesian classifier as shown in Fig. 1. The details will be discussed in the followings.
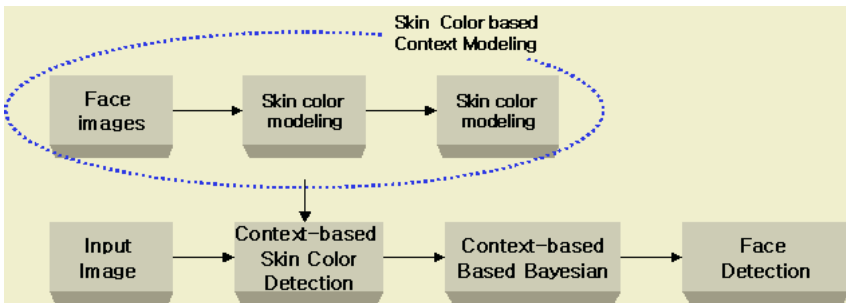
Fig. 1. Face detection system architecture

In the proposed method, the feature space for object detection with multiple scales and varying illumination condition has been partitioned into subspaces properly so that the location error can be minimized. However, the partitioning of multiple scales and varying illumination condition is very subjective and ambiguous in many cases. Another problem is the accumulation of properly labeled training data. The detailed structure of skin color modeling and detection proposed architecture is shown Fig.2.
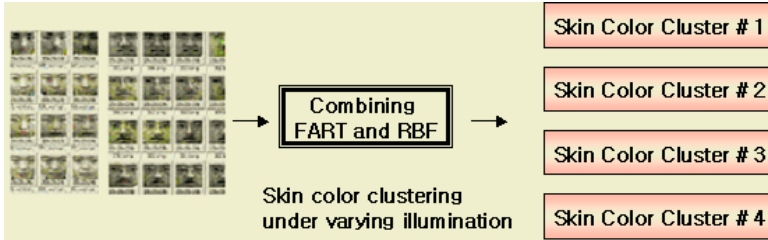
Fig. 2. The detailed structure of skin color modeling

Illumination condition (context) is analyzed using the combined network RBF network and FuzzyART. Fig.2 shows illuminant and face model training method. It is combined method form supervised and unsupervised learning method As showed Fig. 3, the idea is to train the network in two separate stages - first, we perform an unsupervised training (FuzzyART) to determine the Gaussians' parameters (j, j). In the second stage, the multiplicative weights w_{kj} are trained using the regular supervised approach. Input pattern is vectorized for grayscale image size of 20x20 pixels, input node had mosaic of size of 10x10 pixels. The transformation from the input space to the hidden unit space is non-linear, whereas the transformation from the hidden-unit space to the output-space is linear. RBF classifier expands input vectors into a high dimensional space. RBF network has architecture that of the traditional three-layer back-propagation. In this paper, hidden units are trained using FuzzyART network and basis function used are Gaussians. The proposed network input consists of n normalized and rescaled size of 1/2 face images fed to the network as 1 dimension vector. Input unit has floating value [0,1]. The vector value is normalized. FuzzyART performance is best in case used picture itself by input node vectorized.
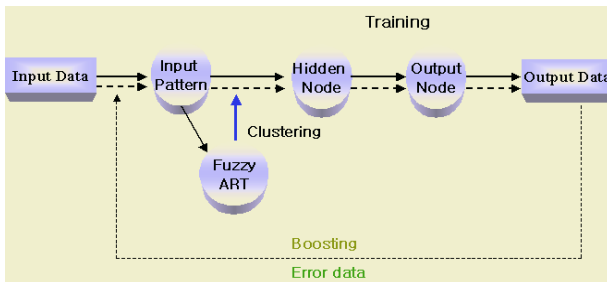
Fig. 3. Illumination clustering method

The proposed adaptive color model using the illumination information determines face skin area. Multi-resolution pyramid generates face candidate region. Pyramid

method based object detection is attractive since it could accumulate the object models by autonomous learning process [9]. In our approach, the object scale is approximated with nine steps. In this module, the search area is reduced by finding a face color region, and 20×20 windows are generated by the pyramid module. A tradeoff must be considered between the size of multiple scale, view representation of the object and its accuracy.

Finally, context-based Bayesian is in a multi classifier combination four Bayesian classifiers. Each Bayesian classifier is trained using face images in each illuminant cluster. Face candidate windows are selected using the multiple Bayesian classifiers, and finally, face window is filtered by the merging process.

In the learning stage, the candidate face regions are clustered into 4 face models, 6 face-like non-face models, non-face models using combined learning algorithm [10], the threshold parameters of multiple Bayesian classifiers are decided. Initially, seed appearance models of face is manually gathered and classified for training the detector module. The detector with prior classes is trained by the initial seed data set.

3   Adaptive Skin Color Modeling Using Context Awareness

The use of color information can simplify the task of face localization in complex environments [11, 12]. Skin–color provides rich information for extracting rough area of the face [6, 9]. Several skin color spaces have been proposed [6]. Difficulties of detecting face skin color comes from no universal skin color model can be determined. Ambient light, un-expectable in many cases, affects skin color. Different cameras produce different colors under same light conditions. Most previous approaches of color based face detection are bordered from unexceptional behavior of the detection scheme when illumination condition cannot be strongly constrained. Even though much research has be performed single statistical color model fails to satisfies for varying illumination condition.

In the face color detection module, the face search space is reduced using global face color context analysis, and the reduced face search space is scanned in the pyramid method. The result of the pyramid module is 20×20 standard face candidate windows. The search space reduction using facial color is performed as follows. Input image is 320×240 size. We define face color when histogram value over a threshold, where the threshold value is determined empirically. Initially, seed appearance samples of face color area are manually gathered and used for modeling face color.

Skin color's area is computed 1,200 face images of 20 x 20 pixel scale. There are several color spaces such YCrCb, HSV, YCbCr, CbCr. Even though which color space is most effective in separating face skin and non-face skin colors, YCbCr has been widely used [6]. We choose CrCb color space which almost approaches the performance of YCbCr color space and provides faster execution time than other color spaces. Fig.4 is shown context-based skin color histogram.

In this paper, we adopted multi-clustered color space for each illumination context, instead of single cluster color space. As shown Fig. 4, face skin color in the CbCr color space is clustered in four regions. We determine a region as skin color region when histogram value over a threshold because face images contain non-skin color. The threshold value is determined empirically.
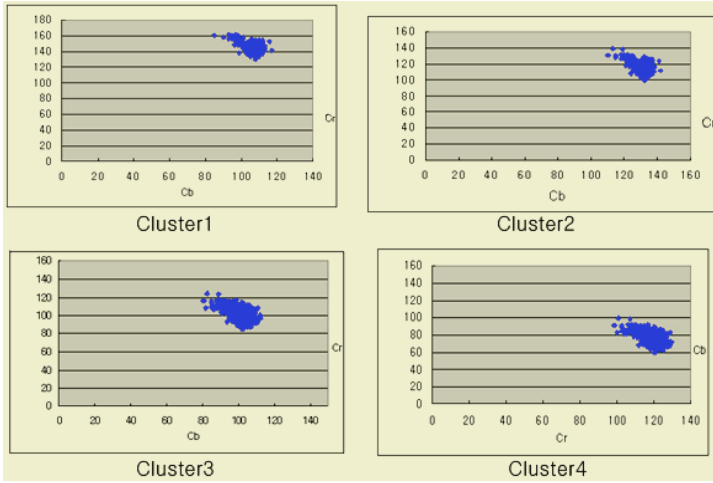
Fig. 4. The skin color histogram distribution for each cluster

Skin color candidate area compute mean and distribution of current skin color model.

$$\mu = \frac{1}{W} \sum_{i \in W} P_i \tag{1}$$

$$\sigma^2 = \sum_{i=0}^{W} (X_i - \mu)/W \tag{2}$$

The detection problem is to find a face candidate region using the skin-color model as formulated in the following equation.

$$Fs = \sum_{i=0}^{W} Face, \quad nFs = \sum_{i=0}^{W} nonFace \tag{3}$$

$$FaceArea = Fs : nFs > threshold \tag{4}$$

(a) Fram1   (b) Fram2
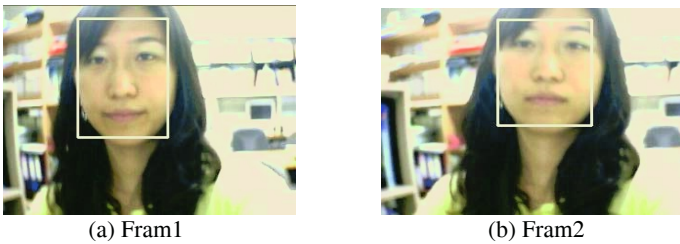
Fig. 5. Skin color detection example in real time system (skin color detection result using adaptive color model.)

4   Context-Based Bayesian Classifiers

We show that illumination context based Bayesian face modeling can be improved by strengthening both contextual modeling and statistical characterization. Multiple Bayesian classifiers are adopted for deciding face and non-face. A Bayesian classifier

is produced for each illumination context. We have modeling four Multi class based Bayesian classifiers are made in 4 group for face.

Initially, seed appearance models of face is manually gathered and classified for training the appearance detector module. Each detector, Bayesian classifier here, with prior classes is trained by the training data set. Training set of faces is gathered and divided into several group in accordance with illumination contexts. Multi-resolution pyramid consists of nine levels by an experiment, and the offset between adjacent levels is established by four pixels. Each training face image is scaled to 20x20 and normalized using the max-min value normalization method, and vectorized. Since the vectorization method affects much to face detection performance, we have tested several vectorization methods and decided 54 dimensional hierarchical vectorization method [13, 14] (see Fig. 6). We use sub-regions for the feature of first classifier. Illumination the training images to a number of predefined clustering goroup by K-means. The detector trained each face illuminant cluster.
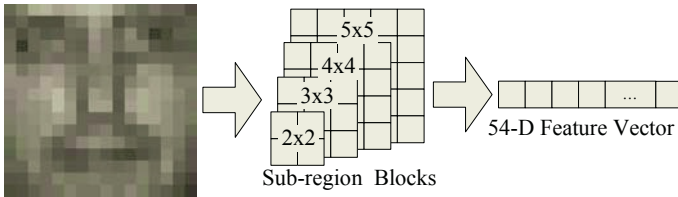
Fig. 6. Face feature extraction method (face image vectorization)

Squared Mahalanobis distance [4], a useful method that measures free model and free pattern relationship using similarity degree, is adopted for the Bayesian classifier. The center of cluster is determined by the mean vector, and the shape of the cluster is determined by the covariance matrix Mahalanobis distance as shown in Eq. 5.

$$r^2 = (x - \mu_x)' \Sigma^{-1} (x - \mu_x), \tag{5}$$

where μ_x is the average vector of face's vector, Σ is the covariance matrix of independent elements. We use 4 face group that is shown Fig. 4. This is generated by proposed learning method. The y axis is cluster number and x axis is sample under each cluster. The advantages of the two-stage training are that it is fast and very easy to interpret. It is especially useful when labeled data is in short supply, since the first stage can be performed on all the data (not only the labeled part).

Fig.7, shows the face candidate of Fig.4: blue box is face candidate area and white box is exact face area.
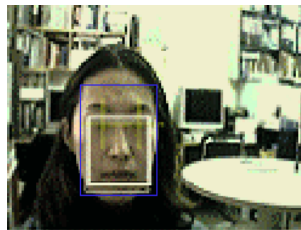
Fig. 7. The face detection using proposed method (skin color area and context based Bayesian)

5 Experiment

The experiment of the proposed method has been performed with images captured in various environments 800 images are captured and used in the experiment. The superiority of the proposed method is discussed in the following.

Fig.8 shows face detection in real-time image by using the proposed cascade of context-based skin color and context-based Bayesian methods.
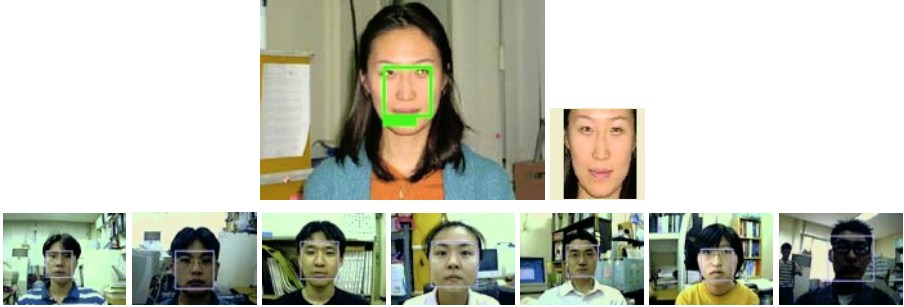
Fig. 8. Examples of bad illuminant face detection

Face detection result shows Table1-5. Images have size of 320 x 240 pixels and encoded in 256 color levels. We resized to various the size using multi-resolution of 9 steps. Rescaled images are transferred. Each face is normalized into a re-scaled size of 20x20 pixels and each data – training image and test images – is preprocessed by histogram equalization and max-min value normalization. Tables compare the detection rates (AR) for the context-based color system and the numbers of context-based Bayesian classifiers for the three boosting-based systems, given the number of false alarms (FA) and false rejection (FR).

Table 1 shows skin color detection result between context based color model and non-context color model. Table 2 shows that the result of face detection between multi context-based Bayesian and single context based Bayesian after multi context based skin color detection.

Table 1. Skin color detection on our lab in color image

Method	Total Faces	AR	FR	FA
Context-based color model	800	791	9	10
Normal Skin color model	800	750	30	50

Table 2. The face detection on our lab in color image using (using propose method adaptive color model and multi-class face model)

Method	Total Image	AR	FR	FA
Context-based Bayesian	801	790	0	0
Normal Bayesian	801	755	35	6

Table 3 shows that the result of face detection between multi context based Bayesian and single context based Bayesian after single context based skin color detection. Table 4 shows that the result of face detection between multi context based Bayesian and single context based Bayesian in gray images.

Table 3. Face detection on Our Lab in Color Image using

Method	Total Image	AR	FR	FA
Context-based Bayesian	850	750	0	0
Normal Bayesian	850	732	18	30

Table 4. Face detection result in FERET Dataset Normal Images ((Proposed method and other method)

Method	Total Image	AR	FR	FA
Context-based Bayesian	1196	1189	5	10
Normal Bayesian	1196	1130	21	40

Table 5. Face detection result in FERET Dataset Dark Images((Proposed method and other method)

Method	Total Image	Accept	FRR	FAR
Context-based Bayesian	194	188	5	3
Normal Bayesian	194	130	25	45
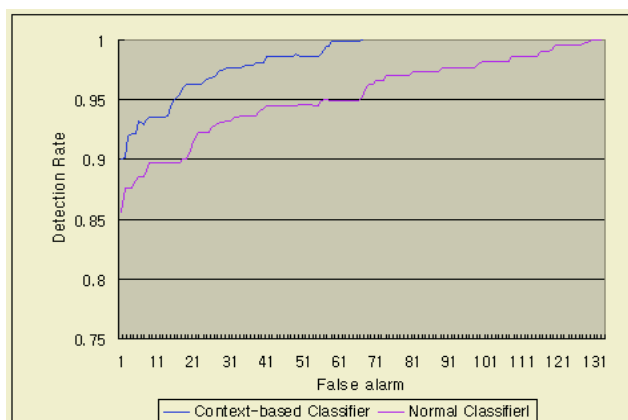
Fig. 9. Comparison context based face detection and non-context based face detection in captured images

From Tables, we know that propose method is good face detection performance other method. By we were going to do studying repeatedly, clustering's performance improved. As seen in tables and figures, error in front side face appears in tilted

image. We must consider about tilted image. We could improve pose classification's performance for face through recursive studying like experiment result. The combined effect of eye lasses is also investigated. In this experiment, the factor of glasses and illumination was considered and experimental images were classified by the factor of glasses and illumination. We classified bad illumination images into the image including a partially lighted face, good images into that including a nearly uniformly lighted face.
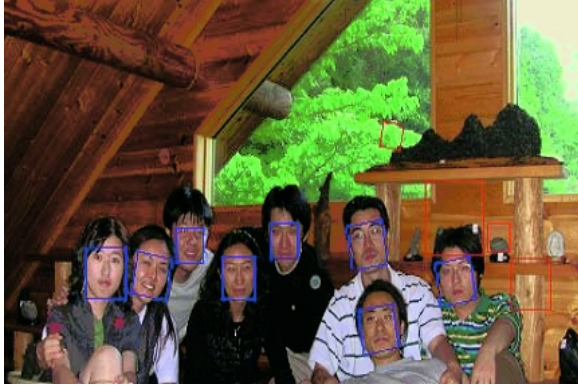
6   Concluding Remarks

This paper discusses a cascade detection scheme by combining the color feature-based method and appearance-based method with the capability of illumination context-awareness. Even though much research has been done for object detection, it still remains a difficult issue. The detection scheme aims at robustness as well as fast execution way under dynamically changing illumination. The cascade architecture of face detection provides fast execution speed and accuracy. Difficulties in detecting face coming from the variations in ambient light, image capturing devices, etc. has been resolved by employing Fuzzy ART and RBFN for the illumination context- awareness. The proposed face detection achieves the capacity of the high level attentive process by taking advantage of the illumination context-awareness. We achieve very encouraging experimental results, especially when illumination condition varies dynamically. Experimental result has shown that the proposed system has detected the face successfully 98% of the offline image under varying illuminant.

Acknowledgement

Reference

1. T. V. Pham, et. Al.,:Face detection by aggregated Bayesian network classifiers. Pattern Recognition Letters 23 (2002) 451-461
2. Li, S.Z.; Zhenqiu Zhang.: FloatBoost learning and statistical face detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 26. (2004) 1112 – 1123
3. B.K.L. Erik Hjelmas: Face Detection: A Survey. Computer Vision and Image Understanding, vol. 3. no.3. (2001) 236-274
4. C. Liu.: A Bayesian Discriminating Features Method for Face Detection. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25. (2003) 725-740
5. H.Schneiderman and T. Kanade, "Object Detection Using the Statistics of Parts," Int'l J. Computer Vision, Vol. 56. no. 3. (2004) 151- 177
6. Cahi, D. and Ngan, K. N.: Face Segmentation Using Skin-Color Map in Videophone Applications. IEEE Transaction on Circuit and Systems for Video Technology, Vol. 9. (1999) 551-564
7. Jones, M.J. ,Rehg and J.M.: Statistical color models with application to skin detection. Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference, Vol. 1. (1999) 23-25
8. Context Driven Observation of Human Activity
9. Yang, M. H., Kriegman, J. and Ahuja, N.: *Detecting Faces in Images: A Survey. IEEE Transaction on Pattern Analysis andMachine Intelligence*, Vol. 24. (2002) 34-58
10. M.Y Nam and P.K Rhee.: *A Scale and Viewing Point Invariant Pose Estimation*. KES2005, (2004) 833-842.
11. D. Maio and D. Maltoni.: Real-time face location on gray-scale static image. Pattern Recognition, Vol. 33. no. 9. (2000) 1525-1539
12. M. Abdel-Mottaleb and A. Elgammal,: Face Detection in complex environments from color images. IEEE ICIP, (1999) 622-626
13. P. Viola and M. Jones,: Robust Real Time Object Detection. IEEE ICCV Workshop Statistical and Computational Theories of Vision, July (2001)
14. S.Z. Li, L. Zhu, Z.Q. Zhang, A. Blake, H. Zhang, and H. Shum.: Statistical Learning of Multi-View Face Detection. Proc. European Conf. Computer Vision, Vol. 4. (2002) 67-81

Adaptive Gabor Wavelet
for Efficient Object Recognition

In Ja Jeon, Mi Young Nam, and Phill Kyu Rhee

Dept. Of Computer Science & Engineering, Inha University
253 Yong-Hyun dong, NamGu, Incheon, South Korea
{juninja,rera}@im.inha.ac.kr, pkrhee@inha.ac.kr

Abstract. This paper describes, using situational awareness and Genetic algorithm, a run-time optimization methodology of the Gabor wavelet parameters so that it produces a feature space for efficient object recognition. Gabor wavelet efficiently extracts the feature space of orientation selectivity, spatial frequency and spatial localization. Most previous object recognition approaches using Gabor wavelet do not include systematic optimization of the parameters for the Gabor kernel, even though the system performance might be much sensitive to the characteristics of the Gabor parameters. This paper explores efficient object recognition using adaptive Gabor wavelet based situational aware method. The superiority of the proposed system is shown using IT-Lab, FERET and Yale face database. We achieved encouraging experimental results.

1 Introduction

Image objects appear in different forms under different vision conditions such as illumination, scale, background, etc. The illumination environment is one of the main factors that influence the precision of object recognition [1]. In order to control the variation of the illumination, object recognition typically is realize under certain restricted conditions. In many applications, it is usual to require the ability to recognize objects under varying environments, even when training images are severely limited. When humans interact with other persons and surrounding environments, they take advantage of situational information. Computers are not as good as humans are in deducing situational information from their environment and in using it in their missions. The use of context or situational information is especially important in an object recognition system.

Much research has devoted on object recognition using Gabor wavelet, however, a few research has dealt with the run-time efficiency by adapting Gabor parameters under varying situations. Traditional system schemes cannot be applicable under dynamically changing situations or environments since they usually require too much computation time or resource to achieve an optimal performance. In this paper, we introduce an adaptive Gabor representation module with the capability of situational awareness. The situational awareness method is applies to manage the variations coming from illumination environments. The role of adaptive Gabor representation module is to find an optimal Gabor feature vectors using situational information so that the recognition system could achieve best performance under varying illuminations. The recognition system using the proposed Gabor parameter adaptation has tested under varying illumination environments. In situational awareness of experimental system, the control process of situational awareness is generally composed of

two phases: the estimation of illumination direction, and the determination of recognition parameters. Reconfigurable architecture of object recognition has two levels that the first level is situational awareness state, and the second level is recognition process that is performed using detected information by situational awareness.

The outline of this paper is as follows. In section 2, we present the situational awareness under varying illumination environment. In section 3, we present the Gabor wavelet definitions and its adaptive parameters. The design methodology of the proposed Gabor kernel and feature space was described in section 4. In section 5, we present run-time optimization of the Gabor parameter and experimental results. Finally, we give concluding remarks.

2   Situational Awareness Under Varying Illumination

Even though many algorithms and techniques are invented, face recognition remains a difficult problem yet, and existing technologies are not sufficiently reliable, especially in the environments of dynamically changing illumination. Situation can be various configurations, computing resource availability, dynamic task requirement, application condition, environmental condition, etc. [2].

In this paper, we approach the problem of face recognition by adapting illumination environments. This step is extract circumstance from input image, and we performed estimation for situation using BPNN algorithm. At this step detected parameter values that it is suitable preprocessing algorithm and parameter using recognition algorithm according to situation. Preprocessing level is transformed optimal image to recognize from current image including illumination. We provide necessary parameter to calculate suitable Gabor wavelet from situational awareness level for effective recognition. We calculate Gabor wavelet to use parameter. At the next step, use calculated value, compare with beforehand stored user's data and measure similarity. Therefore, situational awareness is very important in our system architecture. Because a similarity change according to given parameter at situational awareness.
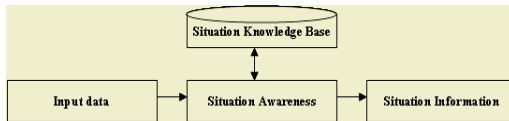
Fig. 1. A module for selecting situational awareness

We have designed the situational awareness module as Figure 1. The situational information has derived by the situational awareness module using back-propagation neural network. After the illumination variation of the face is estimate, the recognition has processed using the Gabor representation adapted in accordance with the identified illumination variation. We start developing our system by first considering the problem of face recognition under varying illumination condition with frontal face. The system accumulates the situational knowledge that guarantees optimal performance for each identified situation. The situational knowledge base stores the expressions of identifiable situations and their matched actions will performed by the situational awareness module. The matched action can decide by either experimental trial-and-error or some automating procedures. In the operation time, the situation

expression is determined from the derived situation information, where the derived situation information is decide from the illumination in image. Figure 2 shows the IT Lab. test images for illumination environment estimation.

Fig. 2. Illumination and noise environment dataset

Recognition state influenced of environmental factors, such as illumination, scale, background, etc. Recognition of face image is due to the influence of illumination, scale, expression, etc. The illumination fluctuation is one of the main factors that influence the precision of face recognition. In order to control the variation of the illumination, face recognition systems is realized under certain restricted conditions. However, the method developed under the restricted environment is not always suitable to uneven environment conditions. Therefore, it is necessary to study the improvement of face images under the variable illumination and noisy environment.

3   Gabor Wavelets

Our objective is to design of situational awareness architecture that recognition is robust to variable illumination and noisy. For recognition used Gabor wavelet. Gabor wavelet efficiently extracts orientation selectivity, spatial frequency, and spatial localization [3],[4]. Gabor wavelet is biologically motivated convolution kernels in the shape of plane waves restricted by Gabor kernel. The convolution coefficient for kernels of different frequencies and orientations starting at a particular fiducial point is calculated. We discuss the discriminating performance of three Gabor Kernels construction method. The first type of Gabor kernels given by are defined as follows [5]:

$$\psi(x,y) = \frac{k_j^2}{\sigma^2} \exp(-\frac{k_j^2(x^2+y^2)}{2\sigma^2}) \bullet [\exp(i\vec{k_j}(x^2+y^2)) - \exp(-\frac{\sigma^2}{2})]$$

$$\vec{k_j} = (k_v \cos\theta_\mu, k_v \sin\theta_\mu)^T, k_v = 2^{\frac{-v+2}{2}}\pi, \theta_\mu = \mu\frac{\pi}{n}$$

$$j = \mu + 8v,\ v = 0,...,m,\ \mu = 0,...,n$$

(1)

The family of Gabor kernels is similar each other since they are generated from one mother wavelet by dilation and rotation using the wave vector. The effect of the DC term vanishes when the parameter $(\vec{k})_{\mu,v}$ has sufficiently high values, where σ determines the ratio of the Gaussian window width to wavelength. Gabor wavelet is usually used at five different frequencies, ν = 0,…, 4, and eight orientations, μ = 0,…, 7 [6],[7]. The kernels show desirable characteristics of spatial locality and orientation selectivity, a suitable choice performed face image feature extraction for classification. We will call this Gabor wavelet LW that is the first letter of the first author name, Laurenz Wiskott. The second type of Gabor kernel was denoted as follows:

$$\psi(x,y) = \frac{\omega_v}{\sqrt{2\pi\kappa}} \exp(-\frac{\omega_v}{8\kappa^2}(4x'^2 + y'^2)) \bullet [\exp(i\omega_v x') - \exp(-\frac{\kappa^2}{2})]$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta_\mu & \sin\theta_\mu \\ -\sin\theta_\mu & \cos\theta_\mu \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (2)$$

$$\omega_v = v\omega_0, \ \theta_\mu = \mu\frac{\pi}{n}, \ v = 0,...,m, \ \mu = 0,...,n$$

Where ω_0 is the radial frequency in radians per unit length and θ is the wavelet orientation in radians. The Gabor wavelet is centered at *(x=0, y=0)* and the normalization factor is such that $<\psi,\psi = 1>$, i.e., normalized by L^2 norm. *x* is a constant, with *x* $\approx$ 2.5 for a frequency bandwidth of 1.5 octaves. We will call this Gabor wavelet TSL that is the first letter of the first author name, Tai Sing Lee. The third type of Gabor kernel is Eq(3).

$$\psi(x,y) = \exp\left\{-\pi\left[\left(\frac{x'^2}{\sigma}\right)^2 + \left(\frac{y'^2}{\alpha\sigma}\right)^2\right]\right\} \bullet \exp\{i[u(x-x_i) + v(y-y_i)]\}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta_\mu & \sin\theta_\mu \\ -\sin\theta_\mu & \cos\theta_\mu \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (3)$$

$$\omega_v = v\omega_0, \ \theta_\mu = \mu\frac{\pi}{n}, \ v = 0,...,m, \ \mu = 0,...,n, \ u = \omega_v\cos\theta_\mu, \ v = \omega_v\sin\theta_\mu$$

Where (x_i, y_i) is the spatial centroid of the elliptical Gaussian window whose scale and aspect are regulated by σ and α, respectively. ω_k and $\theta_l(k,l \in N)$ are the modulation frequency and direction, respectively, and (u,v) are the frequency components in x and y directions, respectively. The scale σ controls the size of the filter as well as its bandwidth, while the aspect ratio α and the rotation parameter ϕl (generally set equal to θl) control the shape of the spatial window and the spectral bandwidth. We will call this Gabor wavelet XW that is the first letter of the first author name, Xing Wu [8].

4   Gabor Kernel and Feature Space

The Gabor wavelet transformation of an image is defined by the convolution of the sub-area of image using a family of Gabor kernels as defined by Eq(4). Let $f(\vec{x})$ be the gray value of an sub-image around pixel $\vec{x} = (x, y)$, and the Gabor wavelet transform of the sub-image is defined as follows:

$$G_{\mu,v}(\vec{x}) = f(\vec{a}) * \varphi_{\mu,v}(\vec{a})$$
$$= \iint f(\alpha,\beta)\varphi_{\mu,v}(x-\alpha, y-\beta)d\alpha\, d\beta \qquad (4)$$

Where $\vec{x} = (x,y)$, and * denotes the convolution operator. Let $G_{\mu,v}^{(e)}(\vec{x})$ denote the normalized vector constructed from $G_{\mu,v}(\vec{x})$ (adapted by the normalization factor *e* and normalized to zero mean and unit variance), the Gabor feature vector $F^{(e)}$ at a fiducial point $\vec{x}$ is defined as follows:

$$F^{(e)}(\vec{x}) = (G_{0,0}(\vec{x})^{(e)^t} G_{0,1}(\vec{x})^{(e)^t} \cdots G_{4,7}(\vec{x})^{(e)^t})^t \qquad (5)$$

The feature vector thus includes all the Gabor transform at the fiducial point $\vec{x}$, $G_{0,0}(\vec{x})$, μ = 0, ... , 20, ν = 0, ..., 20, as it derives an optimal discriminating information for a given external environment using the evolutionary module discussed in the next section. The design of optimal Gabor vector includes to determine a optimal Gabor kernel parameters and select proper fiducial points. This implies the center frequency of the filter is related to the rotation angle of the modulating Gaussian by introducing the new parameter, radio frequency.

The Gabor wavelet by have also three parameters, frequency(ω), orientation(θ) and the ratio of the standard deviation of an elliptical Gaussian to frequency(κ) [5]. It is used at three different frequencies and twenty orientations. κ is fixed for Gabor wavelets of a particular bandwidth. The Gabor wavelet by Malsberg have six parameters, frequency(ω), orientation(θ), size of filter(σ), aspect ratio(α), scaling factor(ρ) and bandwidth-frequency(λ). It is used at seven frequencies and eight orientations. Other parameters of the filters were supposed as α=1, ρ=$\sqrt{2}$, λ=$\pi/4$.

Although Gabor wavelet has the different number of parameters, it is important to optimize parameters for face recognition. Face features are represented by multiple fiducial points, which are characterized by Gabor wavelet transform. Most of these are located at well-defined features, which are likely to be significant, such as pupil, mouth, nose, and so on. A set of fiducial points was marked manually 24 fiducial points that do by standard established to do manually and different 40 fiducial points. The geometries of fiducial points given in Figure 3 are tested. Could acquired effective recognition rate made change for fiducial point number according to size of using frontal face image.

(a)   (b)   (c)   (d)
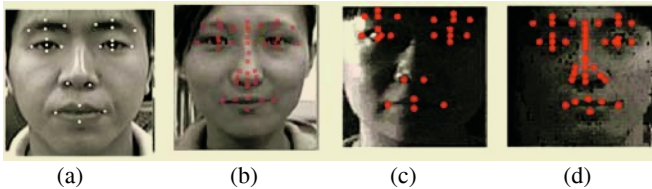
Fig. 3. The geometry of fiducial points: (a) and (b) 24, 40 fiducial points in image of size 128*128 in normal illumination environment, (c) and (d) 24, 40 fiducial points in image of size 128*128 in uneven illumination environment

5   Adaptive of Gabor Parameter Construction

The adaptation of optimal Gabor vector includes to determine optimal Gabor kernel parameters and to select optimal feature space depending on identified situation. In this paper, Gabor kernel adaptation involves determining a 5-tuple (frequency, orientation, width, height and absolute phase) in according to varying situation.

5.1   Operation Parameter Construct

GA is employed to search among the different combination of kernel design parameters, fiducial points, and vector elements of each Gabor feature vector. The chromo-

some represents the all possible combination of kernel design parameters, fiducial points and their Gabor feature vectors. The optimality of the chromosome is defined by classification accuracy and generalization capability. The total Gabor feature vector for all fiducial points, V is evolved from a larger vector set defined as follows:

$$V = \left(F^{(e)}\left(\vec{x_1}\right)F^{(e)}\left(\vec{x_2}\right)\cdots\cdots F^{(e)}\left(\vec{x_n}\right)\right) \tag{6}$$

$\vec{x_1}, \vec{x_2}, \cdots, \vec{x_n}$ are fiducial points. The contribution of fiducial points is in n-dimensional space by a set of weights in the range of (0.0, 1.0). As it searches the genospace, the GA makes its choices via genetic operators as a function of probability distribution driven by fitness function. The genetic operators used here are selection, crossover, and mutation [4].

5.2 Fitness of GA

Evolution or adaption was guide by the fitness function defined for the system performance and the class scattering criterion. The system performance denotes the correctness that the evolvable adaptation has achieved so far, and the class scattering indicates the expected fitness on future generations. The evolutionary module derives the classifier balanced between successful recognition and generalization capabilities. The fitness function can defined as follows:

$$\eta(V) = \lambda_1 \eta_s(V) + \lambda_2 \eta_g(V) \tag{7}$$

Where $\eta_s(V)$ is the term for the system correctness and successful recognition rate and, $\eta_g(V)$ is the term for class generalization. λ_1 and λ_2 are positive parameters that indicate the weight of each term, respectively.

5.3 Optimization Process of Gabor Parameter

The evolutionary approach as previously discussed has employed to decide the Gabor parameter of the autonomous object recognition system for various input environments. The recognition system learns a given input environment, the Gabor parameter and classifier architecture adapt itself by restructuring its structure and optimizing kernel parameters. The image filters employed here is the lighting compensation, histogram equalization, opening operation, boost-filtering [8]. A-nn (Approximate Nearest Neighbor) algorithm is employs for the recognition [9]. The A-nn employed here is deterministic and easy to implement. The reference face points in the Gabor space is preprocessed into a hierarchical decomposition of the space, called balanced box decomposition (BBD) tree. The height of the tree is O(log n), and subdivided space is O(d) complexity, where d is the dimension of the Gabor space. The face points associated the leaf is computed, the closest point found so far is tracked.

6 Experimental Results

Figure 4 shows the proposed system architecture of face object recognition using the situational awareness. The performance evaluation of the proposed method performed using our laboratory, FERET, Yale data sets. We used, as test data, 1000 images of 100 persons from our lab data set, and 2418 images of 1209 persons from FERET

data set. We employed the leave-one-out method in our performance evaluation. We compared the performance of three different Gabor wavelets using images of resolution 128×128 in order to select most proper method in our recognition system. Because each Gabor wavelet may have its own optimal parameters, the comparison is performed by changing the values of it's parameters within the accepted range.
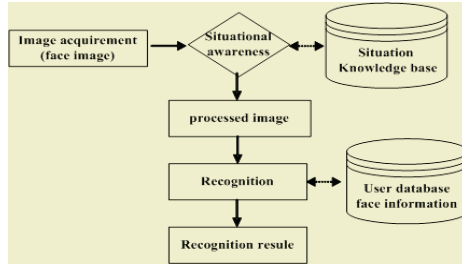
Fig. 4. Object recognition structural design using the situational-awareness

Table 1. The performance comparison of three different Gabor wavelet method

Sigma #of error	π	1.5π
9*9	11	11
17*17	2	3
33*33	2	1

(a) LW

# of error / Sigma	π	1.5π	2π	2.5π	3π
9*9,π	12	32	19	18	18
9*9,2.5	33	26	21	20	22
17*17,π	6	7	2	7	4
17*17,2.5	6	2	8	3	3
32*32,π	2	5	2	2	2
32*32,2.5	3	3	2	2	1

(b) TSL

Omega / Sigma	P	1.5p	2p	2.5p	3p
9*9,p/8	82	66	50	45	42
9*9,p/16	79	61	49	41	39
9*9,p/32	80	58	46	40	40
17*17,p/8	80	62	53	38	32
17*17,p/16	80	57	37	23	19
17*17,p/32	78	55	32	20	16
33*33,p/8	81	65	57	47	36
33*33,p/16	81	60	45	34	24
33*33,p/32	82	58	42	28	21

(c) XW

We reduce the number of fiducial point by 12, 6 and increase it by 36. Images of size 128×128 have 24 fiducial points. We first tested performance of three different Gabor wavelets. The face recognition performance of the LW, TSL and XW is show

in Table 1. One can find that the performance of the LW method is influenced by σ and the size of convolution and when the value of σ is 2π and the convolution has a size of 33×33, the number of error is the smallest (i.e. zero). We found that the performance of the LW method is superior to those of the TSL method and the XW method.

The TSL method has one additional parameter κ than the LW that affects shape of Gabor wavelet. When the σ is above 2 and the convolution has a size of 33×33 regardless of κ, the performance of TSL is almost the same as the LW method. But, when the size of convolution is 9×9, the TSL method is much inferior to the LW method. The performance of the XW isn't acceptable in our application, though it shows the best performance when the σ is 3π and the size of convolution is 17×17.

We have synthesized eight images from one original FERET and IT Lab image (see Fig. 5) in order to observe the effect of bad illuminations. We only rearrange Yale dataset since Yale dataset includes bad illuminant images.

Fig. 5. Bad illumination synthesized test images from FERET fafb set

We selected the Malsberg's kernel from the above experiment. We explored the various orientation sets and the frequency sets using Genetic algorithm. The sizes of the kernel used here are 33x33 and 17x17. The value of σ is 2π, and the resolutions of the images used here are 128x128 and 64x64. The experimental result shows that we can achieve an optimal performance with 8 orientations and 8 frequencies using 128x128 images.

The FERET gallery images of 1209 people are used for registration and 1209 probe_fafb_expression images are used for test. Table 2 shows the performance of Malsberg's kernel and parameters. The experiment in order to be suitable in image used the vector, which is various. It will be able to predict the parameter, which is suitable in illumination image from test.

The test image used the Yale B, IT Lab, FERET database. The Yale B database has the various illumination environments. Illumination directions divided at 9 phases, as showed in Table 2. The situational knowledge base has the information that estimated illumination situation, and corresponding Gabor parameters. Input image include light acquired, and estimated illumination direction, and executed recognition make use of Gabor wavelet using parameter as illustrated in Table 2. The table is the recognition rate test against an illumination 9 phases. Gabor parameters are kernel size 33x33 and 17x17, orientation & frequency that has from 0 to 20. Figure 6 shows the results of recognition rate for situational awareness. Experiment result, when value of frequency and orientation has value between from 5 to 8, and recognition rate was improved.

Table 2. Adaptive Gabor parameter construction for each situation

	Malsberg's kernel 33*33		Malsberg's kernel 17*17	
	Frequency	Orientation	Frequency	Orientation
Original image set	8	8	5	8
Image set of situation 1	7	8	8	8
Image set of situation 2	8	5	7	6
Image set of situation 3	7	7	8	5
Image set of situation 4	5	8	5	8
Image set of situation 5	6	7	6	8
Image set of situation 6	5	8	8	7
Image set of situation 7	8	7	7	8
Image set of situation 8	8	8	8	8

Fig. 6. Comparison of recognition rates using fixed Gabor representation and that using adaptive Gabor representation for each illumination situation

Table 3. Comparison of recognition rates using fixed Gabor representation and that using adaptive Gabor representation for each database

	Recognition rate using fixed Gabor representation	Recognition rate using adaptive Gabor representation
IT Lab database	83.63%	88.74%
Yale database B	73.75%	89.10%
FERER database	55.98%	84.10%

Database that used in our experiment are FERET, IT Lab and Yale B. We used IT Lab database in an experiment using 5 different places, 9 different lightings from 100 persons and acquires 4,500 images, as shown in Figure 6. Yale B database is 135 images using illumination 9 phase from 15 persons. FERET gallery set is illumination of 9 phases in frontal face image of 1,200 persons and about 10,800 images are consisted. FERET gallery set consisted of frontal face image of 1200 persons. FERET image is not various lighting, inserted condition of light artificially, as show in Table 3. Our experiment has FERET image to have condition of light of 9 phases. So, used in an experiment composing about 10,800 images in each image. The recognition rates using adaptive Gabor representation show much superiority to those using fixed

Gabor representation for all three databases as expected (see Table 3). Table 3 shows the recognition rate of over maximum 28.12 % for face dataset using our methods.

7 Concluding Remarks

In this paper, we propose a situational awareness based adaptive Gabor representation for robust object recognition, especially for changing illumination environments. Even though much research has done for object recognition, it remains a difficult problem yet. Furthermore, most existing technologies are not sufficiently reliable under changing illumination environment. The representation of Gabor wavelet is optimized using the situational awareness method and Genetic algorithm. The proposed situational awareness architecture for face recognition adapts itself to varying illumination environments, and shows much robustness. The experiment shows very encouraging result, especially for changing illumination environments. One future research direction is the research on automatic Gabor parameter decision process for high performing face recognition system.

Acknowledgement

References

1. N. Gourier, D. Hall, J. L.: Crowley, Facial Features Detection Robust to Pose, Illumination and Identity. IEEE Transactions on Systems, Man and Cybernetics (2004)
2. Bass E.J., Zenyuh J.P., Small R.L., Fortin S.T.: A context-based approach to training situation awareness. Human Interaction with Complex Systems, 1996. HICS '96. Proceedings., Third Annual Symposium on (1996) 89- 95
3. Bossmaier, T.R.J.: Efficient image representation by Gabor functions - an information theory approach. J.J. Kulikowsji, C.M. Dicknson, and I.J. Murray(Eds.), Pergamon Press, Oxford, U.K (1989) 698-704.
4. Goldberg, D.: Genetic Algorithm in Search, Optimization, and Machine Learning. Addison-Wesley (1989)
5. Wiskott, L., Fellous, J.M., Krüger, N., von der Malsburg, C.: Face Recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7) (1997) 775-779.
6. Field, D.: Relations between the statistics of natural images and the response properties of cortical cells. J. Opt. Soc. Amer. A, 4(12) (1987) 2379-2394.
7. Jones, J., Palmer, L.: An evaluation of the two dimensional Gabor filter model of simple receptive fields in cat striate cortex. J. Neurophysiology (1987) 1233-1258.
8. Gonzalez, R. C., Woods, R. E.: Digital Image Processing. Addison-Wesley Publishing Company (2003)
9. Arya, S., Mount, D. M., Silverman, N. S. Netanyahu. R., Wu, A. Y.: An Optimal Algorithm for Approximate Nearest Neighbor Searching in Fixed Dimensions. Journal of ACM (1994) 1-31
10. Daugman, J.: Two dimensional spectral analysis of cortical receptive field profiles. Vision research 20 (1980) 847-856

11. Faugman, J.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimization by two-dimensional cortical filters. Journal Opt. Soc. Amer. 2(7)(1985) 675-676
12. Liu,C., Wechsler,H.: Evolutionary Pursuit and Its Application to Face recognition. IEEE Trans. on PAMI, vol. 22, no. 6 (2000) 570-582
13. Brunelli, R., Poggio. T.: Face Recognition: Features versus Templates, IEEE Transactions on PAMI, 15(10) (1993) 1042-1052
14. Georghiades,A.S., Belhumeur,P.N., Kriegman,D.J.: From Few to Many: Illumination Cone Models for face recognition under Variable Lighting and Pose. IEEE Trans. on PAMI, vol. 23 no. 6 (2001) 643-660
15. Haiyuan Wu, Yoshida, Y., Shioyama, T.: Optimal Gabor filters for high speed face identification. Pattern Recognition, 2002. Proceedings. 16th International Conference on, Volume: 1, 11-15 (2002) 107 - 110
16. Yavnai A.: Context recognition and situation assessment in intelligent autonomous systems, Intelligent Control, 1993., Proceedings of the 1993 IEEE International Symposium on (1993) 394-399
17. Maestre R., Kurdahi F.J., Fernandez M., Hermida R., Bagherzadeh N., Singh H.m.: Optimal vs. heuristic approaches to context scheduling for multi-context reconfigurable architectures. Computer Design, 2000. Proceedings. 2000 International Conference on (2000) 575-576
18. Yau, S.S., Chandrasekar, D., Dazhi Huang,: An adaptive, lightweight and energy-efficient context discovery protocol for ubiquitous computing environments. Distributed Computing Systems 2004. FTDCS 2004. Proceedings. 10th IEEE International Workshop on Future Trends of (2004) 261 - 267

An Evolvable Hardware System
Under Uneven Environment

In Ja Jeon[1], Phill Kyu Rhee[1], and Hanho Lee[2]

[1] School of Computer Science & Engineering, Inha University
253 Yonghyun-Dong, Nam-Gu, Incheon, Korea
`juninja@im.inha.ac.kr, pkrhee@inha.ac.kr`
[2] School of Information & Communication Engineering, Inha University
`hhlee@inha.ac.kr`

Abstract. This paper proposes an evolvable hardware system with capability of evolution under uneven image environment, which is implemented on reconfigurable field programmable gate array (FPGA) platform with ARM core and genetic algorithm processor (GAP). Parallel genetic algorithm based reconfigurable architecture system evolves image filter blocks to explore optimal configuration of filter combination, associated parameters, and structure of feature space adaptively to uneven illumination and noisy environments for an adaptive image processing. The proposed evolvable hardware system for image processing consists of the reconfigurable hardware module and the evolvable software module, which are implemented using SoC platform board with the Xilinx Virtex2 FPGA, the ARM core and the GAP. The experiment result shows that images affected by various environment changes are enhanced for various illumination and salt & pepper noise image environments.

1 Introduction

Currently, many image processing applications are implemented on general-purpose processors such as Pentiums. Effective processing of images and especially sequences of images in space requires computing architectures that are less complicated, highly flexible and more cost-effective. Recently the complex and fast computation is performed by dedicated hardware instead of software in digital computer because hardware can operate in parallel. The concept of a reconfigurable hardware and evolvable hardware has been studied actively [1, 3, 4]. The reconfigurable architecture can be classified into gate-level and functional-level ones [5, 6]. The evolvable hardware architecture is a functional evolvable module, which can be implemented by reconfigurable field programmable gate arrays (FPGAs). The configurations that make the device output responses optimal to the desired response are combined to make better configurations until an optimal architecture is achieved [8]. The evolvable hardware architecture continues to reconfigure itself in order to achieve a better performance.

This paper presents evolvable hardware system that is effective for implementing adaptive image processing. The proposed system consists of the reconfigurable hardware module and the evolvable software module. The reconfigurable hardware module process the median, histogram equalization, contrast stretching and illumination compensation algorithm, which are implemented on Xilinx Virtex2 FPGA. The evolvable software module consists of genetic algorithm and feature space search block, which are implemented by genetic algorithm processor (GAP) and ARM core, respectively.

Section 2 describes the evolvable hardware system for adaptive image processing applications. The experimental results are given in Section 3. In Section 4, the conclusions are given.

2   Evolvable Hardware System for Image Processing

The proposed evolvable hardware system consists of the reconfigurable hardware module for the image preprocessing algorithms and the evolvable software module for feature space of Gabor representation and fitness evaluation, as shown in Fig. 1.
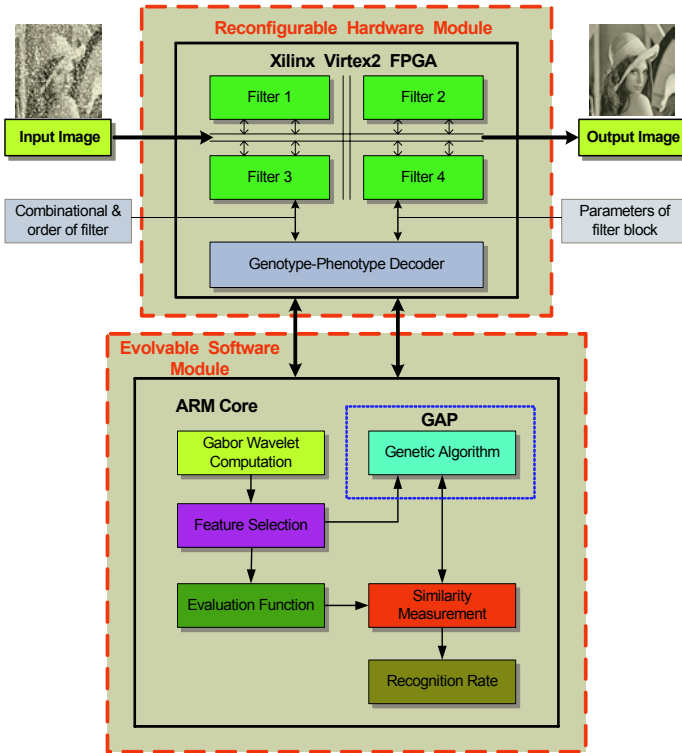
Fig. 1. The block diagram of evolvable hardware system

2.1   Reconfigurable Hardware Module

The reconfigurable hardware module processes the image preprocessing algorithms, which are the median, the histogram equalization, the contrast stretching and the illumination compensation algorithm for object recognition. This module consists of the genotype-phenotype decoder and 4 types of image filters, and it is implemented on Xilinx Virtex2 FPGA. The optimal image filter function can be searched and selected using genetic algorithm running on the genetic algorithm processor (GAP) (see Fig. 2) and the genotype-phenotype decoding.
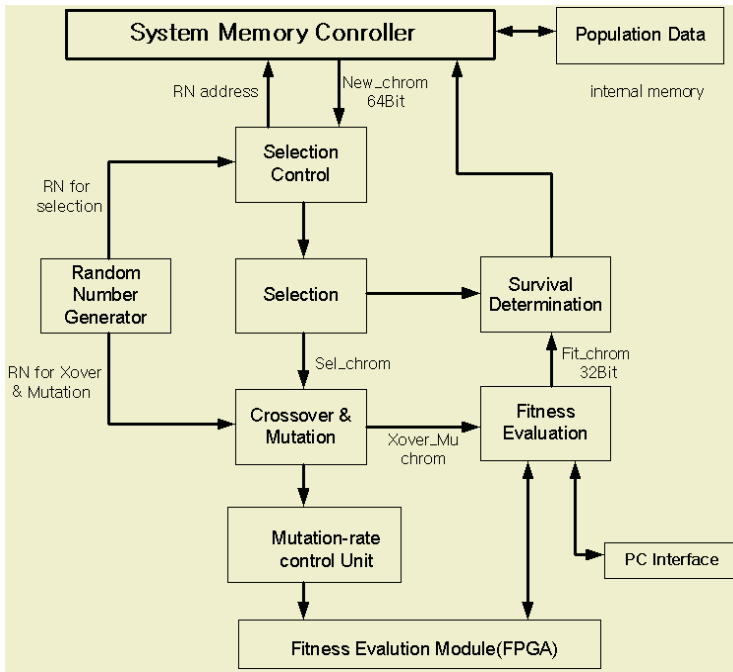
Fig. 2. The structure of the GAP

The input of reconfigurable hardware module is illumination image, which is suitable image to achieve processing. Used image processing algorithms transform image lineally. There is no dependency between the used algorithms. Therefore, the image filter function can be selected using the GAP about used order or used existence, and nonexistence of image that is most suitable to recognize. Used algorithm's order and existence and nonexistence are selected using the GA method. Composite processed image after image preprocessing is accomplished. The image filters are operated in parallel using multiple memory blocks in order to decrease processing time in hardware.

Fig. 3 shows the hardware structure of the reconfigurable image filter banks. The four types of image preprocessing filters are processed in parallel and synchronously. When this parallel processing method is applied to filters in hardware, the processing speed is increased more than about three times. The gene of GAP includes the sequent order and parameter values of filters. Each filter sets their parameter values and images are filtered in the order. The filtered images are sent to the fitness evaluating part. Selected filter gets the parameter values and image. The four types of image filter banks are as follows.

Median Filter
Median filter is useful to remove impulse noise. The filtered gray levels are calculated by taking median of gray level of each pixel in mask set. The six types of median filter masks are horizontal, vertical, cross, block, diamond masks, and genetic algorithms select optimal mask set.
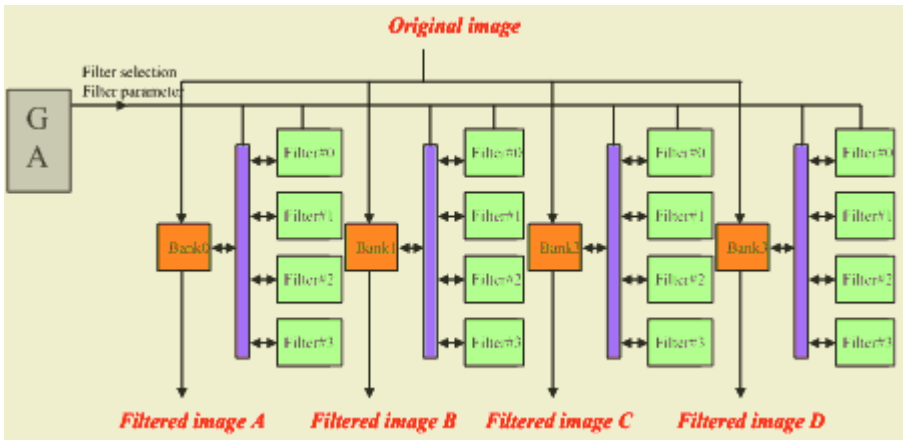
Fig. 3. The structure of the reconfigurable image filter banks

Histogram Equalization Filter
To improve contrast of image, histogram equalization filter is used. If the distribution of gray level was biased to one direction or scaled value was not uniformly distributed, histogram equalization is a good solution for image enhancement.

Contrast Stretching Filter
The image process is viewed as a product of image illumination and scene reflectance. This is exactly what contrast stretching filtering does to enhance details in an image. That is, the contrast-stretching filter reduces brightness and emphasizes contrast in order to improve a reflectance effect and decrease a lighting effect.

Illumination Compensation Filter
Illumination compensation filter is the high pass filter using the local brightness, which means the average of difference between the brightness of central pixel point and the brightness of total window [7]. Brightness value is applied to an image, compensated image become totally dark compared to the pixel value of original image and background illumination modeling function. In order to enhance the dark image, Brightness value compute added to the average of total original image pixel. The shadow part in an image comes to blur and stain in the compensated image computed. So illumination compensation is performed by a ratio of original image to background illumination modeling function. The final compensated image is obtained by multiplying the ratio and weight value.

2.2 Evolvable Software Module

The role of evolvable software module controls a Gabor feature vector of a face for achieving optimal performance of recognition system in varying environment. The evolvable software module consists of a genetic algorithm block and a feature space search block for feature space of Gabor representation and fitness evaluation, which are implemented by software and processed by GAP and ARM processor. The ge-

netic algorithm block evolves the interconnection structure and the parameters of the reconfigurable hardware module in order to remove random noises in the image data. The feature space search block explores the feature space of Gabor representation to find optimal feature representation.

Evolvable computing initiates a population, which is a set of members [6]. Each member is described by a vector, which is called a chromosome. The fitness of each member is evaluated, and only a portion of the population is selected as the population of next generation. It is said survival of the fittest with analogy with biological evolution. This process is repeated until a best member, i.e., a best face recognition architecture is reached, or a generation exceeds a desired criterion value.

Genetic Algorithm Block

GA block has two functionality. The first module of GA block sends the information of the use existence, nonexistence, order, parameter value and window size of median filter to image preprocessing algorithm. The second module of GAP generates the possible combination of fiducial points, Gabor feature vectors and the optimality of the chromosome, which is defined by classification accuracy and generalization capability. The total Gabor feature vector for all fiducial points, V is evolved from a larger vector set defined as follows:

$$V = (F^{(e)}(\vec{x_1}) F^{(e)}(\vec{x_2}),,,, F^{(e)}(\vec{x_n})) \tag{1}$$

$(\vec{x_1}), (\vec{x_2}),,,, (\vec{x_n})$ are fiducial points. The contribution of fiducial points in n-dimensional is space by a set of weights in the range of (0.0, 1.0). As it searches the geno-space, the GA makes its choices via genetic operators as a function of probability distribution driven by fitness function. The genetic operators used here are selection, crossover, and mutation [7].

The Fitness of Genetic Algorithm

The evolvable hardware system needs a salient fitness function to evaluate current population and choose offspring for the next generation. Evolution or adaptation is guided by the fitness function defined for the system performance and the class scattering criterion. The system performance denotes the correctness that the evolvable adaptation has achieved so far, and the class scattering indicates the expected fitness on future generations. The evolvable software module derives the classifier being balanced between successful recognition and generalization capabilities.

3   Experimental Results

The proposed evolvable hardware system using FPGA platform with ARM processor has been tested to adapt the system for the image processing under uneven illumination and noisy. IT Lab., FERET, and Yale databases are used for the performance evaluation of the proposed system. Gray level face images are synthesized by adding illumination and noise factors to original images in order to make the quality of face image more uneven. These databases are used to evaluate the fitness of chromosome. The face recognition block learns changing environments and the evolvable hardware computing image processing is adapted by restructuring its structure. The total experiment procedure is given in the followings:

1. Perform preprocessing using the evolvable hardware, and derives Gabor representation $F(x_i)$ for each fiducial point, and normalizes it.
2. Concatenate the Gabor representations for fiducial points to generate total Gabor vector $V = (F^{(e)}(\vec{x_1})\ F^{(e)}(\vec{x_2}),,,,F^{(e)}(\vec{x_n}))$.
3. Begin the evolvable hardware optimization until a criterion is met, where the criterion is the fitness does not improve anymore or the predefined maximum trial limitation is encountered. Generate new Gabor feature space.
4. Evaluate the fitness function $\eta(v) = \lambda_1\eta_s(v) + \lambda_2\eta_g(v)$ for the newly derived Gabor feature space. Search for the subset of Gabor representations in the Gabor feature space that maximize the fitness function and keep those as the best chromosomes.
5. Apply genetic operators to generate new population of Gabor feature space and perform recognition using the image preprocessing hardware and Gabor vector representation evolved from Step 3.

Preprocessing is performed for providing nice quality images as much as possible using image filtering techniques discussed in the previous session. A-NN algorithm [8] is employed for face identification. In the A-NN, the reference face points in the Gabor space is preprocessed into a hierarchical decomposition of the space, called a balanced box decomposition tree. The height of the tree is $O(log\ n)$, and subdivided space is $O(d)$ complexity, where d is the dimension of the Gabor space. The A-NN begins by locating the leaf containing the query face points in $O(log\ n)$ time, and enumerates the leaf in increasing order of distance from the query point. The face points associated the leaf is computed, and the closest point found so far is tracked, and the decision of identification is made.

For experiments, the number of initial population is set to 32 and each chromosome is evolved with crossover of 0.8 and mutation of 0.03. Fig. 4 shows the enhanced result for left/right-side and full-bright/darkness illumination condition. The sine shaped illumination is synthesized with the half region of input facial image. After 2,000 generations, optimal filter combination is obtained in the order of contrast stretching filter, median filter, illumination compensation filter and histogram equalization filter. Fig. 4 shows that synthesized image and the image impaired by illumination is improved by the optimal filter system. IT Lab. test data of 400 facial images composed of 39 people is used in order to test the optimal filter system. As a result, recognition rate increased 68.01% after evolution. Figure 3 and 4 show that the image corrupted by noise is improved by the optimal filter system. As a testing result, recognition rate increased 65.4% after evolution.

Fig. 4. It was preprocessed image with uneven illumination by the proposed evolvable hardware system
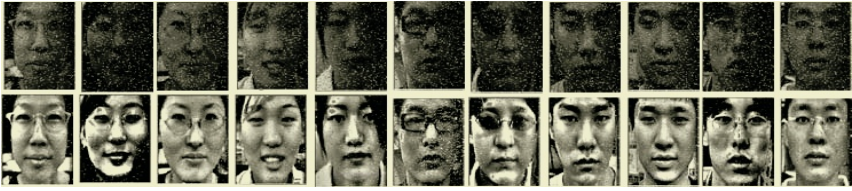
Fig. 5. It was preprocessed image with uneven illumination and noisy by the proposed evolvable hardware system
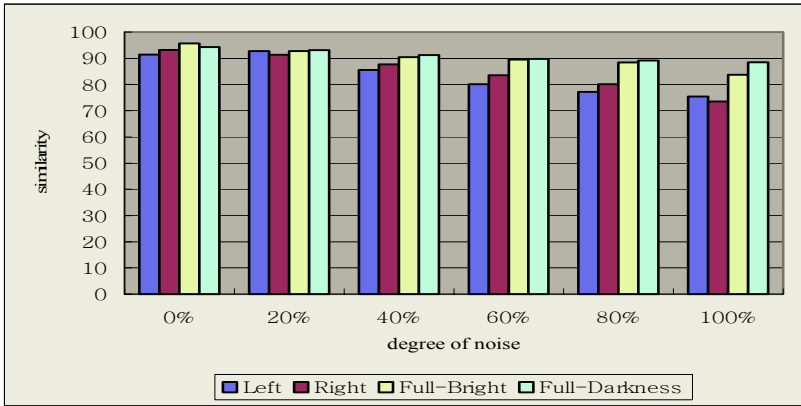
Fig. 6. The results of similarity using varying illumination and noise condition of IM Lab DB

Table 1. Comparison of the performance between the image processing evolvable hardware and traditional image processing methods

	Histogram Equaliza-tion filter only	Median filter only	Illumina-tion com-pensation filter only	Hommor-phic filter only	Proposed evolvable image processing hardware method
IT Lab. database	87.8%	90.8%	86.3%	89.6%	92.3%
FERET database	91.3%	89.5%	88.5%	90.4%	92.5%
Yale data-base	84.5%	90.4%	77.2%	87.5%	91.1%

Fig. 5 shows the enhanced result of combination with noise and illumination. The random noise is synthesized to the universal region of 0% ~ 100% of an image. Optimal filter combination is obtained in the order of contrast stretching filter, median filter and histogram equalization filter. The results of experiment are given in the Fig. 6.

Table 1 shows the experiment result in which we used 4 types of image preprocessing algorithms to compare the performance. As seen in Table 1, we know that

similarity rate is improved when used preprocessing algorithm used preprocessing algorithm variously according to the input image. Also, when the image preprocessing algorithms are implemented on FPGA instead of software method, the execution time was much improved. In the experiments, the recognition rate improved from 0.7% to 13.9%, after use filters to adaptive, the recognition rate is changes.

4   Conclusions

In this paper, we proposed an evolvable hardware system, which is implemented on Xilinx FPGA, ARM processor and GAP, for adaptive image processing applied to the face recognition in object recognition. Most existing technologies are not sufficiently reliable under changing illumination and various noises. The proposed evolvable hardware system performs well especially in changing illumination and noisy environments, since it can adapt itself to external environment.

The face recognition performs by Gabor wavelet, which is intrinsically robust to uneven environments. The face recognition is optimized using evolvable approach. The proposed system for face recognition adapts itself to varying illumination and noisy environments, and shows much robustness especially for changing environments of illumination and noisy.

References

1. T. Higuchi, M. Iwata, W. Liu.: Evolvable Systems: From Biology to Hardware. Tsukuba, Springer, (1996)
2. D.E. Goldberg.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison Wesley. (1989)
3. Adrian Stoica, Ricardo Zebulum, Didier Keymeulen, Raoul Tawel,Taher Daud, Anil Thakoor.: Reconfigurable VLSI Architectures for Evolvable Hardware: From Experimental Field Programming Transistor Arrays to Evolution-Oriented Chip, IEEE Trans. on VLSI Systems, Vol. 9, No. 1 (2001) 227-231
4. N. Tsuda.: Fault-Tolerant Processor Arrays Using Additional Bypass Linking Allocated by Graph-Node Coloring, IEEE Trans. Computers, Vol. 49, No. 5 (2000) 431-442
5. A. Marshall, T. Stansfield, I. Kostarnov.: A Reconfigurable Arithmetic Array for Multimedia Applications. ACM/SIGDA International Symposium on FPGAs (1999) 135-143
6. K. K. Bondalapati.: Modeling and Mapping for Dynamically Reconfigurable Hybrid Architectures. PhD thesis, University of Southern California (2001)
7. D. Goldberg.: Genetic Algorithm in Search, Optimization, and Machine Learning. Addison-Wesley, (1989)
8. J. Faugman.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimization by two-dimensional cortical filters. Jounal Opt. Soc. Amer. 2(7) (1985) 675-676
9. L. Wiskott, J.-M. Fellous, N. Kuiger, C. von der Malsburg.: Face Recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19 (1997) 775-779
10. Bossmaier, T.R.J.: Efficient image representation by Gabor functions - an information theory approach. in J.J. Kulikowsji, C.M. Dicknson, and I.J. Murray(Eds.), Pergamon Press, Oxford, U.K.(2001) 698-704

An Efficient Face Location
Using Integrated Feature Space

Mi Young Nam and Phill Kyu Rhee

Dept. of Computer Science & Engineering, Inha University
Yong-Hyun Dong, Incheon, Korea
rera@im.inha.ac.kr, pkrhee@inha.ac.kr

Abstract. We propose a method for an efficient frontal face detection using skin color, integrated feature space, and post processing. The proposed method reduces the search space by facial color information and detects face candidate windows by integrated feature space. The integrated feature space consists of intensity and texture information. Multiple Bayesian classifiers are employed for selection of face candidate windows on integrated feature space. And we use face and face-like nonface samples to training these Bayesian classifiers. Finally, face regions of the detected candidates are selected by merging and filtering post processing.

1   Introduction

Detecting human face in image frames is an important task in many computer vision applications. Face detection, the first step of an automated face recognition system, determines the location and size of each human face from an input image. Closely related problems are face verification and identification. Face detection is one of the hardest problem in the research area of face recognition, and an important task in many computer vision applications such as human-computer interface, smart environment, ubiquitous computing, multimedia retrieval, etc.

Object detection using a static image can be used in unconstrained environments with complex background [1-6]. Object detection systems from still image can be divided into three major categories [7]: model-based method, image invariant method, and appearance-based learning method. In the model-based method, an object model is defined, and this model is matched to the image. The second one, image invariant method is based on a matching set of image pattern relationships. In the last one, the features of objects of labeled categorized object examples are used. Recently, appearance-based approach has been employed successfully in the computer vision areas [8]. We concentrate on the frontal face detection in this paper.

The major contributions of this paper are 1) integrated feature analysis using hierarchical intensity and texture information, 2) statistical modeling of face and face-like nonface, and 3) efficient classification using multiple Bayesian classifiers.

The outline of this paper is as follows. In section 2, we present the architecture of the proposed face detection system. In section 3, we describe the method that reduces search spaces using skin color. The integrated feature space is discussed in section 4. Bayesian classifier and modeling of face and face-like nonface is explained in section 5. In section 6, we present the post processing that merges and eliminate overlapped face candidate regions. We give experimental results in section 7. Finally, we give concluding remarks in section 8.

2   The Proposed Face Detection Scheme

The proposed scheme consists of the search space reducing module using skin color, the integrated feature extraction module, and multiple Bayesian classifiers for selecting face candidate windows. The final decision of face candidate window is done by post processing. The outline of the proposed method is shown in Fig.1.

Search space is reduced by finding regions that has face color. The remaining regions are searched by multiple Bayesian classifiers using integrated feature space to determine whether the 16×16 sub-window is face or nonface. These classifiers are constructed in cascade form of two steps. The 54-dimensional feature from intensity using sub-regions is used as feature in first step. Haar wavelet transform is used in second step. These two classifiers produce different false alarms when they work alone. Therefore cascade classifier reduces false alarms when each result is merged. Structure of cascade classifier is shown in middle block of Fig.1. Once a sub-window is determined to nonface by first classifier, it don't be tested by second classifier. First, search space can be reduced by skin color module. The remaining regions are searched by multiple Bayesian classifiers using integrated feature space. Final results is produced by post processing
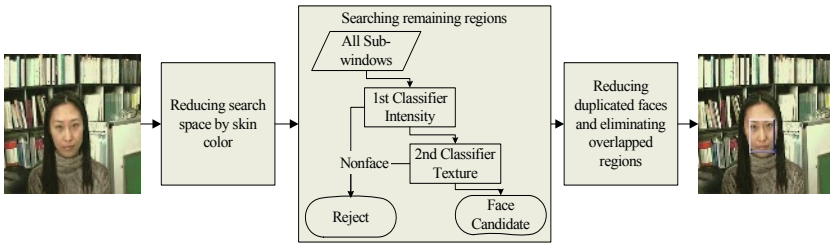
Fig. 1. The proposed scheme used for face detection. Input image has RGB color scale. Reducing Search Space by Skin Color

Skin color provides good information for extracting area of the face [9]. We used YCrCb color model Fig.2 shows the distribution of face color in that color model.
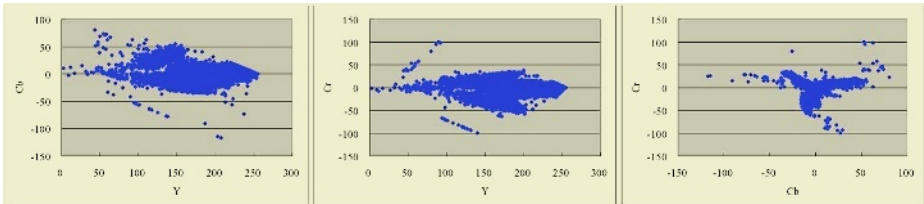
Fig. 2. Ace color distributions in YCrCb color model of 16×16 size face images

Search space reduction using facial color is performed as follows. We define face color as the value over a threshold, where the threshold value is determined empirically. Initially, seed appearance samples of a face color area are manually gathered and used for modeling face color. Fig.3 shows a face candidate region and an extracted face window: outer box is a face candidate region and inner box is an extracted face window.

Fig. 3. Face color region that is reduced by skin color from input image and face window that is detected in that region

3   Integrated Feature Space

Feature extraction method has much affect in face detection performance. The feature extraction method that was optimized by an experiment, as a result, raises good performance. Therefore, we can improve performance through efficient feature extraction and dimension decrease. We use two types of feature extraction method. One uses the 54 sub-regions. Other is the method using Haar wavelet transform.

3.1   Feature from Intensity

We extracted 54-dimensional feature from facial images using sub-regions. Face in search window is encoded to 54-dimensional vector which including hierarchical intensity information (see Fig.4).
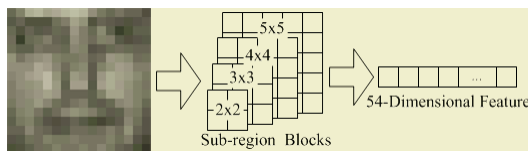
Fig. 4. Face window and its 54-dimensional feature

3.2   Feature from Texture

We can model a face in frequency domain using Haar wavelet transform. It is organize the image into sub-bands that are localized in orientation and frequency. In each sub-bands, each coefficient is spatially localized. We use a wavelet transform based on 3 level decomposition producing 10 sub-bands as shown below in Fig.5. Generally, low frequency component has more discriminate power than higher. Also too high component has some noisy information. Then we use 3 level decomposition of Haar wavelet transform.
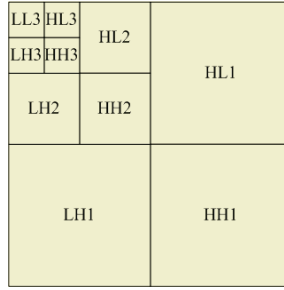
Fig. 5. 3 level Haar wavelet transform of the 16×16 image

4   Bayesian Classifier

Bayesian formula is

$$P(\omega_i \mid \mathbf{x}) = \frac{p(\mathbf{x} \mid \omega_i) P(\omega_i)}{p(\mathbf{x})} \tag{1}$$

$P(\omega_i|\mathbf{x})$, $p(\mathbf{x}|\omega_i)$, $P(\omega_i)$, and $p(\mathbf{x})$ is posterior, likelihood, prior, and evidence respectively. Evidence is not independent to pattern $\mathbf{x}$. Hence we can ignore it. And decision is not changed although we use natural logarithm at posterior. We can classify input pattern $\mathbf{x}$ using below two discriminant functions. If $g_f(\mathbf{x})$ is greater than $g_n(\mathbf{x})$, we can decide input pattern as face. Otherwise, it is nonface.

$$\begin{aligned} g_f(\mathbf{x}) &= \ln p(\mathbf{x} \mid \omega_i) + \ln P(\omega_i) \\ g_n(\mathbf{x}) &= \ln p(\mathbf{x} \mid \omega_n) + \ln P(\omega_n) \end{aligned} \tag{2}$$

But we can use one discriminant function, $g(\mathbf{x})$ as (3) since we try to classify a pattern into two categories. In this case, if $g(\mathbf{x})$ is greater than 0, we can decide input pattern as face. Otherwise, it is nonface.

$$g(\mathbf{x}) = g_f(\mathbf{x}) - g_n(\mathbf{x}) \tag{3}$$

$$\mathbf{x} \in \begin{cases} \omega_f & \text{if } g(\mathbf{x}) > 0 \\ \omega_n & \text{otherwise} \end{cases} \tag{4}$$

4.1   Modeling of Face and Nonface

We assume that distributions of face and nonface samples form the shape of Gaussian distribution. The likelihood of face and nonface can be modeled as multivariate normal distribution as (5).

$$p(\mathbf{x} \mid \omega_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left[-\frac{1}{2} (\mathbf{x} - \mu_i)^t \Sigma_i^{-1} (\mathbf{x} - \mu_i) \right] \tag{5}$$

$\mathbf{x}$ is the d-dimensional feature, Σ_i is the covariance matrix of the samples, and μ_i is the d-dimensional mean vector. Then two discriminant functions has the following form:

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^t \, \Sigma_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| + \ln P(\omega_i) \; . \tag{6}$$

First term of the formula in (6) is the squared Mahalanobis distance. And rest terms are not dependent of input pattern $\mathbf{x}$. Therefore we can use variable r and α to simplify the formula. r is the Mahalanobis distances and α is the constant.

$$g_i(\mathbf{x}) = r_i^2 + \alpha_i \tag{7}$$

$$\begin{aligned} r_i^2 &= (\mathbf{x} - \boldsymbol{\mu}_i)^t \, \Sigma_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) \\ \alpha_i &= d \ln 2\pi + \ln|\Sigma_i| - 2\ln P(\omega_i) \end{aligned} \tag{8}$$

Therefore we can classify input pattern into face and nonface using two squared Mahalanobis distances, r_f^2 and r_n^2. First term of each α_f and α_n, is the same and we can ignore.

$$\mathbf{x} \in \begin{cases} \omega_f & \text{if } r_f^2 < r_n^2 + \tau \\ \omega_n & \text{otherwise} \end{cases} \tag{9}$$

$$\tau = -\ln|\Sigma_f| + \ln|\Sigma_l| + 2\ln\frac{P(\omega_f)}{P(\omega_l)} \tag{10}$$

τ must be decided empirically. We cannot decide $P(\omega_f)$ and $P(\omega_n)$ because nonface is the rest of the world except faces. We cannot decide how many faces and nonfaces are in the world. If τ is increased, then false alarm is decreased and also face detection rate is decreased. Similarly, τ is decreased, false alarm is increased and face detection rate is also increased.

4.2   Choosing Face-Like Nofaces

We choose face-like nonface samples from 23 natural images. The method choosing face-like nonface samples is based on squared Mahalanobis distance from face class. We briefly describe it below:

1. Construct covariance matrix, Σ_f from face training set, $\mathbf{T}_f$,
2. Calculate squared Mahalanobis distance, $d_{f,i}$ from face class to each $f_i \in \mathbf{T}_f$,
3. Set max$\{ d_{f,i} \,|\, 1 \leq i \leq f \}$ to θ,
4. Calculate Mahalanobis distance, $d_{f,j}$ from nonface class to each $n_j \in \mathbf{T}_n$,
5. Collect n_j that $d_{f,j} < \theta$ under face-like nonface set, $\mathbf{T}_l$,
6. And construct covariance matrix, Σ_n from $\mathbf{T}_l$.

Therefore we can decide which category the input pattern is as follow method:

$$\mathbf{x} \in \begin{cases} \omega_f & \text{if } \left(r_f^2 < \theta\right) \text{and} \left(r_f^2 < r_l^2 + \tau\right) \\ \omega_n & \text{otherwise} \end{cases} \; . \tag{11}$$

5   Merging and Filtering Process of Candidate Windows

Merging heuristic that removes the overlapped face candidate windows is powerful process to remove most false detections and reduce multiple detections of a single face to one [10]. The merging heuristic used in this paper is as follows:

1. Count the neighborhoods of each face candidate windows,
2. Classify the candidate as a face if the number of its neighborhoods is above a threshold,
3. And eliminate overlapped face windows as the number of neighborhoods. We can preserver the location with more number of detection, and remove locations with less detections.

This pose processing can produce the results as Fig.6.

(a)   (b)   (c)
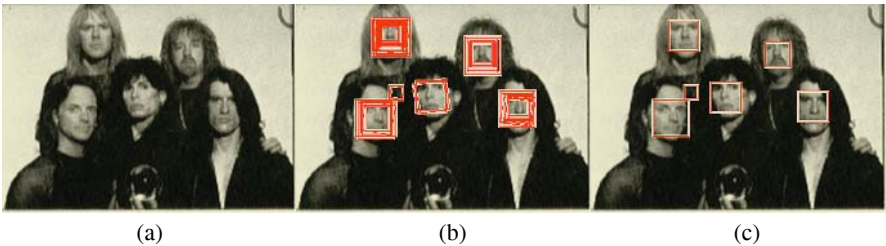
Fig. 6. Post processing. (a) Input image. (b) Before post processing. (c) After post processing

6   Experiment

We used 3,816 FERET frontal face images as face data and 23 natural images as non-face data for training data. Because we include mirror images with the face data, total face data consist of 7,632 frontal face images.

Fig. 7. Some images in FERET database

In order to normalize size and position of eyes, each face image was rotated and scaled to 16×16 size as shown in Fig.8.

Fig. 8. Face samples that are normalized in 16×16 scale

Nonface data was gathered from 23 natural images. Because the face-like nonface samples were chosen by each Bayesian classifier with two different types of feature extraction methods. Thus each classifier uses different number of face-like nonface images.

Fig. 9. Natural images that are not contained any face image

We performed the experiments on two test sets constituted from images of MIT-CMU frontal face images and captured images. First set consists of 130 images include 507 faces. 117 images contain frontal faces and remains don't. Second set consists of 2,539 images that were captured in various environments. We show the experiment result in Table 1 and Table 2.

Table 1. Frontal face detection result using single step method

Test Set	Face Detected	False Detections
MIT-CMU Test Set	283	150
Our lab test set	2299	43

Table 2. Frontal face detection results of MIT-CMU test set and our lab test set using the proposed method

Test Set	Face Detected	False Detections
MIT-CMU Test Set	383	50
Captured Images	2321	21
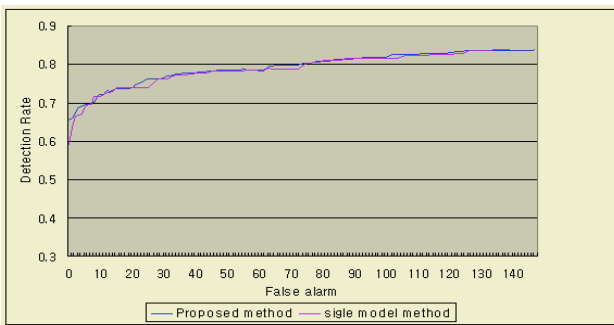
Fig. 10. Face detection ROC curve

The first set has the some faces that are not frontal. Thus the result in this set is not good. But in the second set, the images have only frontal faces and our skin color module worked. The result in this set is very good. We show the detected images in Fig.11.

Fig. 11. Detected images of MIT-CMU data set

7   Concluding Remarks

In this paper, we propose the efficient face detection method. It consists of skin color module, multiple Bayesian classifiers using integrated feature space, and post processing. In color images such that from cameras skin color module can reduce search space successfully. Then the remaining space is searched by multiple Bayesian classifiers using features from intensity and texture information and it results face candidate regions. These are decided through post processing, and we can get satisfied final result. This proposed method can detect 91.41% of faces in captured images. And 75.54% of face in MIT-CMU data set that include gray scale images is achieved without skin color module.

There are some works that can improve the proposed method. The second feature of our method, Haar wavelet transformation must be improved. If its dimensionality can be reduced, our method will spend less search time.

Acknowledgement

References

1. Papageorgiou, C., Oren, M., and Poggio, T.: A General Framework for Object Detection. In Proceedings of the International Conference on Computer Vision, Bombay, India, January (1998)
2. C. Liu.: A Bayesian Discriminating Features Method for Face Detection. IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 25. (2003) 725-740

3. T. V. Pham, et. Al.,: Face detection by aggregated Bayesian network classifiers. Pattern Recognition Letters 23, (2002) 451-461
4. H.Schneiderman and T. Kanade.: Object Detection Using the Statistics of Parts. Int'l J. Computer Vision, Vol. 56. no. 3. (2004) 151- 177
5. B.K.L. Erik Hjelmas, "*Face Detection: A Survey*," Computer Vision and Image Understanding, vol. 3. (2001) 236-274
6. Wren, C., Azarbayejani, A., Darrell, T., and Pentland, A.: Pfinder: Real-Time Tracking of the Human Body. Technical Report 353, Media Laboratory, Massachusetts Institute of Technology, (1995)
7. Mohan, A., Papageorgiou, C., and Poggio, T.: Example-Based Object Detection in Images by Components. IEEE Transactions on PAMI, Vol. 23. No. 4. April (2001) 349-361
8. Kasuba, T.: Simplified Fuzzy ARTMAP. AI Expert, November 19-25, (1993)
9. Yang, J. and Waibel, A.: A Real-Time Face Tracker. Proceeding of the Third Workshop on Applications of Computer Vision, (1996) 142-147
10. Rowley, H.A., Baluja, S., and Kanade, T.: Neural Network-Based Face Detection. IEEE Transaction on PAMI, Vol. 20, No. 1, January (1998) 23-38

Fuzzy Predictive Preferential Dropping for Active Queue Management

Lichang Che and Bin Qiu

School of Computer Science and Software Engineering
Monash University Wellington Road, Vic 3800, Australia
{carski,bq}@csse.monash.edu.au

Abstract. This paper proposes an Active Queue Management(AQM) scheme referred to as Fuzzy Predictive Preferential Dropping (FPPD). Two contributions are made in this work. Firstly, a fuzzy predictor is employed to improve the accuracy of traffic prediction. Secondly, a novel congestion index, predicted traffic intensity from fast flows, is used to derive packet dropping probability for AQM. The FPPD safely detects real congestion caused by large flows while leaving transient traffic burst from short-lived flows alone. Furthermore, a preferential dropping mechanism is adopted to treat packets from long-term fast flows and short-lived flows differently. Simulations show that the proposed FPPD reduces packet drop ratio and utilizes link bandwidth more efficiently than other AQM schemes. It also improves the quality of service perceived by web users.

1 Introduction

The research community has been advocating the deployment of active queue management mechanisms at routers since the proposal of Random Early Detection (RED) [1]. It is known that the effectiveness of an AQM scheme depends heavily on how well the scheme detects congestion and computes the random dropping probability accordingly. The RED and its variants use average queue length to compute random dropping probability. Although average queue length based schemes can mask transient traffic burst, it is argued that using queue length solely for both congestion and performance indices makes it hard to achieve high link utilization and low packet loss ratio simultaneously [2]. This has motivated other AQM schemes to introduce additional measures such as link utilization [3] and input traffic rate [4] to detect congestion and compute random dropping probability.

Recently AQM schemes based on traffic prediction have received much interest because predictive AQMs generally yield small queue variation and high link utilization. This work follows the practice of predictive queue management and makes the following two contributions. A) A fuzzy inference system is proposed for the prediction of future traffic on the basis of our previous work. B) A novel congestion index, predicted traffic intensity from long-term fast (LTF) flows, is used to compute packet dropping probability. The above two points consists of

the core idea of the proposed Fuzzy Predictive Preferential Dropping (FPPD) scheme. The FPPD scheme aims to reduce packet loss ratio by masking transient traffic change, while still timely reacting to real congestion caused by LTF flows. Simulation studies carried out with NS-2 [5] demonstrate that the FPPD achieves its design goal by outperforming other AQM schemes with lower packet drop ratio and higher link utilization. It also improves the quality of service (QoS) for web users by reducing complete time for web transactions.

The next section explains why predicted traffic from LTF flows should be used for congestion detection and control. Section 3 introduces the FPPD scheme in detail. The simulation results and related discussion are presented in Section 4. Section 5 concludes the paper and gives our future work.

2   Motivation

To realize the objectives of AQM such as maximizing link utilization and reducing delay and delay jitter, it is necessary to stabilized packet queue length around a small target length q_t. However, for a control loop with variable and generally large delay like the Internet, it is extremely difficult to stabilize output to a set point (target queue length in this case) without predictive control. Predictive queue management is required for stabilizing queue length in the sense that by predicting the traffic intensity in the future control horizon(s), early adjustment can be made to prevent large queue variation. Supposing that packet drop ratio is adjusted in every control interval (denoted as T seconds hereafter) and the current queue length is q. If the arriving traffic in the next control interval is T_a and the capacity of the link is C, the packet or byte dropping ratio during the next interval is [6]

$$p_d = \frac{q + T_a - C - q_t}{T_a} \tag{1}$$

to minimize the queue variation.

The scheme would work well for stabilizing queue length if the traffic on the Internet was all controllable. However, according to the heavy-tailed distribution of the Internet traffic, there exists a vast number of short-lived flows on the Internet, which mostly last for only a few seconds and are too short to control. It has been reported that transient traffic change caused by short-lived or slow flows (SLS) has adverse impact on the stabilizing queue length [1]. This work thus aims to compute random dropping probability on the basis of traffic from LTF flows. In this way, the transient traffic noise caused by SLS flows diminishes and the impact of the noise on queue performance can be reduced.

There is one more reason to compute random dropping probability using traffic from LTF flows. Because LTF flows are the main contributors of Internet congestion [7], more weight should be put on the traffic of LTF flow instead of treating all the traffic equally. This can be further justified by the observation that even the current queue length is not low, as long as there is no much traffic coming from LTF flows, the network congestion is not severe and packet

dropping ratio should not be increased much. Otherwise, it may lead to packet over-dropping but only marginal performance gain in congestion control.

Based on the above understanding, this paper initiates to derive packet drop probability using predicted traffic intensity from LTF flows to mask traffic noise. A Fuzzy Predictive Preferential Dropping (FPPD) scheme is then proposed to treat SLS flows more favorably. Next, the FPPD scheme is presented in detail.

3   Fuzzy Predictive Preferential Dropping

The FPPD scheme consists of the following three components.

- Collection of traffic from LTF flows
- Fuzzy prediction of traffic from LTF flows
- Preferential dropping

It is understood that to collect traffic from LTF flows for traffic prediction, the first step is to differentiate LTF flows from other flows. Our previous L-LRU proposal [8] for LTF flow detection is adopted in this paper. The L-LRU cache only requires to maintain flow state for a small number of LTF flows. Simulations driven by Internet traces have proved that the L-LRU cache is effective in the detection LTF flows. Due to space limit, only the fuzzy predictor and the preferential dropping mechanism are described in the paper.

3.1   Fuzzy Prediction of Traffic from LTF Flows

The ubiquitous existence of Long Range Dependence (LRD) among traffic aggregates at various time scales implies that traffic prediction is viable[1]. Previous research has shown that the predictability of Internet traffic can be exploited to improve network efficiency as well as QoS. Moreover, our previous work has shown that fuzzy prediction outperforms traditional linear prediction model with better prediction accuracy [9]. This might be explained by the fact that fuzzy prediction does not require to understand the distribution of system noise and observation noise in advance. So in this work, the fuzzy Kalman filter proposed in [9] is adopted for traffic prediction.

The input of the fuzzy predictor is a sequence of aggregated traffic from LTF flows, X_t, collected by the L-LRU cache. The time scale used in this work to aggregate incoming traffic is 0.01s. For a one-step-ahead prediction, the predicted traffic in the next control interval is given by

$$\hat{X}_t = \alpha \hat{X_{t-1}} + (1 - \alpha)X_{t-1} \qquad (2)$$

where α gives the exponential averaging weight of the past history. α is inferred by a fuzzy system on the basis of normalized prediction error, which is defined by

$$error = \frac{|X_t - \hat{X}_t|}{X_t} \qquad (3)$$

---

[1] Similar LRD trend can also be found for traffic from LTF flows, according to our experiments carried out with Internet traces.

The fuzzy inference system adopts the following linguistic rules.

- If *error* is large, then α is small
- If *error* is medium, then α is medium
- If *error* is small, then α is large

The membership function of the α and *error* is given in Figure 1.
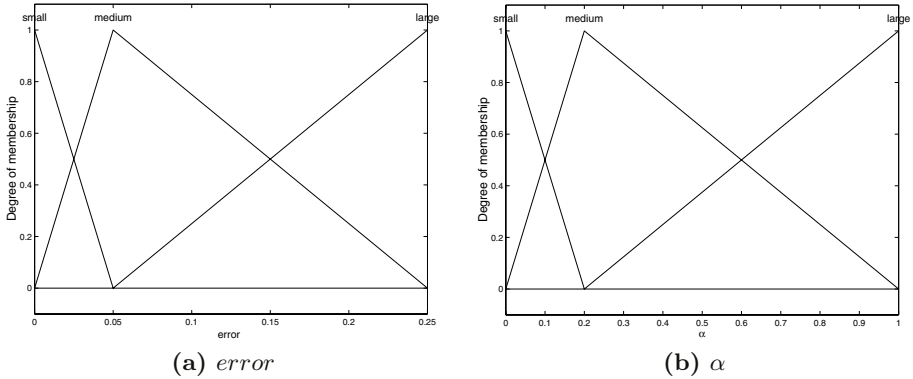
(a) *error* **(b)** α

Fig. 1. Membership function of the fuzzy predictor (a) error (b) α

The output of the predictor is the predicted traffic aggregate for the next control interval. The predicted traffic intensity is then used to compute the packet dropping probability based on (1). It should be noted that the system performance has certain correlation with the span of the control interval used. Generally the better performance comes with a smaller T, however, more computation is required. The simulation results presented in this paper used a control interval of 0.01s, which is the same as the time scale used for traffic aggregation. Since the computation overhead of the fuzzy inference engine is trivial, it has no problem for online computation.

3.2   Predictive Preferential Dropping

By preferential dropping, this work means to drop packets from LTF flows more vigorously during congestion. That is, if the packet count of flow i is longer than a threshold P_{ltf} (long-term) and its flow rate is higher than a threshold r_{ltf} (fast), packets from flow i will be dropped more heavily. On the contrary, packets from short-lived or slow flows will not be dropped, because dropping packets from those flows has not much effect on queue dynamics and little performance gain for congestion control. It is also highly desirable in the sense that dropping packets from cached flows only helps reduce the time to complete web sessions and allocate slow flows more bandwidth, because short-lived or slow flows are less likely to be cached. The preferential dropping mechanism works as follows.

For any incoming packet, if the flow which the packet belongs to is found in the cache, the FPPD randomly drops the packet only if the following two conditions are met. A) Flow rate (bandwidth) higher than a threshold r_{ltf} and B) Total packet count higher than a threshold P_{ltf}. If the two criteria are met for an LTF flow, its packets will be dropped with a probability given by

$$p_{id} = p_d * \frac{r_i}{r_{ltf}} \tag{4}$$

where r_i is the sending rate of flow i. Please note that the predictive packet dropping probability, p_d, still follows (1). However, the incoming traffic intensity is replaced by the predicted future traffic intensity using the above fuzzy predictor.

To implement the above preferential dropping, the total packet count for each cached flow is counted by the L-LRU cache, while the flow rate is approximated by the number of packets arrived during a fixed period of time T. Note that the threshold r_{ltf} is simply given by management policies, say if a flow is sending at a bandwidth above 0.1% of link bandwidth, its packets are more likely to be dropped. The P_{ltf} can be policy-based as well, for instance, most web sessions on nowadays Internet is less than 100KB, P_{ltf} can be set to 100 for a fixed packet size of 1KB with the aim to reduce the packet drop ratio for web sessions.

4 Simulation Results and Discussion

The FPPD scheme is implemented for NS-2 evaluation. Three other schemes RED, AFD [10] and LRURED [11] are used for performance comparison. Since the RED is not equipped with preferential dropping mechanism, its performance is used as a reference. The following simulations are done with a dumbbell topology given in [12], with a bottleneck link of 15Mb/s bandwidth and 10ms delay. The Round Trip Time distribution and other network configurations follow [12]. The queue limit is set to 300 packets with a fixed packet size P of 1000 bytes/packet. The traffic model used for simulation is based on the Internet traces, which contain both UDP and TCP flows. The TCP flows includes both long-term FTP flows and short-term web sessions. The Packmime [13] is used for web sessions simulation. All the simulations run for 600 seconds and are launched for 6 times. The results presented as follows are given with 95% confidence interval. Due to space limit, only the result under typical scenarios is presented in this paper, although all the simulation results proved the effectiveness of the FPPD scheme.

The performance of all the schemes under various traffic load level is presented as follows. 5 FTP flows and 5 constant rate UDP flows are running as long-term traffic, while the number of web sessions launched by Packmime is 150, 180 and 200 per second respectively to vary the traffic load level. The target delay is set to 16ms (target queue length 30 packets) for all the schemes. The packet count threshold for the FPPD is set to 50 to let any flows smaller than 50KB pass without any packets dropped. The rate threshold is set to 10

Table 1. Performance with Varied Traffic Load

Scheme	Link Utilization			Packet Loss Ratio		
	150	180	200	150	180	200
RED	0.86	0.94	0.96	0.82%	5.24%	9.56%
AFD	0.78	0.93	0.94	0.64%	3.57%	4.64%
LRURED	0.86	0.93	0.95	0.68%	3.66%	8.78%
FPPD	0.88	0.94	0.96	0.60%	2.22%	4.21%

and rate update period is 0.25, which means that only flows with a rate above $\frac{10*8Kb}{0.25sec} = 320Kbps$ may lose their packets during congestion period.

Table 1 gives the packet loss ratio and link utilization under the three scenarios for different schemes. It can be seen that the FPPD achieves highest link utilization and lowest packet loss ratio in all cases. The packet loss ratio is notably lower than other schemes, because of the predictive control with masked traffic noise. The benefit of traffic prediction is also revealed in the improved link efficiency.

In addition to network efficiency, the FPPD also improves the QoS for web users. As shown in Figure 2, where the cumulative distribution of time to complete a web session is given. It can be seen that the FPPD finishes 98% of all the web sessions less than 1.5 secs, by adopting the threshold based preferential dropping, while all other schemes spend longer to complete web transactions.
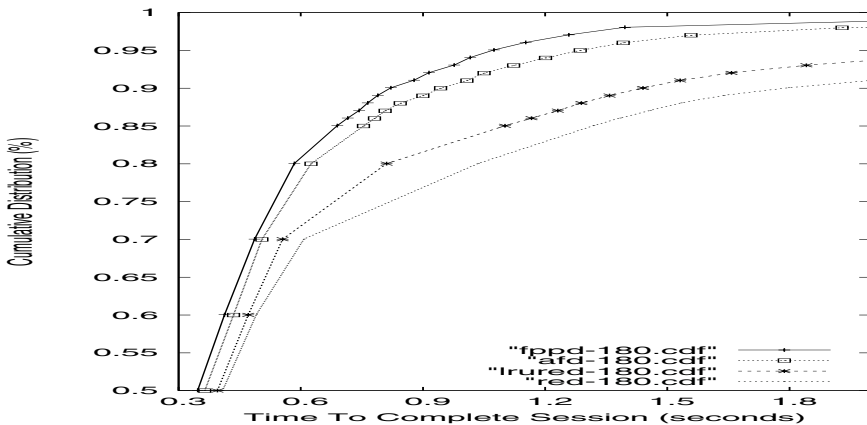
Fig. 2. Web Response Time

5   Conclusion and Future Work

A Fuzzy Predictive Preferential Dropping scheme for AQM is proposed in this paper. Incorporated with the fuzzy prediction and the predictive control, the FPPD scheme uses predicted traffic intensity from fast flows as the congestion

control index. It ignores transient traffic changes caused by short-lived flows while still promptly responding to real congestion. Its embedded preferential dropping mechanism treats SLS flows favorably. Simulation results prove that the proposed FPPD scheme outperforms other AQM schemes with improved link efficiency and reduced packet loss ratio. It also reduces the transaction completion time for web sessions. In the future, multiple-step ahead prediction will be investigated to further improve the performance of the FPPD scheme.

References

1. Sally Floyd and Van Jacobson. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1(4):397–413, 1993.
2. S. Athuraliya, S. H. Low, V. H. Li, and Q. Yin. Rem: Active queue management. *IEEE Network Magazine*, 15(3):48–53, May/June 2001.
3. W. Feng, D.Kanklur, and D.Saha. A self-configuring red gateway. In *Proceedings of INFOCOM*, March 1999.
4. C. V. Hollot, Vishal Misra, Donald F. Towsley, and Weibo Gong. Analysis and design of controllers for AQM routers supporting TCP flows. *IEEE Transactions on automatic control*, 47(6):945–959, June 2002.
5. ns 2 Network Simulator. Vint project group. available at: `http://www.isi.edu/nsnam/ns/`.
6. Yuan Gao, Guanghui He, and Jennifer C. Hou. On exploiting traffic predictability in active queue management. In *Proceedings of INFOCOM*, volume 3, pages 1346–1355, 2002.
7. R. Mahajan, S. Floyd, and D. Wetherall. Controlling high-bandwidth flows at the congested router. In *Proceedings of ICNP*, Sept 2001.
8. Lichang Che and Bin Qiu. Landmark filter: An efficient scheme for the detection of long-term fast internet flow. *Submmitted to IEEE Communication Letters*, 2005.
9. Bin Qiu. A predictive fuzzy congestion avoidance scheme. In *IEEE Global Telecommunications Conference (Globecomm 97)*, pages 967–971, 1997.
10. R. Pan, L. Breslau, B. Prabhakar, and S. Shenker. Approximate fairness through differential dropping. *ACM SIGCOMM Computer Communication Review*, 33(2), April 2003.
11. A. Smitha and A.L N. Reddy. Lru-red: An active queue management scheme to contain high bandwidth flows at congested routers. In *Proceedings of Globecom*, 2001.
12. S. Floyd and E. Kohler. Internet research needs better models. In *Proceedings of HorNets-I*, October 2002.
13. J. Cao, W.S. Cleveland, Y. Gao, K. Jeffay, F.D. Smith, and M.C. Weigle. Stochastic models for generating synthetic http source traffic. In *Proceedings of IEEE INFOCOM*, 2004.

A Fuzzy Method for Measuring Efficiency Under Fuzzy Environment

Hsuan-Shih Lee[1], Pei-Di Shen[2], and Wen-Li Chyr[3]

[1] Department of Shipping and Transportation Management
National Taiwan Ocean University
Keelung 202, Taiwan
[2] Department of Information Management
Ming Chung University
Taipei 111, Taiwan
[3] Graduate School of Management
Ming Chung University
Taipei 111, Taiwan

Abstract. DEA (data envelopment analysis) is a non-parametric technique for measuring and evaluating the relative efficiencies of a set of decision-making units (DMUs) in terms of a set of common inputs and outputs. Traditionally, the data of inputs and outputs are assumed to be measured with precision, i.e., the coefficients of DEA models are crisp value. However, this may not be always true. There are many circumstances where precise inputs and outputs can not be obtained. Under such situations, data of inputs and outputs can be represented by fuzzy numbers. Based on the dual program of DEA models, we propose fuzzy DEA models for CCR and BCC models. Our fuzzy DEA models provide crisp efficiency with fuzzy input and output data.

1 Introduction

Since the advent of the work of Charnes et at. [2], data envelopment analysis (DEA) is a methodology that has been widely used to measure relative efficiencies within a group of decision making units (DMUs) that utilize several inputs to produce a set of outputs. It has been applied to evaluate schools, hospitals and various organizations with multiple inputs and outputs. Unlike regression analysis, DEA uses only a single set of observations per DMU. Moreover, DEA is sensitive to outliers. Therefore, it is very difficult to evaluate the efficiency of DMU with uncertain inputs by conventional DEA models.

The DEA models with fuzzy data can more realistically represent real-world problems where data may be missing or uncertain than traditional DEA models. Sengupta [9] proposed a fuzzy DEA model and solved it by Zimmermann's method [11]. Triantis et al. [10] solved their fuzzy model with the parametric approach proposed by Carlsson et al. [1]. Kao and Liu [6, 7] transformed the fuzzy DEA model to a family of the conventional crisp DEA models by applying the α-cut approach. Hougaard [5] extended scores of technical efficiency to

fuzzy intervals to allow decision maker to use scores in combination with other sources of available performance information. Entani et al. [3] developed a DEA model with an interval efficiency. Guo et al. [4] proposed a fuzzy DEA model to deal with efficiency evaluation problem with given fuzzy input and output data, and extended the efficiency to be a fuzzy number. A crisp efficiency will be more helpful for decision makers because it provides clearer information. In this paper, we are going to derive fuzzy DEA models from this perspective.

2   The Data Envelopment Analysis Model

DEA is a mathematical model that measures the relative efficiency of DMUs with multiple inputs and outputs with no obvious production function to aggregate the data in its entirety. Relative efficiency is defined as the ratio of total weighted output over weighted input. By comparing n units with s outputs denoted by y_{rk}, $r = 1, \ldots, s$ and m inputs denoted by x_{ik}, $i = 1, \ldots, m$, the efficiency measure for DMU k is

$$
\begin{aligned}
h_k = Max \sum_{r=1}^{s} u_r y_{rk} \\
s.t. \qquad \sum_{i=1}^{m} v_i x_{ik} = 1, \\
\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0 \quad for \ j = 1, \ldots, n, \\
u_r \geq 0 \quad for \ r = 1, \ldots, s, \\
v_i \geq 0 \quad for \ i = 1, \ldots, m.
\end{aligned}
\tag{1}
$$

Mode (1), often referred to as the input-oriented CCR (Charnes Cooper Rhodes) model [2], assumes that the production function exhibits constant returns-to-scale. The following is referred to as the output-oriented CCR model:

$$
\begin{aligned}
\frac{1}{g_k} = Min \sum_{i=1}^{m} v_i x_{ik} \\
s.t. \qquad \sum_{r=1}^{s} u_r y_{rk} = 1, \\
\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0 \quad for \ j = 1, \ldots, n, \\
u_r \geq 0 \quad for \ r = 1, \ldots, s, \\
v_i \geq 0 \quad for \ i = 1, \ldots, m.
\end{aligned}
\tag{2}
$$

The BCC (Banker Chang Cooper) model (2) adds an additional constant variable, c_k, in order to permit variable returns-to-scale. The following is input-oriented BCC model:

$$
\begin{aligned}
h_k = Max \sum_{r=1}^{s} u_r y_{rk} + c_k \\
s.t. \qquad \sum_{i=1}^{m} v_i x_{ik} = 1, \\
\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} - c_k \geq 0 \quad for \ j = 1, \ldots, n, \\
u_r \geq 0 \quad for \ r = 1, \ldots, s, \\
v_i \geq 0 \quad for \ i = 1, \ldots, m.
\end{aligned}
\tag{3}
$$

Model (4) is referred to as the output-oriented BCC model:

$$
\begin{aligned}
\frac{1}{g_k} = Min \sum_{i=1}^{m} v_i x_{ik} + c_k \\
s.t. \qquad \sum_{r=1}^{s} u_r y_{rk} = 1, \\
\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} + c_k \geq 0 \quad for \ j = 1, \ldots, n, \\
u_r \geq 0 \quad for \ r = 1, \ldots, s, \\
v_i \geq 0 \quad for \ i = 1, \ldots, m.
\end{aligned}
\tag{4}
$$

If a DMU proves to be inefficient, a combination of other efficient units can produce either greater output for the same composite of inputs, use fewer inputs to produce the same composite of outputs or some combination of the two. A hypothetical decision making unit, k', can be composed as an aggregate of the efficient units referred to as the efficient reference set for inefficient unit k. The solution to the dual problem of above models directly computes the multipliers required to compile aggregated efficient unit k'. Therefore, we will focus on the dual models. The dual of model (1) is:

$$
\begin{aligned}
& h_k = Min\theta_k \\
& s.t. \sum_{j=1}^{n} \lambda_{kj} x_{ij} - \theta_k x_{ik} \leq 0 \quad for \ i = 1, \ldots, m, \\
& \qquad \sum_{j=1}^{n} \lambda_{kj} y_{rj} \geq y_{rk} \ for \ r = 1, \ldots, s, \\
& \qquad \lambda_{kj} \geq 0 \quad for \ j = 1, \ldots, n.
\end{aligned}
\tag{5}
$$

The dual of model (2) is:

$$
\begin{aligned}
& \frac{1}{g_k} = Max\theta_k \\
& s.t. \sum_{j=1}^{n} \lambda_{kj} x_{ij} - x_{ik} \leq 0 \quad for \ i = 1, \ldots, m, \\
& \qquad \sum_{j=1}^{n} \lambda_{kj} y_{rj} \geq \theta_k y_{rk} \ for \ r = 1, \ldots, s, \\
& \qquad \lambda_{kj} \geq 0 \quad for \ j = 1, \ldots, n.
\end{aligned}
\tag{6}
$$

The dual of model (3) is:

$$
\begin{aligned}
& h_k = Min\theta_k \\
& s.t. \sum_{j=1}^{n} \lambda_{kj} x_{ij} - \theta_k x_{ik} \leq 0 \quad for \ i = 1, \ldots, m, \\
& \qquad \sum_{j=1}^{n} \lambda_{kj} y_{rj} \geq y_{rk} \ for \ r = 1, \ldots, s, \\
& \qquad \sum_{j=1}^{n} \lambda_{kj} = 1.
\end{aligned}
\tag{7}
$$

The dual of model (4) is:

$$
\begin{aligned}
& \frac{1}{g_k} = Max\theta_k \\
& s.t. \sum_{j=1}^{n} \lambda_{kj} x_{ij} - x_{ik} \leq 0 \quad for \ i = 1, \ldots, m, \\
& \qquad \sum_{j=1}^{n} \lambda_{kj} y_{rj} \geq \theta_k y_{rk} \ for \ r = 1, \ldots, s, \\
& \qquad \sum_{j=1}^{n} \lambda_{kj} = 1.
\end{aligned}
\tag{8}
$$

3   Fuzzy DEA Models

In this section, we are going to extend models (5)-(8) to fuzzy environment.

Definition 1. *The α-cut of fuzzy set $\tilde{A}$, $\tilde{A}^\alpha$, is the crisp set $\tilde{A}^\alpha = \{x | \mu_{\tilde{A}}(x) \geq \alpha\}$. The support of $\tilde{A}$ is the crisp set $Supp(\tilde{A}) = \{x | \mu_{\tilde{A}}(x) > 0\}$. $\tilde{A}$ is normal iff $sup_{x \in U} \mu_{\tilde{A}}(x) = 1$, where U is the universal set.*

Definition 2. *A fuzzy subset $\tilde{A}$ of real number R is convex iff*

$$
\mu_{\tilde{A}}(\lambda x + (1 - \lambda)y) \geq (\mu_{\tilde{A}}(x) \wedge \mu_{\tilde{A}}(y)), \forall x, y \in R, \forall \lambda \in [0, 1],
$$

where $\wedge$ denotes the minimum operator.

Definition 3. $\tilde{A}$ *is a fuzzy number iff* $\tilde{A}$ *is a normal and convex fuzzy subset of R.*

Definition 4. *A triangular fuzzy number* $\tilde{A}$ *is a fuzzy number with piecewise linear membership function* $\mu_{\tilde{A}}$ *defined by*

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x - a_L}{a_M - a_L}, & a_L \leq x \leq a_M, \\ \frac{a_R - x}{a_R - a_M}, & a_M \leq x \leq a_R, \\ 0, & otherwise, \end{cases}$$

which can be denoted as a triplet (a_L, a_M, a_R).

Definition 5. *Let* $\tilde{A}$ *and* $\tilde{B}$ *be two fuzzy numbers. Let* $\circ$ *be a operation on real numbers, such as* $+$, $-$, $*$, $\wedge$, $\vee$, *etc. By extension principle, the extended operation* $\circ$ *on fuzzy numbers can be defined by*

$$\mu_{\tilde{A} \circ \tilde{B}}(z) = \sup_{x,y : z = x \circ y} \{ \mu_{\tilde{A}}(x) \wedge \mu_{\tilde{B}}(y) \}. \tag{9}$$

Definition 6. *Let* $\tilde{A}$ *and* $\tilde{B}$ *be two fuzzy numbers.* $\tilde{A} \leq \tilde{B}$ *if only if* $\tilde{A} \vee \tilde{B} = \tilde{B}$.

Definition 7. *Let* $\tilde{A}$ *be a fuzzy number. Then* $\tilde{A}_{\alpha}^{L}$ *and* $\tilde{A}_{\alpha}^{U}$ *are defined as*

$$\tilde{A}_{\alpha}^{L} = \inf_{\mu_{\tilde{A}}(z) \geq \alpha} (z) \tag{10}$$

and

$$\tilde{A}_{\alpha}^{U} = \sup_{\mu_{\tilde{A}}(z) \geq \alpha} (z) \tag{11}$$

respectively.

Assume inputs and outputs of DMUs are represented by triangular fuzzy numbers. Fuzzy DEA models correspond to (5)-(8) are shown as follows. Fuzzy model of (5) is

$$\begin{aligned} h_k &= Min\theta_k \\ s.t. \; \textstyle\sum_{j=1}^{n} \lambda_{kj} \tilde{X}_{ij} - \theta_k \tilde{X}_{ik} &\leq 0 \quad for \; i = 1, \ldots, m, \\ \textstyle\sum_{j=1}^{n} \lambda_{kj} \tilde{Y}_{rj} &\geq \tilde{Y}_{rk} \; for \; r = 1, \ldots, s, \\ \lambda_{kj} &\geq 0 \quad for \; j = 1, \ldots, n. \end{aligned} \tag{12}$$

Fuzzy model (6) is

$$\begin{aligned} \tfrac{1}{g_k} &= Max\theta_k \\ s.t. \; \textstyle\sum_{j=1}^{n} \lambda_{kj} \tilde{X}_{ij} - \tilde{X}_{ik} &\leq 0 \quad for \; i = 1, \ldots, m, \\ \textstyle\sum_{j=1}^{n} \lambda_{kj} \tilde{Y}_{rj} &\geq \theta_k \tilde{Y}_{rk} \; for \; r = 1, \ldots, s, \\ \lambda_{kj} &\geq 0 \quad for \; j = 1, \ldots, n. \end{aligned} \tag{13}$$

Fuzzy model (7) is

$$\begin{aligned} h_k &= Min\theta_k \\ s.t. \; \textstyle\sum_{j=1}^{n} \lambda_{kj} \tilde{X}_{ij} - \theta_k \tilde{X}_{ik} &\leq 0 \quad for \; i = 1, \ldots, m, \\ \textstyle\sum_{j=1}^{n} \lambda_{kj} \tilde{Y}_{rj} &\geq \tilde{Y}_{rk} \; for \; r = 1, \ldots, s, \\ \textstyle\sum_{j=1}^{n} \lambda_{kj} &= 1. \end{aligned} \tag{14}$$

Fuzzy model (8) is

$$
\begin{aligned}
\frac{1}{g_k} &= Max\theta_k \\
s.t. \ \sum_{j=1}^{n} \lambda_{kj}\tilde{X}_{ij} - \tilde{X}_{ik} &\leq 0 \qquad for \ i = 1,\ldots,m, \\
\sum_{j=1}^{n} \lambda_{kj}\tilde{Y}_{rj} &\geq \theta_k\tilde{Y}_{rk} \ for \ r = 1,\ldots,s, \\
\sum_{j=1}^{n} \lambda_{kj} &= 1.
\end{aligned}
\tag{15}
$$

Because of the convexity of fuzzy numbers, we can have the following definition equivalent to definition 6.

Definition 8. *Let $\tilde{A}$ and $\tilde{B}$ be two fuzzy numbers. $\tilde{A} \leq \tilde{B}$ if only if $\tilde{A}_\alpha^L \leq \tilde{B}_\alpha^L$ and $\tilde{A}_\alpha^R \leq \tilde{B}_\alpha^R$ for $\forall \alpha \in [0,1]$.*

Lemma 1. *Let $\tilde{A}$ and $\tilde{B}$ be two triangular fuzzy numbers denoted by (a_L, a_M, a_R) and (b_L, b_M, b_R) respectively. Then $\tilde{A} \leq \tilde{B}$ if only if $a_L \leq b_L$, $a_M \leq b_M$ and $a_R \leq b_R$.*

Proof: By definition 4 and 7, we have

$$
\begin{aligned}
\tilde{A}_\alpha^L &= \alpha(a_M - a_L) + a_L, \tilde{A}_\alpha^R = a_R - \alpha(a_R - a_M), \\
\tilde{B}_\alpha^L &= \alpha(b_M - b_L) + b_L, \tilde{B}_\alpha^R = b_R - \alpha(b_R - b_M),
\end{aligned}
$$

If for all $\alpha \in [0,1]$, $\tilde{A}_\alpha^L \leq \tilde{B}_\alpha^L$ and $\tilde{A}_\alpha^R \leq \tilde{B}_\alpha^R$ then

$$
\alpha(a_M - a_L) + a_L \leq \alpha(b_M - b_L) + b_L \ and \ a_R - \alpha(a_R - a_M) \leq b_R - \alpha(b_R - b_M).
$$

Let $\alpha = 1$. We have $a_M \leq b_M$. Let $\alpha = 0$. We have $a_L \leq b_L$ and $a_R \leq b_R$.

If $a_L \leq b_L$, $a_M \leq b_M$ and $a_R \leq b_R$ then

$$
\alpha(a_M - a_L) + a_L \leq \alpha(b_M - b_L) + b_L
$$

and

$$
a_R - \alpha(a_R - a_M) \leq b_R - \alpha(b_R - b_M)
$$

for $\alpha \in [0,1]$. That is, $\tilde{A}_\alpha^L \leq \tilde{B}_\alpha^L$ and $\tilde{A}_\alpha^R \leq \tilde{B}_\alpha^R$ for all $\alpha \in [0,1]$. $\qquad \square$

Let $\tilde{X}_{ij} = (x_{ij_L}, x_{ij_M}, x_{ij_R})$ and $\tilde{Y}_{rj} = (y_{rj_L}, y_{rj_M}, y_{rj_R})$. Following lemma 1, we can rewrite fuzzy model (12) as follows:

$$
\begin{aligned}
h_k &= Min\theta_k \\
s.t. \ \sum_{j=1}^{n} \lambda_{kj}x_{ij_L} - \theta_k x_{ik_L} &\leq 0 \qquad for \ i = 1,\ldots,m, \\
\sum_{j=1}^{n} \lambda_{kj}x_{ij_M} - \theta_k x_{ik_M} &\leq 0 \qquad for \ i = 1,\ldots,m, \\
\sum_{j=1}^{n} \lambda_{kj}x_{ij_R} - \theta_k x_{ik_R} &\leq 0 \qquad for \ i = 1,\ldots,m, \\
\sum_{j=1}^{n} \lambda_{kj}y_{rj_L} &\geq y_{rk_L} \ for \ r = 1,\ldots,s, \\
\sum_{j=1}^{n} \lambda_{kj}y_{rj_M} &\geq y_{rk_M} \ for \ r = 1,\ldots,s, \\
\sum_{j=1}^{n} \lambda_{kj}y_{rj_R} &\geq y_{rk_R} \ for \ r = 1,\ldots,s, \\
\lambda_{kj} &\geq 0 \qquad for \ j = 1,\ldots,n.
\end{aligned}
\tag{16}
$$

Fuzzy model (13) can be equivalently rewritten as model (17):

$$
\begin{aligned}
\frac{1}{g_k} &= Max\theta_k \\
s.t. \quad & \sum_{j=1}^{n} \lambda_{kj} x_{ij_L} - x_{ik_L} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} x_{ij_M} - x_{ik_M} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} x_{ij_R} - x_{ik_R} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_L} \geq \theta_k y_{rk_L} && for\ r = 1, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_M} \geq \theta_k y_{rk_M} && for\ r = 1, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_R} \geq \theta_k y_{rk_R} && for\ r = 1, \ldots, s, \\
& \lambda_{kj} \geq 0 && for\ j = 1, \ldots, n.
\end{aligned}
\tag{17}
$$

Fuzzy model (14) can be equivalently rewritten as model (18):

$$
\begin{aligned}
h_k &= Min\theta_k \\
s.t. \quad & \sum_{j=1}^{n} \lambda_{kj} x_{ij_L} - \theta_k x_{ik_L} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} x_{ij_M} - \theta_k x_{ik_M} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} x_{ij_R} - \theta_k x_{ik_R} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_L} \geq y_{rk_L} && for\ r = 1, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_M} \geq y_{rk_M} && for\ r = 1, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_R} \geq y_{rk_R} && for\ r = 1, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_{kj} = 1.
\end{aligned}
\tag{18}
$$

Fuzzy model (15) can be equivalently rewritten as model (19):

$$
\begin{aligned}
\frac{1}{g_k} &= Max\theta_k \\
s.t. \quad & \sum_{j=1}^{n} \lambda_{kj} x_{ij_L} - x_{ik_L} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} x_{ij_M} - x_{ik_M} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} x_{ij_R} - x_{ik_R} \leq 0 && for\ i = 1, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_L} \geq \theta_k y_{rk_L} && for\ r = 1, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_M} \geq \theta_k y_{rk_M} && for\ r = 1, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_{kj} y_{rj_R} \geq \theta_k y_{rk_R} && for\ r = 1, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_{kj} = 1.
\end{aligned}
\tag{19}
$$

To solve fuzzy models (12)-(15), we solve their equivalent linear programming problems (16)-(19).

Theorem 1. *Efficiencies given by fuzzy models (12)-(15) are all no greater than 1. That is, $h_k \leq 1$ and $g_k \leq 1$.*

Proof: For models (12) and (14), let $\lambda_{kj} = \begin{cases} 1\ if\ j = k \\ 0\ if\ j \neq k. \end{cases}$ Then $\theta_k = 1$. Let θ_k^* be the optimal objective value of models (12) and (13). We have

$$
\theta_k^* \leq \theta_k = 1.
$$

That is, $h_k \leq 1$.

Similarly, for models (13) and (15), let

$$\lambda_{kj} = \begin{cases} 1 \ if \ j = k \\ 0 \ if \ j \neq k. \end{cases}$$

Then $\theta_k = 1$. Let θ_k^* be the optimal objective value of models (13) and (15). We have

$$\theta_k^* \geq \theta_k = 1.$$

That is, $g_k \leq 1$. □

4   Conclusions

Traditional DEA deals with crisp data. When data are difficult to measure, fuzzy numbers may be introduced to represent the vagueness or uncertainty of the data. In this paper, we extended the duals of DEA models to fuzzy environment. In our extension, the efficiency measured with fuzzy data is a crisp value which enables decision makers to decide whether a DMU is relative efficient or not very clear. It is also obvious that our fuzzy models encompass conventional DEA models.

Acknowledgement

References

1. C. Carlsson, P. Korhonen, A paramet4ric approach to fuzzy linear programming, Fuzzy Sets and Systems 20 (1986) 17-30.
2. A. Charnes, W.W. Cooper, E. Rhodes, Measuring the efficiency of decision-making units, European J. Oper. Res. 2 (1978) 429-444.
3. T. Entani, Y. Maeda, H. Tananka, Dual models of interval DEA and its extension to interval data, European J. Oper. Res. 136 (2002) 32-45.
4. P. Guo, H. Tanaka, Fuzzy DEA: a perceptual evaluation method, Fuzzy Sets and Systems 119 (2001) 149-160.
5. J.L. Hougaad, Fuzzy scores of technical efficiency, European J. Oper. Res. 115 (1999) 529-541.
6. C. Kao, S.T. Liu, Data envelopment analysis with missing data: an application to University libraries in Taiwan, Journal of the Operational Research Society 51 (2000) 897-905.
7. C. Kao, S.T. Liu, Fuzzy efficiency measures in data envelopment analysis, Fuzzy Sets and Systems 113 (2000) 429-437.
8. S. Lerworasirikul, S.-C. Fang, J.A. Jonies, H. L.W. Nuttle, Fuzzy data envelopment analysis (DEA): a possibility approach, Fuzzy Sets and Systems 139 (2003) 379-394.
9. J.K. Sengupta, A fuzzy systems approach in data envelopment analysis, Comput. Math. Appl. 24 (1992) 259-266.
10. K. Triantis, O. Girod, A mathematical programming approach for measuring techincal efficiency in a fuzzy environment, J. Prod. Anal. 10 (1998) 85-102.
11. H.J. Zimmermann, Description and optimization of fuzzy system, Interat. J. General System 2 (1976) 209-216.

Anytime Iterative Optimal Control
Using Fuzzy Feedback Scheduler

Feng Xia and Youxian Sun

National Laboratory of Industrial Control Technology
Institute of Modern Control Engineering
Zhejiang University, Hangzhou 310027, China
{xia,yxsun}@iipc.zju.edu.cn

Abstract. From a viewpoint of integrating control and scheduling, the impact of resource availability constraints on the implementation of iterative optimal control (IOC) algorithms is considered. As a novel application in the emerging field of feedback scheduling, fuzzy technology is employed to construct a feedback scheduler intended for anytime IOC applications. Thanks to the anytime nature of the IOC algorithm, it is possible to abort the optimization routine before it reaches the optimum. The maximum iteration number within the IOC algorithm is dynamically adjusted to achieve a desired CPU utilization level. Thus a tradeoff is done between the available CPU time and the quality of control. Preliminary simulation results argue that the proposed approach is effective in managing the inherent uncertainty in control task execution and delivers better performance than traditional IOC algorithm in computing resource constrained environments.

1 Introduction

Since almost all real-world systems are inherently nonlinear, the control issue in nonlinear systems is one of the hot topics attracting a lot of research efforts from the control community. Still, the controller design remains difficult due to the increasing complexity of these systems. As a successful methodology, many researchers treated the control of a nonlinear system as an optimization problem, and solved it using iterative procedures, e.g. [1-3]. These algorithms form the basic methodology of iterative optimal control (IOC). Up to now, however, little work is dedicated to developing efficient real-time implementations of IOC algorithms in real-world applications. In practice, the control community generally assumes that the real-time platform used to implement the control algorithms can provide deterministic timing behaviors (e.g. constant execution time) and sufficient computing resources (CPU time) as needed. It is quite clear that this is not always the truth, especially in an embedded control system that features resource constraints and workload uncertainty [4,5]. With recently increasing adoption of embedded controllers, modern control applications representatively feature computing resource constraints. As a consequence, when moving from the research laboratories to real-world applications, control tasks become subject to the availability of computing resources. As for IOC, it is intuitive that more iterations lead to higher accuracy of the solution, since it searches an optimal control input from iteration to iteration. On the other hand, increase in iterations demands more computation time. In a resource-constrained environment,

this will cause unexpected task overruns. When the control task becomes unschedulable, the quality of control will be reversely degraded.

In this paper, we consider the impact of computing resource constraints on IOC algorithms. As an emerging technique for integrating control and scheduling, feedback scheduling seems to be a promising methodology to address the abovementioned problem. With the help of a feedback scheduler, the CPU demand can be dynamically adjusted through manipulating timing parameters, e.g. execution time of the control task [6,7]. It is interesting that IOC algorithms are of anytime nature. That is to say, they always generate a result, but with a quality level that increase with the execution time. Therefore, it is possible to abort the optimization routine before it has reached the optimum. In this way, a tradeoff is done between the available computation time, i.e., how long time the controller may spend calculating the new control signal, and the control performance. In order to cope with the non-deterministic characteristics of iterative algorithms' execution times as well as measurement noises, we employ fuzzy logic control technology [8] to construct feedback schedulers for anytime IOC applications, with the primary goal of improving control performance as much as possible under resource constraints.

This paper is structured as follows. The problem we consider is described in Section 2. We present the architecture and algorithm of the fuzzy feedback scheduler for anytime iterative optimal control in Section 3. And its performance is evaluated via preliminary simulation studies in Section 4. Section 5 concludes this paper.

2   Problem Statement

For the sake of simple description, we employ the following IOC algorithm from [3] as an example to illustrate our methodology.

Consider a SISO (single-input and single-output) nonlinear system described as:

$$y(k+1) = f[y(k),...,y(k-n+1),u(k),...,u(k-m+1)] \tag{1}$$

where $y(k)$ and $u(k)$ are the system output and input respectively, $f(\bullet)$ is an unknown nonlinear function. The reference model is given by:

$$y_m(k+1) = A_m y_m(k) + B_m u_m(k) \tag{2}$$

where $A_m(z) = a_0 + a_1 z^{-1} + ... + a_{n-1} z^{-n+1}$, $B_m(z) = b_0 + b_1 z^{-1} + ... + b_{m-1} z^{-m+1}$. As an output tracking problem, the control objective is to find a control input $u(k)$ that satisfies $\lim_{k \to \infty} | y_m(k) - y(k) |_{u(k)} = 0$.

The referred iterative optimal control law is based on a neural network (NN) model of the plant, and follows an iterative learning manner. In each sample, the control signal is gradually refined via iterations. Let $u_m(k)$ denote the reference input, and the NN to approximate the SISO system is described as $y_n(k+1) = NN(X(k),W)$. The iterative algorithm operates as follows.

- 0) Train the NN model offline to provide good approximation of the plant dynamics. Initialize the system with the upper bound of the number of iterations I_{max} and the allowable maximum control error E_{max}.
- 1) At each instant k, let the iteration $i = 0$, and sample $y(k)$.

- 2) Calculate $u_m(k)$ and $y_m(k+1)$.
- 3) Calculate $u(k, i)$ according to

$$u(k,i+1) = u(k,i) + \lambda [y_m(k+1) - NN_f(k,i)] \times [NN'_f(k,i)]^{-1}$$
$$u(k,0) = u(k-1)$$

(3)

where $\lambda > 0$ is the iterative learning factor.
- 4) For the updated input $u(k, i+1)$, evaluate $y_n(k+1, i+1)$ via the NN model.
- 5) Let $i = i +1$. If $i > I_{max}$, go to next step. Otherwise, if $| y_n(k+1, i+1) - y_m(k+1) | > E_{max}$, return to step 3).
- 6) Issue the control input onto the plant, and wait for the next instant $k = k +1$.

The above mathematical description clearly shows us the anytime nature of IOC algorithms, which fits nicely into the general framework of scheduling of imprecise computations [9]. Generally speaking, as the number of iterations increases, the resulting control performance becomes better. However, the execution time as well as the requested CPU utilization U of the algorithm will increase at the same time. Traditionally, the maximum iteration number is chosen regardless of the computing resource availability. In a resource sufficient environment, good performance may be achieved, e.g., with an adequately large I_{max}. For a resource-constrained environment with time-varying workloads, however, such a pre-specified maximum iterations may result in bad performance. The reason behind this is that traditional IOC says little about the real-time implementation, treating control and scheduling separately and causing the control task to be unschedulable sometimes. As we know, much more iterations are needed to reach the optimum when the system is in transient process than in equilibrium. This makes the IOC algorithm to be more sensitive to the resource availability constraints.

From a viewpoint of integrating feedback control and real-time scheduling, we formulate the problem as

$$\min_C J = \eta(C) \quad \text{s. t. } U = C / h \le U_{sp}$$

(4)

where C and h are the execution time and sampling period of the iterative controller, J is the cost function with respect to C, and U_{sp} is the available CPU utilization. The sampling period h is assumed to be constant. It is typical that $\eta(\bullet)$ is a monotonic decreasing function. Therefore, following the methodology of feedback scheduling, the control cost will be simply minimized by setting C to be $h*U_{sp}$. However, this is not so easy as it seems. Although it may be feasible to restrict the execution time of the algorithm through regulating the upper bound of iteration number, the relationship between execution time and the iteration number is uncertain. The computation time of each iteration may vary significantly, e.g. due to the changes in the data to be processed. Since measurement noises inevitably exist, the available timing parameters of the control task are inaccurate. In addition, the disturbance from other concurrent tasks further makes the execution of the IOC task uncertain.

3 Fuzzy Feedback Scheduler

The motivation for using fuzzy logic for feedback scheduling of IOC tasks is mainly raised from its capacity to formalize control algorithms that can tolerate imprecision

and uncertainty [8,10]. It releases us from the mathematical system models, and emulates the cognitive process of human beings. Moreover, its simplicity makes it suitable for real-time applications. The architecture of fuzzy feedback scheduling methodology is given in Fig. 1. From the control perspective, the scheduler can be viewed as a feedback controller. The controlled variable is the CPU utilization while the upper bound of iteration numbers within the IOC task acts as the manipulated variable. The role of the fuzzy feedback scheduler is to control the CPU utilization to the desired level and to maximize the upper bound of iteration numbers under dynamic resource constraints, thus optimizing the control performance.
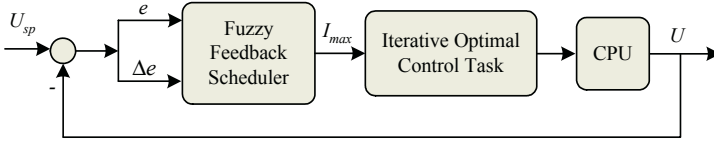
Fig. 1. Fuzzy logic based feedback scheduling structure
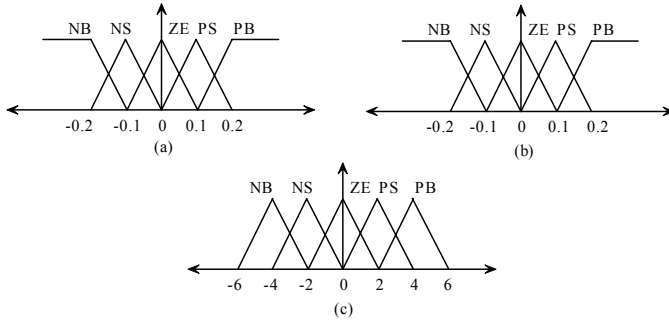
Fig. 2. Membership functions: (a) *E*, (b) *DE*, (c) *DI*

As we can see from Fig. 1, the fuzzy feedback scheduler can be designed as a multi-input-single-output fuzzy logic controller. The input variables are the error between desired CPU utilization and actual (measured) one, $e = U_{sp} - U$, and the change in error, $\Delta e = (e_N - e_{N-1})$, where N denotes the scheduler's sampling instance. The associated linguistic variables are E and DE, respectively. The output is the linguistic variable DI that represents the increment of the maximum iteration number ΔI_{max}. Fig.2 shows the membership functions of the input and output linguistic variables, which are utilized in the simulations in Section 4. The set of linguistic values is the same for each of them, which is {NB, NS, ZE, PS, PB}.

The rule base is constructed based on our experience on simulation studies and well-established knowledge from both control community and real-time scheduling community. Table 1 describes all possible rules. And the fuzzy sets representing the conclusions are obtained using the max-min inference mechanism.

To produce a numeric value $\Delta I_{max,N}$ from the fuzzy sets qualification of the conclusions, the center of gravity method is used in the defuzzification procedure. Since the number of iterations should be an integer value, the final output of the scheduler is calculated by

Table 1. Rule base for fuzzy feedback scheduler

DI		DE				
		NB	NS	ZE	PS	PB
	NB	NB	NB	NB	NS	ZE
	NS	NB	NS	NS	ZE	PS
E	ZE	NS	ZE	ZE	ZE	PS
	PS	NS	ZE	PS	PS	PB
	PB	ZE	PS	PB	PB	PB

$$I_{\max,N} = I_{\max,N-1} + [\Delta I_{\max,N}] \tag{5}$$

where $[\bullet]$ denotes the nearest integer. In this way, the upper bound of iteration number within the IOC algorithm will be dynamically adjusted with respect to the variations in use of available computing resources. Note that the actual iteration number is different from the specified upper bound.

Before practical application of the fuzzy feedback scheduler, one more problem must be addressed. When the controlled system is in equilibrium, the needed iteration number is very small. Thus the input E will always be PB, which in turn leads to continuous increase in I_{max}. This is of no sense since the actual iteration number is always smaller than I_{max} in equilibrium and may causes overload in case of reference changes and external disturbances. Therefore, we introduce a simple *activation condition* for the fuzzy feedback scheduler as follows. Let T be the sampling period of the feedback scheduler.

```
In each time interval [(N-1)T N*T]
  FLAG = true
  For each control period
    If actual iteration number < I_max and U_sp > U
    then FLAG = false
  Next
At time instance N*T
  If FLAG = true then activate fuzzy feedback scheduler
```

4   Example

In this section, we carry out preliminary simulations to evaluate the performance of the proposed fuzzy feedback scheduler. Consider the following nonlinear system to be controlled using the IOC algorithm.

$$y(k+1) = 0.5y(k)\sin(u(k)) + u^2(k) \tag{6}$$

The reference model is described by $y_m(k+1) = 0.6y_m(k) + 0.4u_m(k)$. We use a three-layer feedforward network to approximate the nonlinear system. It is trained using the Levenberg-Marquardt back-propagation algorithm. The IOC algorithm is implemented as a real-time task, which executes periodically with a sampling period of 10 ms. It is assumed that the underlying system schedules the task in a way that a new instance is not allowed to start execution until the previous instance has completed. If an instance does not start executing when a new one is activated, it will be

discarded. For all simulations, the desired CPU utilization $U_{sp} = 0.5$. The reference input $u_m(k)$ is given as a square wave with a frequency of 0.5 rad/s.

In the first case, the nonlinear system is controlled using the traditional IOC algorithm given in Section 2. Involved parameters are chosen as follows: $\lambda = 0.01$, $I_{max} = 30$, and $E_{max} = 0.01$. The resulting control performance is shown in Fig. 3. In the second case, the anytime IOC approach with the fuzzy feedback scheduler we present in Section 3 is implemented. The upper bound of iteration number is online adjusted in response to workload variations. The scheduler acts as a periodic task with a period of 30 ms and the scheduling overhead is ignored. The corresponding output tracking performance is given in Fig.4.

It is clear that the proposed approach outperforms the traditional one. Thanks to the dynamic adjustment of maximum iteration number, unexpected task overruns can be avoided while the use of available CPU time is maximized. With the help of the fuzzy feedback scheduler, the impacts of computing resource constraints and task execution uncertainty on the quality of control are reduced significantly.
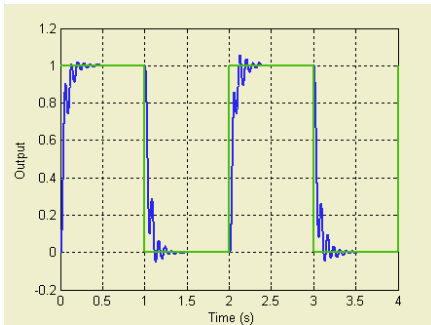
Fig. 3. System output tracking performance in the first case
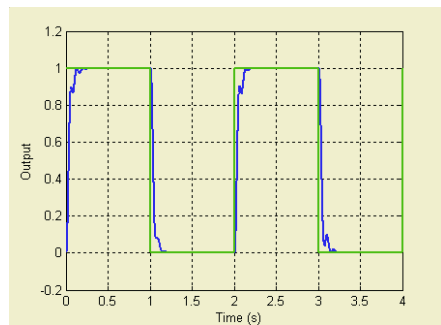
Fig. 4. System output tracking performance in the second case

5 Conclusions

In this paper, we demonstrate a novel application of fuzzy technology in the newly emerging field of feedback scheduling. The anytime nature of iterative optimal control algorithms is exploited in the context of integrated feedback control and real-time scheduling. A simple while effective fuzzy feedback scheduler is proposed to cope with the impact of workload uncertainty on the IOC applications in resource constrained environments. It is found that better performance can be achieved through dynamically assigning the upper bound of iteration number within the IOC algorithm, in response to workload variations. Preliminary simulation results argue that the proposed fuzzy feedback scheduler can effectively manage the inherent uncertainty within IOC task execution.

References

1. P. D. Roberts: Two-Dimensional Analysis of an Iterative Nonlinear Optimal Control Algorithm. IEEE Trans. on Circuits and Systems - I, 49:6 (2002) 872-878

2. Fuli Wang and Mingzhong Li: Optimal Iterative Control of Unknown Nonlinear Systems Using Neural Networks, in Proc. 36th IEEE CDC, San Diego, California USA (1997) 2201-2206
3. Yubin He, Xinzhong Li: Neural Network Control Technology and Its Applications. Beijing: Science Press (2000)
4. Feng Xia, Zhi Wang, and Youxian Sun: Integrated Computation, Communication and Control: Towards Next Revolution in Information Technology, Lecture Notes in Computer Science, Vol. 3356, Springer-Verlag, Heidelberg (2004) 117-125
5. K.-E. Årzén, A. Cervin, D. Henriksson: Resource-Constrained Embedded Control Systems: Possibilities and Research Issues, in Proc. CERTS: Co-design of Embedded Real-Time Systems Workshop, Porto, Portugal (2003)
6. D. Henriksson, A. Cervin, Johan Åkesson, Karl-Erik Årzén: Feedback Scheduling of Model Predictive Controllers. In Proc. 8th IEEE RTAS, San Jose, CA (2002) 207-216
7. Feng Xia and Youxian Sun: NN-based Iterative Learning Control under Resource Constraints: A Feedback Scheduling Approach, Lecture Notes in Computer Science, Vol. 3498, Springer-Verlag, Heidelberg (2005) 1-6
8. Kevin M. Passino and Stephen Yurkovich: Fuzzy Control, Addison Wesley Longman, Menlo Park, CA (1998)
9. J. W.S. Liu, K. Lin, W. Shin, A. Chuang-shi Yu: Algorithms for Scheduling Imprecise Computations, IEEE Computer 25:3 (1991) 58-68
10. Hong Jin, Hongan Wang, Yong Fu, Qiang Wang, Hui Wang: A Fuzzy Feedback Control Real-Time Scheduling Algorithm, Journal of Software, 15:6 (2004) 791-798 (in Chinese)

A Coupled Fuzzy Logic Control for Routers' Queue Management over TCP/AQM Networks

Zhi Li and Zhongwei Zhang

University of Southern Queensland, Toowoomba, QLD, Australia

Abstract. Significant efforts in developing active queue management (AQM) in gateway routers in a TCP/IP network have been made since random early detection (RED) in 1993, and most of them are statistical based. Our approach is to capitalize on the understanding of the TCP dynamics to design an effective AQM scheme. In this paper, two FL-based AQM algorithms are proposed with the deployment of traffic load factor for early congestion notification. Extensive experimental simulations with a range of traffic load conditions have been done for the purpose of performance evaluation. The results show that the proposed two FLAQM algorithms outperform some well-known AQM schemes in terms of both user-centric measures and network-centric measures.

1 Introduction

The control design philosophy of the current Internet is that reliable data transfer must be provided by TCP operating at the endpoints, not in the network. As a result, end hosts deploy sophisticated TCP end-to-end flow and congestion avoidance algorithms, while routers in the network simply perform routing and forwarding, with which the routers accommodate each incoming packet unless the output buffer is full, and serve the packets in the buffer with FIFO order. However, under heavy-loaded traffic, the current technique of queue management, Drop Tail, causes not only high loss rate and low network throughput but also long packet delay and lengthy congestion conditions. Research shows that routers have the potential of knowing traffic conditions and thereby playing an active role in traffic control instead of operating reactively. As a result, the extension of the current queue management scheme to a method realizing traffic and congestion control has thus been widely discussed. Random Early Detection (RED) [1] is one of the first queuing schemes to realize active queue management (AQM) that aims to proactively drop packets in such a way that the congestion avoidance strategy of TCP works most effectively by means of detecting congestion at its early stage. However, the main problem of RED is *parameter tuning*. Due to the fact that Internet traffic is ever-changing or dynamic, this parameter tuning issue thus has been a major hurdle to the deployment of RED in reality. Efforts have been made in the auto-configuration of RED parameters such as ARED [2]. Meanwhile, the problem of queue management has been addressed from different standpoints. Some have addressed it in an empirical way such as

BLUE [3]. Some solutions regard it as an optimization problem such as REM [4]. A high proportion of work in this area is based on classical control theory such as PI [5].

In this paper, inspired by the success of fuzzy logic (FL) in the robust control of nonlinear complex systems, we apply artificial intelligence technologies, particularly FL, to design an effective AQM scheme, FL-based AQM or FLAQM. The rest of this paper is organized as follows. Section 2 will introduce the congestion indicators used in FLAQM. Section 3 presents two FLAQM algorithms, FLAQM and improved FLAQM. Section 4 provides the performance evaluation results of FLAQM compared to some well-known AQM schemes, while conclusions are given in Section 5.

2   TCP/AQM Modelling and Traffic Load Factor

Our hypothesis is that the Internet is one of the most complex systems. It is based on the fact of the nonlinear nature of TCP dynamics, and network complicity such as varying round-trip time (RTT) and different traffic patterns, and various kinds of TCP
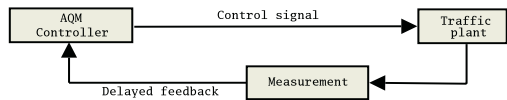
Fig. 1. Block diagram of a generic AQM control system

variants. Routers and endpoints cooperate to transmit data streams. Therefore, an AQM can be viewed as a controller, and the control object is the traffic plant representing the TCP/AQM system. Within the control-loop depicted in Figure 1, the AQM controller makes decisions on dropping probability of arrival packets based on the feedback from the traffic plant and then sends a control signal to the traffic plant informing it of the dropping probability. Note that valid feedback from the plant is delayed at least one RTT for each connection.

Precise detection of congestion levels is a key to the success of an AQM design. Most existing AQM schemes have deployed queue length or input rate, or both, for congestion indication. We choose *traffic load factor*, inspired by its successful use in a rate-based scheme for controlling ABR flows in ATM networks in [6]. The traffic load factor, denoted as z, is the ratio of input rate to target capacity, where target capacity or expected traffic input rate is the leftover link capacity after draining the remaining packets in an output buffer.

3   Fuzzy Logic Coupled Controller

Observing the TCP dynamics and congestion situations, we can see TCP/AQM loop feature. If traffic load is higher, the buffer is likely full, the possibility of dropping packets is high. However, if traffic load is low, even the buffer is full, the likelihood of dropping packets is not high. Based on these observations, we would be able to design a new AQM which coupled two FL controllers (FLCs). Two FLAQM algorithms are proposed in this study, namely FLAQM and improved

FLAQM. FLAQM directly uses traffic load factor z and its change Δz as inputs, whereas improved FLAQM inputs the values of the reciprocal of z, z' and its change $\Delta z'$.

3.1   Structure of FLAQM

It is clear that the set-point for the measured plant output in FLAQM, z, is 1 in that the input rate equals the target link capacity. Thus, the steady-state operating region toward which FLAQM attempts to drive the closed-loop feedback system is in the neighborhood of $z = 1$, $[1, 1+\delta]$, where δ is a constant. The FLAQM controller copes with three cases as shown in
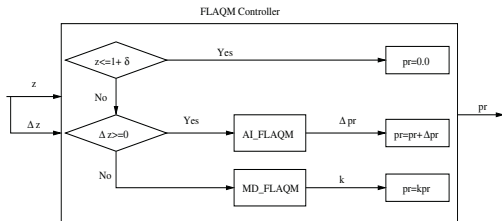
Fig. 2. The structure of FLAQM

Figure 2. If load factor z is beyond the set-point of the system, multiplicative decrease (MD) action is taken with negative Δz by using the MD_FLAQM controller, while additive increase (AI) is applied with positive Δz by using the AI_FLAQM controller. Otherwise, no changes are applied to the current values of pr.

Membership functions for both FLCs: MD_FLAQM and AI_FLAQM are shown in Figure 3. Note that the *sup-min* FL inference and the *center of gravity* defuzzification methods are adopted. Tables 1 and 2 show the fuzzy if-then rules in MD_FLAQM and AI_FLAQM respectively.

3.2   Discussion About FLAQM

We can improve FLAQM by using different traffic variables and control rules. For instance, improved FLAQM uses z' and $\Delta z'$ as process states and the dropping probability pr as control output. The improved FLAQM controller also consists of two FLCs, MD_FLAQM and AI_FLAQM which are activated when $z' < 1+\delta$, where δ is a constant.

For the MD_FLAQM controller, scaling or normalization of input variables has been implicitly achieved. The inputs z' and $\Delta z'$ are limited in the range of $[0, 1+\delta]$. For the AI_FLAQM controller, however, the input $\Delta z'$ could vary from negative infinite to zero despite the effective universe of discourse of z' in the range of $[0, 1+\delta]$. We argue that $\Delta z' < -(1+\delta)$ is a rather rare event as proved in the simulations. In addition, although it really happens, the previous z' must be larger than $1 + \delta$ and the previous control value of dropping probability pr is zero. In this case, caution is needed to additively increase pr to accommodate bursty traffic and maintain high link utilization. Therefore, in the controller, we take the same action for any value of $\Delta z'$ less than $-(1+\delta)$ with $\Delta z' = -(1+\delta)$. Due to space limit, we will not show the graphs of the membership functions and fuzzy rules of improved FLAQM.

(a) z in MD_FLAQM and AI_FLAQM

(b) Δz in MD_FLAQM

(c) k in MD_FLAQM

(d) Δz in AI_FLAQM

(e) Δpr in AI_FLAQM
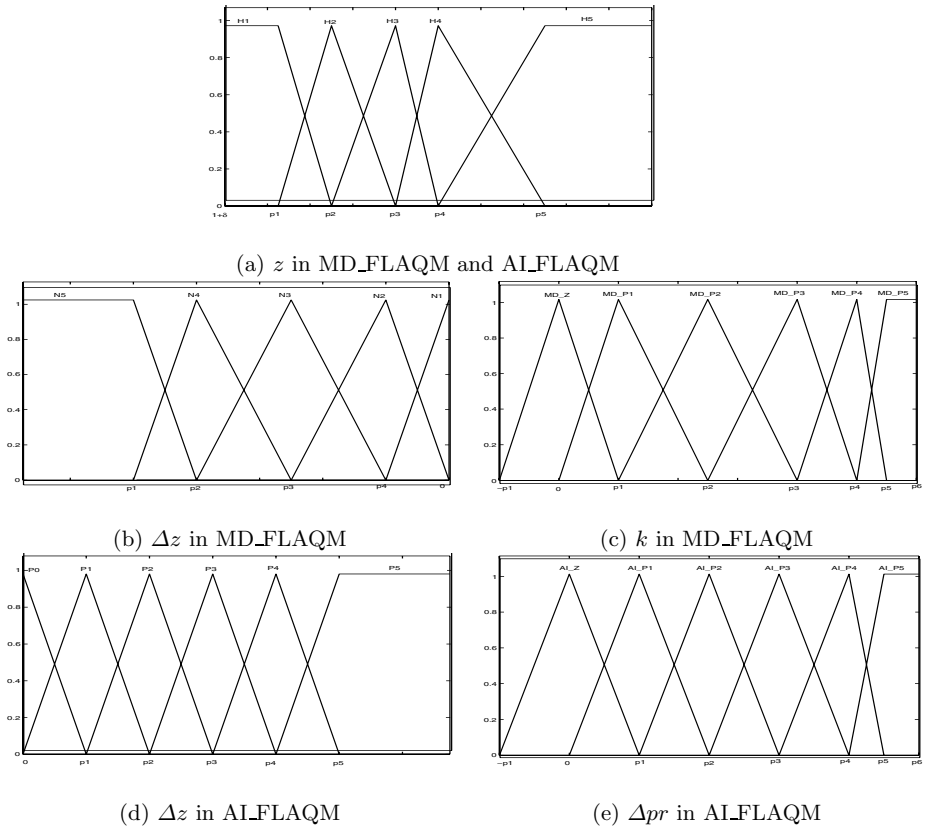
Fig. 3. Membership functions in FLAQM

Table 1. Rules of MD_FLAQM in FLAQM

$\Delta z/z$	H1	H2	H3	H4	H5
N5	MD_P0	MD_P1	MD_P2	MD_P2	MD_P3
N4	MD_P1	MD_P2	MD_P3	MD_P3	MD_P4
N3	MD_P2	MD_P3	MD_P3	MD_P4	MD_P5
N2	MD_P3	MD_P3	MD_P4	MD_P5	MD_P5
N1	MD_P4	MD_P4	MD_P5	MD_P5	MD_P5

Table 2. Rules of AI_FLAQM in FLAQM

	H1	H2	H3	H4	H5
P0	AI_Z	AI_Z	AI_P1	AI_P1	AI_P2
P1	AI_Z	AI_P1	AI_P2	AI_P2	AI_P3
P2	AI_P1	AI_P2	AI_P2	AI_P3	AI_P4
P3	AI_P1	AI_P2	AI_P3	AI_P3	AI_P4
P4	AI_P2	AI_P2	AI_P3	AI_P4	AI_P5
P5	AI_P2	AI_P3	AI_P4	AI_P4	AI_P5

4 Applying FLAQM on a Simulated Network

We investigate the performance of the proposed FLAQM algorithms compared with that of the traditional Drop Tail (or DT), RED [1], and ARED [2] via simulations conducted on the NS2 network simulator [7].

4.1   Network Topology and Traffic Pattern

A dumbbell network topology as depicted in Figure 4 is used to simulate a network where a bottleneck link lies between a premise in the network client side and an edge router in an ISP side as in reality. As shown in Figure 4, the bottleneck link is from $R2$ to $R3$ with bandwidth of 1.5Mbps.
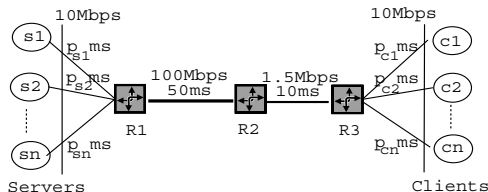
Fig. 4. Network topology

We apply different queue management schemes in the output queue of router $R2$ towards $R3$, while using Drop Tail elsewhere. Reno TCP is used with the buffer size of $R2$ set as 160 packets (1000 bytes per packet). In addition, all the traffic experiences the same propagation delay on the links from the server side to the client side with 40ms from the serves to $R1$, and 1ms from $R3$ to the clients.

In the simulations, traffic is generated to cause congestion on the bottleneck in the selected network topology. Traffic from server s1 to client c1 models Web service. The traffic is generated by the Poisson-Pareto traffic model[1] with the truncated threshold of Pareto distribution of 1000 packets. On the other hand, the traffic from the other servers simulates extremely long bursts which last the whole simulation period, and the number of such servers is varied ranging from 5 to 40 to create different traffic conditions. Note that the simulation duration is 1000s and the first half is set as a warmup period.

4.2   Simulation Results

Compared with Drop Tail, RED, and ARED, our proposed FLAQM (FLAQM(I)) and improved FLAQM (FLAQM(II)) demonstrate their robustness in a various traffic load conditions as follows.

– The performance comparison in terms of user-centric measures, including user goodput[2] and response time[3], is shown in Figure 5. Weighted average

---

[1] Flow arrivals form a Poisson process and flow lengths are distributed according to a Pareto distribution.

[2] The ratio of the amount of the packets received by the destination, excluding duplicated packets, to the time spent.

[3] The time spent to complete a response; namely, the elapsed time interval from the server sending out the first packet of a response to the client receiving the last packet of the response.

is calculated by taking account of the number of flows for a given flow size. Figure 5 (a) shows that improved FLAQM outperforms the others in terms of TCP goodput, while FLAQM achieves goodput performance comparable to ARED, and the goodput performance of Drop Tail is far more behind the others. Figure 5 (b) shows that Drop Tail obtains the highest weighted average flow latency in all the investigated traffic conditions and also flow latency increases dramatically with the number of extremely long bursts compared to those of the AQM schemes. Among the AQM schemes, improved FLAQM still performs the best in terms of flow latency, while FLAQM and ARED are comparable.
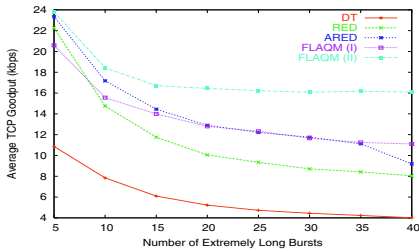
(a) Weighted average user TCP goodput comparison
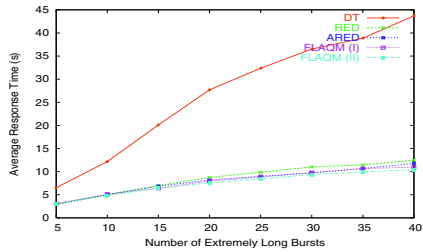
(b) Weighted average user response time comparison

Fig. 5. User performance

− The network throughputs of FLAQM and improved FLAQM are the best as plotted in Figure 6, with almost 100% link utilization for all the schemes including Drop Tail.
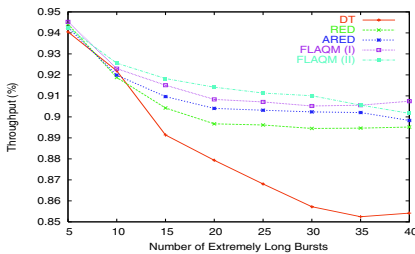
Fig. 6. Network throughput versus number of FTP extremely long bursts
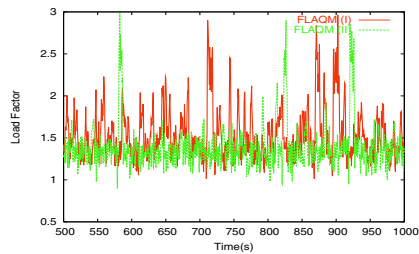
Fig. 7. Load factor comparison between FLAQM and improved FLAQM with n=20

On the other hand, not only is the queue dynamics of two FLAQM algorithms steady compared to that of RED and ARED, but also improved FLAQM is capable of quickly driving the system towards the set-point of traffic load factor

z without much oscillation compared to FLAQM as shown in Figure 7 with the dynamics of traffic load factor z when n=20.

Due to its achievement of better performance, improved FLAQM is set as a default option for the FLAQM controller. Note that we also have conducted simulations with changed traffic load conditions, which also prove the capability of FLAQM, especially improved FLAQM. Due to space limit, we omit the simulation results.

5   Conclusions

In this paper, we have applied FL to implement our knowledge-based control on the complex nonlinear TCP/AQM system. Two algorithms, FLAQM and improved FLAQM, have been presented. Performance evaluation of the proposed schemes is carried out via extensive experimental simulations under a variety of traffic conditions against a number of performance measures: TCP goodput and flow latency from the user perspective and network throughput, link utilization, and loss rate from the ISP perspective. By applying the proposed FLAQM scheme to queue management, the simulation results show that the traffic performance has been improved dramatically compared to that of both Drop Tail and some well-known AQM schemes in all the investigated network circumstances. More importantly, the fuzzy controllers, in particular improved FLAQM, are robust in the face of traffic fluctuation and different network conditions. It thus demonstrates the feasibility of applying FL to the area of traffic control in the Internet. Autoconfiguration of FLAQM and its improved version by coupling neural networks (NNs) and FL is our future work in order to obtain better control.

References

1. Floyd, S., Jacobson, V.: Random Early Detection Gateways for Congestion Avoidance. IEEE/ACM Transactions on Networking (1993), Vol. 1(4), 397–413
2. Floyd, S., Gummadi, R., Shenker, S.: Adaptive RED: An Algorithm for Increasing the Robustness of RED. Technical report (2001).
   http://citeseer.nj.nec.com/floyd01adaptive.html
3. Feng W.: BLUE: A New Class of Active Queue Management Algorithms. Tchnical report (1999). http://citeseer.nj.nec.com/article/feng99blue.html
4. Athuraliya, S. et al: REM: Active Queue Management. IEEE Network (2001), Vol. 15(3), 48 – 53
5. Hollot, C.V. et al: On Designing Improved Controllers for AQM Routers Supporting TCP Flows. Proceedings of IEEE INFOCOM (2001), 1726-1734
6. Kalyanaraman, S. et al: The ERICA Switch Algorithm for ABR Traffic Management in ATM Networks. IEEE/ACM Transactions on Networking (200), Vol. 8(1), 87–98
7. The Network Simulator - ns-2. http://www.isi.edu/nsnam/ns/

Iris Pattern Recognition Using Fuzzy LDA Method

Hyoun-Joo Go, Keun-Chang Kwak, Mann-Jun Kwon, and Myung-Geun Chun[*]

School of Electrical and Computer Engineering,
Chungbuk National University, Cheongju, Korea
mgchun@chungbuk.ac.kr

Abstract. This paper proposes an iris pattern recognition algorithm as one of biometric techniques applied to identify a person using his/her physiological characteristics. Since the iris pattern of human eye has an unique and invariant texture, we can use it as a biometric key. First, we obtain the feature vector from the fuzzy LDA after performing 2D Gabor wavelet transform. And then, we compute the similarity measure based on the correlation. Here, since we use four matching values obtained from four different directional Gabor wavelets and select the maximum value among them, it is possible to reduce the recognition error. To show the usefulness of the proposed algorithm, we applied it to an iris database consisting of 300 iris patterns extracted from 50 subjects and finally got more higher than 90% recognition rate.

1 Introduction

Recently, the electronic commerce has been popular according to the rapid spread of Internet and now we are living in a networked society. This tightly networked society rises serious problems related to the information security to detect imposters and protect genuine users. One of the most conventional methods for system security is using password, which is very simple and does not require any special device. However, it can be easily divulged to others. So, we need more safe verification and authentication methods to prove user's identity. One of the most competitive technologies to meet this requirement is a biometric method. Biometrics uses personal physiological and behavior characteristics to identify a person. There are various biometric researches based on the fingerprint, face, retina, iris, voiceprint, and signature.

The fingerprint has been used for various identifications in real life since it is well known that its pattern is unique and permanent for all life. However, most of people are not favor for this technique because fingerprint recognition has been used for forensic affairs. On the other hand, face recognition shows high acceptance but has low performance due to make-up and aging process. The retinal vasculature is rich in structure and is supposed to be unique for each person. To scan a retinal pattern, however, one should peer into an eye-piece and focus on a specific spot so that a predefined part of the retinal vasculature could be imaged. Though high security can be obtained from this technology, it is not popular because of user's discomfort and fear. On the other hand, there is an iris pattern in front of eyeball that exists between cornea and crystalline lens. Obverse of iris has irregular rudiments and circular pattern that is near pupillary margin. It is known that iris pattern is unique and permanent such as fingerprint[1][2]. Figure 1 shows various iris patterns from 8 different subject.

---

[*] Corresponding Author

Daugman studied an iris pattern recognition method using 2-D Gabor filter and Boles used the wavelet transform[3][4] to extract feature vectors. However, in case of Daugman's method, the iris code is 256 bytes which is rather large template size. Also, since the feature vector of Boles's method is not rich, it cannot be applied to larger population. The organization of this paper is as follows. In Section II, some properties of the iris and Gabor filter are briefly introduced. And then we describe FLDA(fuzzy linear discriminant analysis) and experiment results in Section III. Finally, some concluding remarks are given in Section IV.
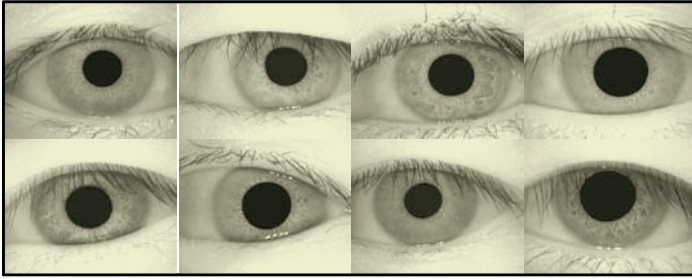
Fig. 1. Iris pattern images for eight different subjects

2   Feature Extraction of an Iris Pattern

2.1   Preprocessing of Iris Images

At the acquisition stage of iris images, the eyebrows hide some part of pupil and also the iris image is often bulged by make-up and lighting. These cause difficulties to find the pupil of iris and center of pupil. Therefore, the iris image was preprocessed by a median filter to remove noises causing by the eyebrows and eyelid, and so on. Here, we implemented a segmentation algorithm to extract the pupil by using the binarization process based on the threshold value computed from the histogram. Consequently, we could find the center of pupil from binary image by using the projection method and the component labeling algorithm. From this pupil's center, we can calculate the radius and circumference of the pupil. Figure 2 (a) shows the binary image after detecting the pupil and Figure 2 (b) shows a gray level image after applying histogram equalization.
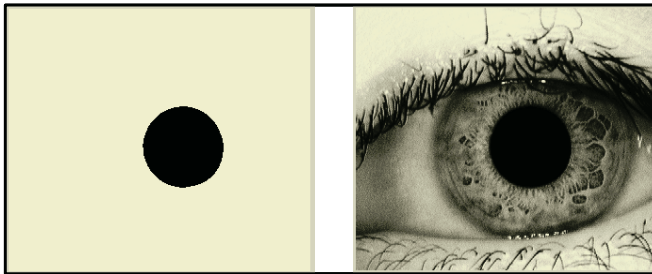
(a) Histogram equalized image   (b) Binary image of pupil

Fig. 2. Preprocessed iris images

2.2   Pattern Extraction Using the Gabor Wavelet

The Gabor wavelets, which capture the properties of spatial localization, orientation selectivity, spatial frequency selectivity and quadrature phase relationship, seem to be a good approximation to the filter response profiles encountered experimentally in cortical neurons. The Gabor wavelets have been found to be particularly suitable for image decomposition and representation when the goal is the derivation of local and discriminating features.

The Gabor wavelets are used for image analysis because of their biological relevance and computational properties, whose kernels are similar to the 2-D receptive field profiles of the mammalian cortical simple cells, exhibit strong characteristics of spatial locality and orientation selectivity, and are optimally localized in the space and frequency domains. The Gabor wavelets can be defined as follows[5].

$$W(x, y, \theta, \lambda, \varphi, \sigma, \gamma) = e^{\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} \cos(2\pi \frac{x'}{\lambda} + \varphi) \qquad (1)$$

$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta$$

where θ specifies the orientation of the wavelet and it rotates the wavelet about its center. The orientation of the wavelets dictates the angle of the edges or bars for which the wavelet will respond. In most cases θ is a set of values from 0 to π. The λ specifies the wavelength of the cosine wave, or inversely the frequency of the wavelet. Wavelets with a large wavelength will respond to gradual changes in intensity in the image. The φ specifies the phase of the sinusoid. Typically Gabor wavelets are based on a sine or cosine wave. In the case of this algorithm, cosine wavelets are thought to be the real part of the wavelet and the sine wavelets are thought to be the imaginary part of the wavelet. The σ specifies the radius of the Gaussian. The size of the Gaussian is sometimes referred to as the wavelet's basis of support. The Gaussian size determines the amount of the image that effects convolution. The γ specifies the aspect ratio of the Gaussian. This parameter was included such that the wavelets could also approximate some biological models. For illustration, Figure 3 shows iris images through a Gabor filter of four directions, $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$
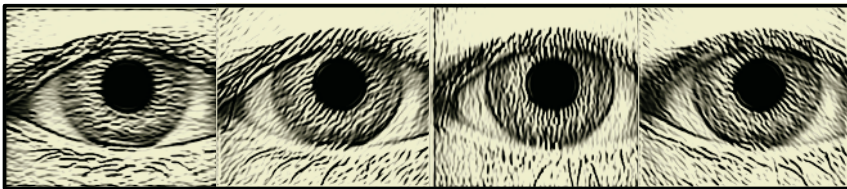
Fig. 3. Iris images after Gabor filtering

After applying the Gabor filter, we performed so-called polar mapping to extract the region of interest(ROI) for the filtered iris image shown in Figure 3. Since we performed the process for four types of Gabor filter ($\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$), we could get four types of polar mapped images as shown in Figure 4. At the first step, we extracted a part of iris pattern from a predefined inner radius to an outer radius,

which looks like doughnut. Here, we should consider the variation of the radius of a pupil for lighting condition. If we get the iris image under bright condition, then the size of pupil becomes smaller. On the contrary, if we get the iris image under dark condition, then the   size of pupil becomes larger. These mean the variation of the size of ROI. So, the recognition algorithm should be robust to endure this variation. This is our motivation to adopt the fuzzy concept in our recognition algorithm. After performing resize process with a linear interpolation, we got the 40*500 polar mapped images.
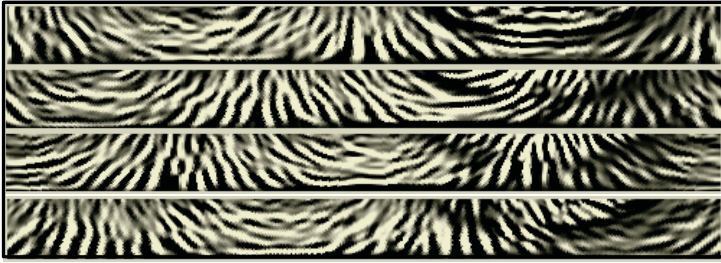
Fig. 4. Polar mapped iris pattern

3   Fuzzy-Based Linear Discriminant Classifier for Iris Recognition

The Linear Discriminant Analysis (LDA) is used to find optimal projection from feature vectors. Rather than finding a projection that maximizes the projected variance, LDA determines a projection, $V = W_{FLD}^{T} X$ (W_{FLD}^{T} is the optimal projection matrix), that maximizes the ratio between the between-class scatter and the within-class scatter matrix. One the other hand, the fuzzy-based LDA method assigns feature vectors to fuzzy membership degree based on the quality of training data[6][7].

In this work, we used two iris database so-called CASIA and CBNU. CASIA(Chinese Academy of Science) database[8] contains eight iris images for one hundred and eight subjects. Among them, we used six iris images for fifty individuals. On the other hand, the CBNU (Chungbuk National University) database contains six iris images for fifty individuals. Thus, the total number of iris images becomes three hundreds for each database and the image size of an iris images is 280*360 having gray levels ranged between 0 and 255. Among them, 150 images were used for training and others for testing.

As described earlier, the iris images were transformed into time-scale space by using the Gabor wavelet to enhance the iris pattern. And then, feature extraction is performed by fuzzy-LDA after reducing the dimensionality of original data space by PCA.

In this paper, we used the correlation value to establish the classifier instead of conventional Euclidean based k-NN classifier. The correlation coefficients are calculated as follows[9].

$$\rho_{x,y} = \frac{\text{cov}[X,Y]}{\sigma_X \sigma_Y} = \frac{E[XY] - \mu_X \mu_Y}{\sigma_X \sigma_Y}, \left| \rho_{X,Y} \right| \le 1 \qquad (2)$$

where the correlation coefficient $\rho_{x,y}$ is a measure of the correlation between the feature vectors X and Y (actually these are LDA coefficients). For an iris image, we calculate 4 correlation coefficients for each polar mapped image after performing the Garbor wavelet transform. And then, we compute a total matching score by adding the computed correlation coefficients. The final decision is made by taking one having the maximum score.
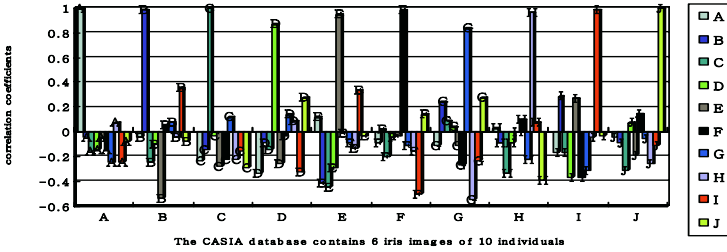
Fig. 5. Correlation coefficients by using the Fuzzy-LDA (CASIA)

As you can see in Fig. 5, the fuzzy LDA shows better recognition results. That is, the difference between the highest matching score and the second score for a subject is always large. We performed the iris pattern recognition under several conditions of changing the numbers of eigenfaces and fisherfaces for CASIA database as shown in Fig. 6. The PCA method shows the recognition rates between 83% and 86%. The LDA shows recognition rates between 91.3% and 92.7%. On the other hand, the Fuzzy LDA method shows recognition rates between 92.7% and 94.7%. For the reference, we also performed ICA(Independent Component Analysis)[10] and got the recognition rate between 90.7% and 92.0%. We can conclude that our proposed method shows better results than others at most cases.
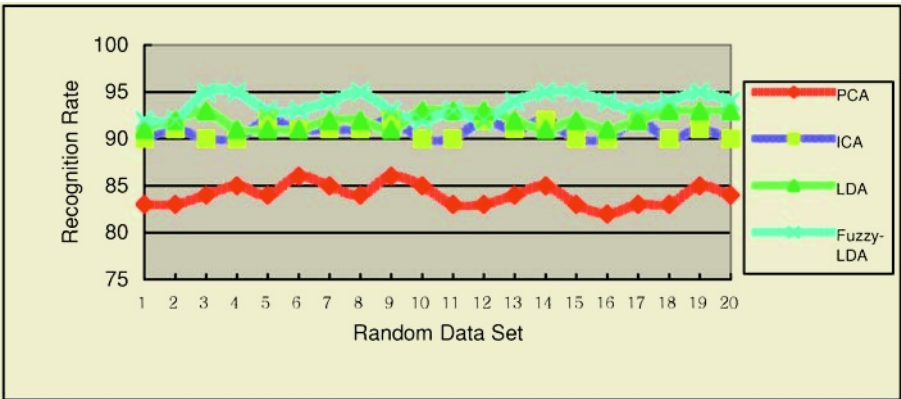
Fig. 6. Recognition rates for various feature extraction methods(CASIA case)

The same experiments are carried out for the CBNU database. Fig. 7 shows overall recognition rates. Among them, the proposed method has the highest recognition rate ranges from 96.0% to 97.33%.
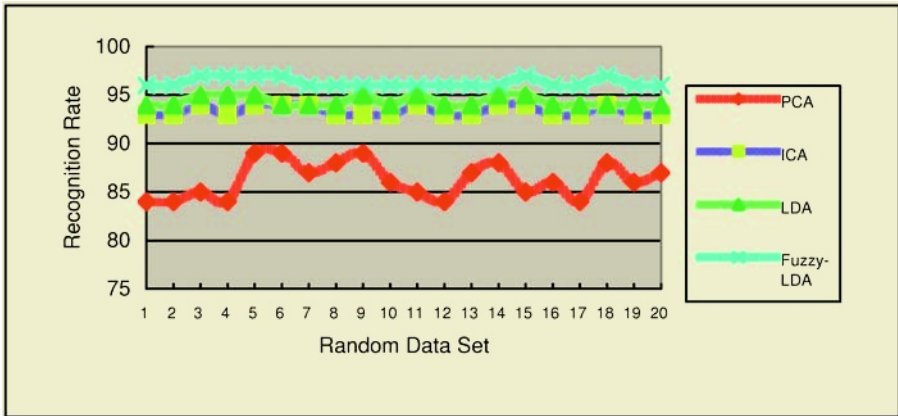
Fig. 7. Recognition rates for various feature extraction methods (CBNU case)

4 Concluding Remarks

This paper dealt with iris pattern recognition as one of biometric techniques which are applied to identify a person using his/her congenital characteristics. First, we obtain the feature vector from the fuzzy LDA after performing 2D Gabor wavelet transform. And then, we compute the similarity measure based on the correlation. Here, since we use four matching values obtained from four different directional Gabor wavelets and select the maximum value among them, it is possible to have robust recognition results. Through various experiments using CASIA and CBNU iris database, we found that the proposed algorithm shows more favorable recognition results than others.

Acknowledgements

References

1. Fasel, I.R, Bartlett, M.S, Movellan, J.R, "Automatic Face and Gesture Recognition", Proceedings, Fifth IEEE International Conference on, 2002.
2. Abdel Alim, O, Sharkas, M, "Texture Classification of the Human Iris using Artificial Neural Networks", Electrotechnical Conference, 2002.
3. J. G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression", IEEE Trans, on Acoustics, Speech, and Signal Processing, Vol. 36, No.7, pp.1169-1179, 1988.
4. W. W. Boles and B. Boashash, "A Human Identification Technique Using Images of the Iris and Wavelet Transform", IEEE Trans, on Signal Processing, Vol. 46, No. 4, pp.1185-1188, 1998.
5. N. Petkov and P. Kruizinga, Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: Bar and grating cells, 1997, pp. 83-96.

6. Keun-Chang Kwak and Witold Pedrycz, Face recognition using fuzzy Integral and wavelet decomposition method, Systems, Man and Cybernetics, Part B, IEEE Transactions on, Vol. 34, No. 4, pp.1666 – 1675, 2004

7. Dae Jong Lee, Keun Chang Kwak, Jun Oh Min, and Myung Geun Chun, Multi-modal biometrics System Using Face and Signature, ICCSA, LNCS 3043, pp.635-544, 2004

8. CASIA Iris Image Database, Chinese Academy of sciences, http://www.sinobiometrics.com/casiairis.htm

9. Childers, Probability And Random Process, Prentice-Hall, 1997

10. M. Bartlett, Face Image Analysis by Unsupervised Learning and Redundancy Reduction, PhD thesis, Univ, of California, 1998

Precision Tracking Based-on Fuzzy Reasoning Segmentation in Cluttered Image Sequences

Jae-Soo Cho[1], Byoung-Ju Yun[2], and Yun-Ho Ko[3]

[1] School of Internet-Media Engineering
Korea University of Technology and Education, Cheonan, South Korea
jaesoo27@kut.ac.kr
[2] Dept. of Information and Communication, Kyungpook National University
Daegu, South Korea
bjisyun@ee.knu.ac.kr
[3] Dept. of Mechatronics Engineering, Chungnam National University
Daejeon, South Korea
koyh@cnu.ac.kr

Abstract. In our previous work [7], we presented a robust centroid target tracker based on new distance features in cluttered image sequences. A real-time adaptive segmentation method based on new distance features was proposed for the binary centroid tracker. The target classifier by the Bayes decision rule for minimizing the probability error should properly estimate the state-conditional densities. In this correspondence, the proposed target classifier adopts the fuzzy-reasoning segmentation using the fuzzy membership functions instead of the estimation of the state-conditional probability densities. Comparative experiments also show that the performance of the proposed fuzzy- reasoning segmentation is superior to that of the conventional thresholding methods. The usefulness of the method for practical applications is demonstrated by considering two sequences of real target images. The tracking results are good and stable without difficulty of the estimation.

1 Introduction

Tracking algorithms include many different methods of target location estimation which have been developed. The most common (and probably the best known) of these in the military applications (fire control and guidance) is the centroid tracker [1]-[4], which determines a target aim point by computing the geometric or intensity centroid of the target object based on a target segmentation method.

Many kinds of segmentation algorithms have been developed in various application areas. The problem of segmenting an object from a given field of view is basic to many image processing applications. Examples of such applications include medical ones, pattern recognition, or military targets in infrared and visible images, etc.. There may not be an universal segmentation algorithm in all kinds of applications of image segmentation. An appropriate segmentation algorithm should be selected or developed for a particular application.

A segmentation-based tracker needs to extract a target from the background in the tracking window [7]. Image thresholding is one of the most common applications in image analysis. Among the image thresholding methods, bilevel thresholding separates the pixels of an image into two regions (i.e., the object and the background). The previous methods of target segmentation [1]-[4] in the military tracking applications are a two-step process. First, the original gray scale image is statistically analysed and transformed to a binary image by the rule of the hard limiter through a determined intensity threshold:

$$\beta_i = \begin{cases} 1, & I_i \ \geq \ Threshold \ (or \ I_i \ \leq \ Threshold) \\ 0, & otherwise, \end{cases} \tag{1}$$

where β_i denotes the ith segmented pixel value and I_i an intensity value.

2   Target Segmentation Based on Fuzzy Reasoning

In the military targets of infrared images, the problem of segmentation is to partition the target window into two regions, where all the points in one region corresponding to target, and all points in the other corresponding to background. Due to the limit of real-time application in the field of target tracking, few methods have been employed to extract the target in cluttered images. Usually, simple algorithms which determine a threshold of hard limiter have been used for this kind of segmentation-based tracking algorithm [7].

Although the centroid tracker works well in some tracking environments, it performs poorly in others. The limitations of the centroid tracking stem fundamentally from its inability to exploit fully the target signature information present in the image. The moving target segmentation which separates the target from the background images is a critical step in the centroid tracker.

In the previous studies [7], we proposed new features in addition to the conventional intensity feature in order to segment effectively a moving target in the cluttered background images. These novel features were distances between the predicted center pixel of a target object by a tracking filter and each pixel in extraction of a moving target. The proposed centroid target tracker is depicted in Figure 1 (a).

Allowing of the use of more than one intensity feature merely requires replacing the scalar feature z by the feature vector $Z = [z_1 \ z_2 \ z_3 \ \cdots]^T$ which denotes the distinctive features to identify the target pixels from the background pixels. Our pixel classifier model uses three features, $Z = [z_1 \ z_2 \ z_3]^T$ for classifying a target: pixel intensity z_1, distance metric values, z_2 and z_3, which can be defined as:

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \equiv \begin{bmatrix} I(x(i), y(i)) \\ x(i) - \hat{x}(k|k-1) \\ y(i) - \hat{y}(k|k-1) \end{bmatrix}, \quad 1 \leq i \leq N \tag{2}$$

where $(\hat{x}(k|k-1), \ \hat{y}(k|k-1))$ is a predicted center pixel of a moving target at the kth frame conditioned upon $k-1$ image sequences $(1, 2, \cdots, k-1)$, and N
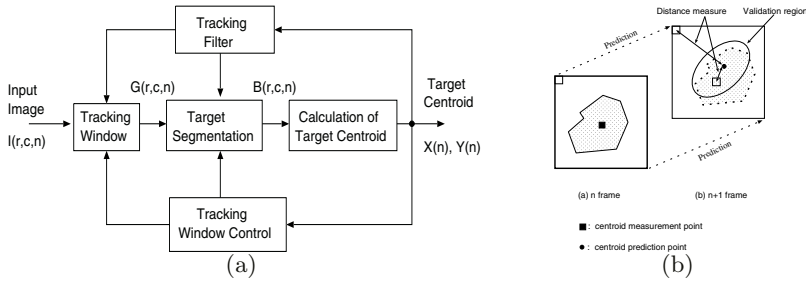
Fig. 1. (a) Block diagram of the proposed centroid tracker, (b) Schematic diagram of prediction of tracking window

denotes the total number of pixels in the target window. $I(x(i), y(i))$ means the gray intensity value in the ith pixel $(x(i), y(i))$.

Figure 1 (b) shows the schematic diagram of tracking window prediction of the kth frame based on the $k-1$ previous image frames. The distance metric features are determined by two dimensional characteristics of a target, and sensor noises.

We let ω denote the state of a pixel, with $\omega = \omega_1$ for the target pixel and $\omega = \omega_2$ for the background pixel in the segmentation region. Because the state of a pixel is unpredictable, we consider ω to be a random variable. We also assume that *a priori* probabilities, $P(\omega_1)$ and $P(\omega_2)$, are known. Let $p(Z|\omega_j)$ be the state-conditional probability density function for the feature vector Z given that the state of nature is ω_j.

Suppose that we know both the *a priori* probabilities $P(\omega_j)$ and the conditional densities $p(Z|\omega_j)$. The gray level image is transformed into a binary image, with a new binary value $B(x(i), y(i))$ in pixel $(x(i), y(i))$, by the Bayes decision rule for minimizing the probability of error:

$$Decide \quad B(x(i), y(i)) = \begin{cases} 1, & if \quad p(\omega_1|Z) > p(\omega_2|Z) \\ 0, & otherwise. \end{cases} \tag{3}$$

We can express the rule in terms of the conditional and *a priori* probabilities.

$$Decide \ B(x(i), y(i)) = \begin{cases} 1, & if \quad p(Z|\omega_1) \cdot P(\omega_1) > p(Z|\omega_2) \cdot P(\omega_2) \\ 0, & otherwise. \end{cases} \tag{4}$$

If we assume that the feature variables z_1 (pixel intensity), z_2 and z_3 are independent each other, we compute the conditional probabilities using the estimated probabilities as follows:

$$p(Z|\omega_1) = p(z_1|\omega_1) \cdot p(z_2|\omega_1) \cdot p(z_3|\omega_1), \tag{5}$$
$$p(Z|\omega_2) = p(z_1|\omega_2) \cdot p(z_2|\omega_2) \cdot p(z_3|\omega_2). \tag{6}$$

The computation of *a posteriori* probabilities $p(\omega_j|Z)$ and the *a priori* probabilities lie at the heart of the Bayesian classification. In the previous study [7], we

proposed a proper adaptive estimation technique of these probability functions. In this paper, we will use the fuzzy membership function for the state-conditional probability density functions instead of estimation as followings:

$$p(z_i|\omega_j) \equiv fm(z_i|\omega_j), \quad i = 1, 2, 3 \; and \; j = 1, 2. \tag{7}$$

where $fm(z_i|\omega_j)$ means the fuzzy membership function for the feature z_i given that the state of nature is ω_j.

Fuzzy set theory has been applied to many areas of image processing including the image thresholding only by using the intensity feature [5]. The fuzzy membership function which represents the degrees of belonging of each region for the two classes of pixels fully reflects the nature of the state-conditional probability density function.

Let $h(z_1)$ denote the number of occurrences at the gray level z_1 in the target window image. Given a determined threshold value T, the average gray levels and the class variances of the target and the background can be, respectively, obtained as follows:

$$\mu_1 = \sum_{z_1=T}^{L-1} z_1 \cdot h(z_1) \Big/ \sum_{z_1=T}^{L-1} h(z_1), \quad \mu_2 = \sum_{z_1=0}^{T-1} z_1 \cdot h(z_1) \Big/ \sum_{z_1=0}^{T-1} h(z_1). \tag{8}$$

and

$$\sigma_1^2 = \frac{\sum_{z_1=T}^{L-1}(z_1 - \mu_1)^2 \cdot h(z_1)}{\sum_{z_1=T}^{L-1} h(z_1)}, \quad \sigma_2^2 = \frac{\sum_{z_1=0}^{T-1}(z_1 - \mu_2)^2 \cdot h(z_1)}{\sum_{z_1=0}^{T-1} h(z_1)} \tag{9}$$

where L means the overall gray level. The average gray levels, μ_1 and μ_2, can be considered as the reference intensity values of the target and the background for the given threshold value T.

The membership functions between a pixel and its belonging region should intuitively depend on the difference of its gray level from the reference value of its belonging region. The membership functions which evaluate the belonging relationship for a pixel can be defined as:

$$fm(z_1|\omega_1) \equiv \frac{1}{\sqrt{2\pi\sigma_1^2}} \; exp \; [-\frac{1}{2}(\frac{z_1 - \mu_1}{\sigma_1})^2], \tag{10}$$

$$fm(z_1|\omega_2) \equiv \frac{1}{\sqrt{2\pi\sigma_2^2}} \; exp \; [-\frac{1}{2}(\frac{z_1 - \mu_2}{\sigma_2}))^2], \tag{11}$$

where $fm(z_1|\omega_1)$ and $fm(z_1|\omega_2)$ denote the exponential fuzzy membership functions of the target and the background for the intensity feature in a given pixel, respectively. The fuzzy membership functions are shown in Figure 2 (a) for the intensity feature z_1.

Next, we should define the fuzzy membership values of the distance metric features for each class. Consider a target that is in track, i.e., its filter has at least been initialized. One may define statistical membership values in the measurement space, as shown in Figure 2 (b) for the distance metric features between the predicted center point, $(\hat{x}, \hat{y})$ and a classified pixel point, $(x(i), (y(i))$.
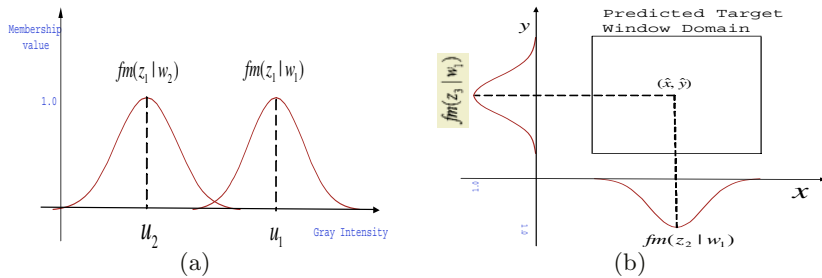
Fig. 2. Membership functions for each feature variable, (a) intensity feature(z_1), (b) distance metric features(z_2, z_3)

$$fm(z_2|\omega_1) \equiv \frac{a_1}{\sqrt{2\pi\sigma_x^2}} \, exp \, [-\frac{1}{2}(\frac{z_2}{\sigma_x})^2] = \frac{a_1}{\sqrt{2\pi\sigma_x^2}} \, exp \, [-\frac{1}{2}(\frac{x(i) - \hat{x}}{\sigma_x})^2], \quad (12)$$

$$fm(z_3|\omega_1) \equiv \frac{a_2}{\sqrt{2\pi\sigma_y^2}} \, exp \, [-\frac{1}{2}(\frac{z_3}{\sigma_y})^2] = \frac{a_2}{\sqrt{2\pi\sigma_y^2}} \, exp \, [-\frac{1}{2}(\frac{y(i) - \hat{y}}{\sigma_y})^2], \quad (13)$$

and,

$$fm(z_2|\omega_2) \equiv 1 - fm(z_2|\omega_1), \quad fm(z_3|\omega_2) \equiv 1 - fm(z_3|\omega_1), \quad\quad (14)$$

where the scalars a_1 and a_2, and the variances σ_x and σ_y, are able to be properly determined according to the size of the target window. Equations (12), (13) and (14), denote the membership degrees of the distance metric features, z_2 and z_3, for the target and the background, respectively.

3 Experimental Results

Moving targets are generally tracked after detecting moving objects, that is also named as the initial step of a tracking algorithm. The initial step may include a recognition process of wanted targets, or other initial step for each tracking algorithm. Two real image sequences are taken by an active Forward-Looking InfraRed (FLIR) system undergoing different types of motion. The first image sequence (referred to tank image) is composed of 148-frame taken by 15 frame/sec with 640×480 dimensions. The second sequence (man image) consists of 200-frames. Each sample frame is shown in Figure 3.

We used a Kalman filter [7] as the tracking filter which is a nearly constant velocity model as the dynamics model. The Kalman filter is used to obtain the state estimates for the tracking filter. To validate the usefulness and the tracking accuracy, the instantaneous tracking error (Terror(k)) used in the study [7] is also used in this paper. For each image sequence, we manually determined a real target center point on the target object and its position tracked over the entire sequences for a reference as precise as possible, similarly in [6], [7].

We compared the performance of the proposed segmentation algorithms with that of the conventional segmentation-based tracking methods, similarly in [7].

Fig. 3. Sample images, (a): a tank image, (b): a man image
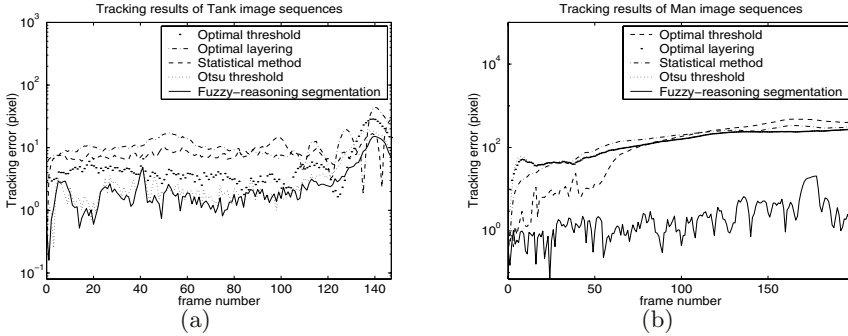
Fig. 4. Tracking results, (a): The tank image, (b): The man image

Considering the man image, the target intensity of a man is much correlated with the background one compared with the tank image. Even though the whole image sequences are not presented in this paper, it is very difficult to track the man target due to the correlated background images by using the conventional methods.

Figure 4 (a) and (b) show the instantaneous tracking errors for the tank and the man images respectively. As shown in the comparative simulation results, even though most of the applied methods have the reasonable tracking results for the tank image, the conventional tracking methods can not tract the man target from the cluttered background for the man sequences.

Figure 5 and Figure 6 show some segmentation results of the tank and the man images. The comparative experiment results show that the fuzzy-reasoning segmentation which has the additional feature (distance metric from the tracking filter) has better performance than the previous methods which use the intensity feature only. Most of the conventional methods can not exclude the clutters of target-like intensity in the low contrast image sequences. But the proposed fuzzy-reasoning segmentation can also restrict the clutters from entering the tracking window very efficiently.

4   Conclusions

In the previous study [7], we proposed an adaptive segmentation based on the Bayes decision rule using the additional distance features. The most distinc-
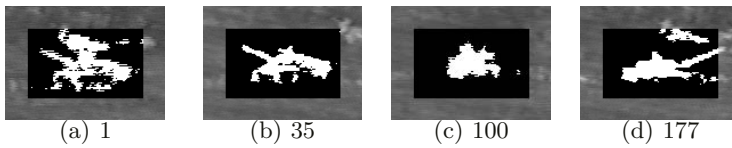
Fig. 5. Tracking results of the tank image using fuzzy-reasoning segmentation
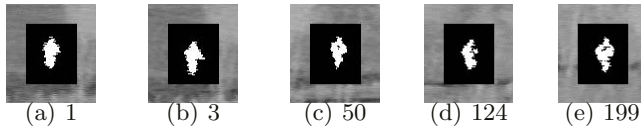
Fig. 6. Tracking results of the man image using fuzzy-reasoning segmentation

tive difference between the traditional segmentation methods and the proposed method was that the proposed segmentation method has additional features of the distance metric as the decision rule, and of low computational complexity for the real-time applications compared with other complex segmentation methods.

In this paper, a fuzzy-reasoning segmentation using the fuzzy membership function is proposed without estimating the state-conditional probability densities. Through other experiments, the proposed fuzzy-reasoning segmentation has good and stable tracking results, and is easy implementable compared to the adaptive estimation approach. So, if we have difficulty in the exact estimation of the state-probability density functions, this fuzzy-reasoning segmentation is a good alternative in cluttered image sequences.

References

1. M. C. Dudzik, "Electro-optical systems design, analysis, and testing," *The Infrared and Electro-Optical Systems Handbook*, SPIE, Vol. 4, pp. 245-298.
2. A. Kumar, Y. Bar-Shalom and E. Oron, "Precision tracking based on segmentation with optimal layering for imaging sensors," *IEEE Trans. on Pattern Anal. Mach. Intell.,* Vol. 17, No. 2, pp. 182-188, Feb. 1995.
3. E. Oron, A. Kumar, and Y. Bar-Shalom, "Precision tracking with segmentation for imaging sensors," *IEEE Trans. on Aerospace Electron. Syst.,* Vol. 29, No. 3, pp. 977-986, July 1993.
4. R. Venkateswarlu, K. Sujata and B. Ven. Rao, "Centroid tracker and aim point selection," in *SPIE Vol. 1697 Acquisition, Tracking, and Pointing VI*, pp. 520-529.
5. L. K. Huang and M. J. Wang, "Image thresholding by minimizing the measures of fuzziness," *Pattern Recognition,* Vol. 28, No. 1, pp. 41 - 51, 1995.
6. A. M. Peacock, S. Matsunaga, D. Renshaw, J. Hannah, and A. Murray, "Reference block updating when tracking with block matching algorithm," Electronics Letters, vol. 36, no. 4, pp. 309-310, Feb. 2000.
7. J.S. Cho, D.J. Kim and D.J. Park, "Robust centroid target tracker based on new distance features in cluttered image sequences," *IEICE Trans. INF. and SYST.,* Vol. E83-D, No. 12, Dec. 2000.

Fuzzy Lowpass Filtering

Yasar Becerikli[*], M. Mucteba Tutuncu, and H. Engin Demiray

Kocaeli University, Computer Engineering Department, Izmit, Turkey
{becer,hdemiray}@kou.edu.tr

Abstract. In this paper, the aim is to solve the effect of a filter on an input signal by using fuzzy logic. We used the characteristic of lowpass filter for this application. If the input frequencies are very high, it will be reduced. We obtained a rule-base for emulating this characteristic. Two methods are used for this problem. In the first method, fuzzy inputs are signal values and the differences of the signal values. This is called "the difference method". The other approach is to calculate the average of two discrete values of the signal. As known, averaging is simple lowpass filter which smoothes out high frequency variations in a signal. But fuzzy approach is used for averaging of the signal. And this is called "the average method".

1 Introduction

The systems and events are defined by crisp mathematical models in engineering and other sciences. The behaviors and state of the system in the future are estimated. However, in the day life, large majority of the problems can't be modeled or can't express a crisp state because of complexity. For research and analysis of these problems we can use fuzzy logic. [1,2] For example, consider the task of driving a car. You notice that the stoplight ahead is red and the car ahead is braking. Your mind might go through the thought process, "I see that I need to stop. The roads are wet because it's raining and there is a car only a short distance in front of me. Therefore I need to apply a significant pressure on the brake pedal." This is all subconscious, but that's the way we think - in fuzzy terms. Do our brains compute the precise distance to the car ahead of us and the exact coefficient of friction between our tires and the road, and then use a lowpass filter to derive the optimal pressure which should be applied to the brakes? Of course not. We use common-sense rules and they seem to work pretty well. So although we think in fuzzy, noncrisp ways, our final actions are crisp. The process of translating the results of fuzzy reasoning to a crisp, nonfuzzy action is called defuzzification.[3] Finally, fuzzy logic was conceived as a better method for sorting and handling data but has proven to be a excellent choice for many control system applications since it mimics human control logic. It can be built into anything from small, hand-held products to large computerized process control systems. It uses an imprecise but very descriptive language to deal with input data more like a human operator.

2 Fuzzy Logic Systems

Fig. 1 depicts a fuzzy logic system that is widely used in fuzzy logic controllers and signal processing applications. Fuzzy logic systems are also known as a fuzzy infer-

---

[*] Corresponding author

ence system, fuzzy rule-based system, fuzzy model, fuzzy associative memory, or a fuzzy controller when it is used as a controller. A fuzzy logic system maps crisp inputs into crisp outputs. It contains four components: rules, fuzzier, inference engine and defuzzier. Once the rules have been established, a fuzzy logic system can be viewed as a mapping from crisp inputs to crisp outputs. [4]
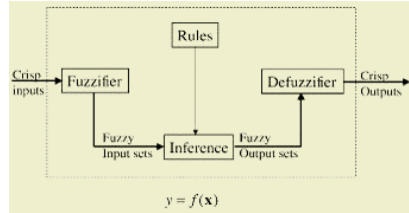
Fig. 1. Fuzzy Logic System

3   Digital Filters

In signal processing, the function of a filter is to remove unwanted parts of the signal, such as random noise, or to extract useful parts of the signal, such as the components lying within a certain frequency range.

There are two main kinds of filter, analog and digital. They are quite different in their physical makeup and in how they work.

A digital filter uses a digital processor to perform numerical calculations on sampled values of the signal. The processor may be a general-purpose computer such as a PC, or a specialized DSP (Digital Signal Processor) chip. DSP technology is nowadays commonplace in such devices as mobile phones, multimedia computers, video recorders, CD players, hard disc drive controllers and modems, and will soon replace analog circuitry in TV sets and telephones. An important application of DSP is in signal compression and decompression. [5]

4   Fuzzy Lowpass Filter Design

The Fuzzy Lowpass filter is designed to see success of the Fuzzy Logic that is used in everything nowadays. Fuzzy lowpass filter can be designed by two methods.

The first method is, choosing the amplitude and difference between signal inputs. This method is called "the difference method". Taking the average of input signals by using fuzzy rules is the second method, and this method is called "the average method"., In the first method we use two signals and the difference between them. In the second method, which is called "average method", we use two signals and average of them. More information about the methods will be given in the next.

4.1   Choosing the Membership Functions for the Difference Method

When we designed a lowpass filter by the way of without using coefficients with fuzzy logic, we should use the characteristic of lowpass filter. This technique is called "pure fuzzy filtering". According to the difference method, the signal amplitudes and the difference between the input signal values are used. Membership functions of the

fuzzy system inputs can be constructed as shown in Fig.2 and in Fig.3. The Fig.2 shows the membership functions for amplitude input. The amplitude is the first input of the fuzzy logic system. It is used for calculating the approximate input.
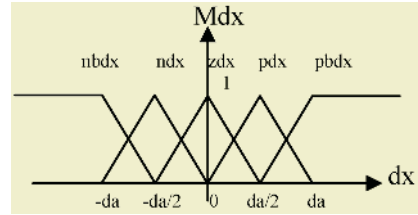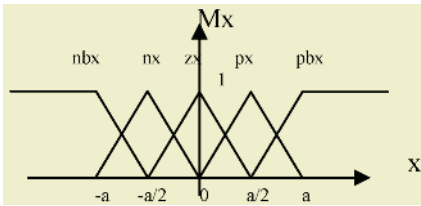
Fig. 2. Membership functions for amplitude input

Fig. 3. Membership functions for difference input

The amplitude input is divided into five parts. The first part is called 'nbx', which means 'negative big x'. It represents the lowest inputs of the signal. The second part is 'nx'. 'nx' means 'negative x'. It represents the low inputs of the signal. We use 'zx' for the signals about zero. The 'zx' means 'zero x'. 'px', which means 'positive x', is used for the high inputs of the signal. The latest part represents the highest inputs. It is shown by 'pbx'.

The equations of the amplitude input memberships are given in equations (1), (2), (3), (4) and (5).

$$\mu_{nbx} = \begin{cases} 1 & x \le -a \\ \dfrac{-a-2x}{a} & a < x \le -a/2 \\ 0 & -a/2 < x \end{cases} \tag{1}$$

$$\mu_{nx} = \begin{cases} \dfrac{2(x+a)}{a} & -a \le x \le -a/2 \\ \dfrac{-2x}{a} & -a/2 < x \le 0 \\ 0 & \text{else} \end{cases} \tag{2}$$

$$\mu_{zx} = \begin{cases} \dfrac{2x+a}{a} & -a/2 \le x \le 0 \\ \dfrac{a-2x}{a} & 0 < x \le a/2 \\ 0 & \text{else} \end{cases} \tag{3}$$

$$\mu_{px} = \begin{cases} \dfrac{2x}{a} & 0 \le x \le a/2 \\ \dfrac{2(a-x)}{a} & a/2 < x \le a \\ 0 & \text{else} \end{cases} \tag{4}$$

$$\mu_{pbx} = \begin{cases} 0 & x \le a/2 \\ \dfrac{2x-a}{a} & a/2 < x \le a \\ 1 & a < x \end{cases} \tag{5}$$

The variable 'a' is maximum value of the signal amplitude For example, for y=3sin(x), value of 'a' is 3 and for y=0.5sin(3x), value of 'a' is 0.5 ,etc. So, the membership functions universes of discourse are changeable according to the maximum value of input signal. This obtains us to see the better results of the system.

The second input which is the difference between discrete values, has five membership functions as shown in Fig. 3. The membership 'nbdx' is used for the biggest difference between signals. It means 'negative big difference'. The 'ndx' which means 'negative difference', represents the small difference between signals which is on the negative section. The 'zdx' is used for the smallest difference or the same inputs. These memberships equations can be written as same as the equations of amplitude input memberships.

In Fig.3, the variable of 'da' is the difference between discrete values of the signal. 'da' is used as a variable to observe the behavior of the system. We can say that maximum value of 'da' is '2a'. Because if the value of a signal is 'a' and the value of the following signal is '-a', the difference would be '2a'. And it is the maximum difference that would be.
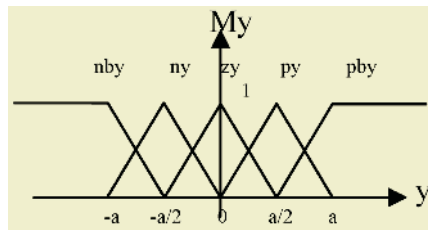
Fig. 4. Membership for output

For output of system five membership functions can be constructed as shown in Fig.4. These are same as the first input that is amplitude input. The output is divided into five parts. The membership 'nby' is used for the negative big outputs. 'ny' means 'negative output'. It represents the low outputs of the signal. We use 'zy' for the outputs about zero. The 'zy' means 'zero output'. 'py', which means 'positive output', is used for the high outputs of the signal. The latest part represents the highest outputs. It is shown by 'pby'.

4.2   Forming Rule Base and Fuzzification

We know we have two inputs, the amplitude and the difference. Both the amplitude input and the difference input have five memberships. We match these inputs memberships and construct the rule table. When the rule table is constructed, lowpass filter characteristic is used. According to this characteristic, a lowpass filter extinguishes the high frequency and passes the low frequency. For example, our amplitude input may be 'nbx' and difference input may be 'nbdx'. That shows us there is a noise and it must extinguish. So the output may be 'ny'. Or, the amplitude may be 'px' and the difference may be 'zdx'. We can say that the output may be 'zy' according to the characteristic of the lowpass filter.

The rule base is designed as in Table 1 for the difference method.

Table 1. Rule table for difference method

dx / x	nb	n	z	p	pb
nb	nb	nb	nb	n	z
n	nb	n	n	z	z
z	n	z	z	z	p
p	z	z	p	p	pb
pb	z	p	pb	pb	pb

The table must be translated to linguistic language for solving the problem by using mathematical formulas. We expressed the rule table in linguistic language as following;

Rule₁- IF x=nbx AND dx=nbdx THEN y=nby
Rule ₂- IF x=nbx AND dx=ndx THEN y=nby
.
Rule ₂₄- IF x=pbx AND dx=pdx THEN y=pby
Rule ₂₅- IF x=pbx AND dx=pbdx THEN y=pby

We use Mamdani Method for this problem. Because Mamdani method is the most useful method in fuzzy logic. It is simple and gives the good results for problems. For inference system with Mamdani Method, the stage of the method can be expressed as following.

According to the rules, the grade values of membership functions F_1, F_2, F_3, F_4....F_{25} can be obtained for the input signal values and difference signal values. The minimum value of membership grade in a rule is taken, because of the Mamdani Method.

F_1=min (Mnbx,Mnbdx), F_2=min (Mnbx,Mndx), F_3= min (Mnbx, Mzdx)
.
F_{24}= min (Mpbx, Mpdx), F_{25}=min (Mpbx,Mpbdx)

Therefore, the membership grades that effect or fire the membership functions can be obtained, separately. In that way we can see which of the rules affects the memberships. So

F_i's for 'nby': F_1,F_2 ,F_3 ,F_6
F_i's for 'ny': F_4 ,F_7 ,F_8 ,F_{11}
F_i 's for 'zy': F_5 , F_9 , F_{10} ,F_{12} ,F_{13} ,F_{14} , F_{16} , F_{17} , F_{21}
F_i's for 'py': F_{15} ,F_{18} ,F_{19} ,F_{22}
F_i's for 'pby': F_{20} ,F_{23}, F_{24}, F_{25}

Then we should choice the maximum values in those results. So we find a value for each output membership functions. These values represents the cut level of output membership functions.

R1 for Mnby = max (F_1 ,F_2 ,F_3 ,F_6)
R2 for Mny = max (F_4 ,F_7 ,F_8 ,F_{11})
R3 for Mzy =max (F_5 ,F_9 ,F_{10} ,F_{12} ,F_{13} ,F_{14} ,F_{16} ,F_{17} ,F_{21})
R4 for Mpy = max (F_{15} ,F_{18} ,F_{19} ,F_{22})
R5 for Mpby = max (F_{20} ,F_{23}, F_{24}, F_{25})

In this result, the R1 is the effecting value of 'Mnby' membership function. In the same manner, the fire grade or membership value of rules on the output membership functions can be found. A defuzzification method must be applied on these outputs because these outputs are fuzzy outputs. By applying defuzzification methods on calculated results, crisp outputs can be obtained. The defuzzification methods will be told later.

4.3 Average Method for Designing Lowpass Filter

We called average method to the second method to design fuzzy lowpass filter. It is about to take the average of input signals by using fuzzy rules. As known, averaging is a simple type of lowpass filter as it tends to smooth out high frequency variations in a signal. In this method, inputs are only amplitudes of input signal. We calculate the average of two input values. For inputs we can use one kind of membership function. The input signals values are between "–a" and "a" as shown in Fig. 2. "a" is the maximum value of input signal and "-a" is the minimum value of input signal. The membership functions of inputs are shown as in Fig. 2. We can constitute the rule base for this method as shown in Table 2.

Table 2. Rule table for avarage method

x_i \ x_{i+1}	nb	n	z	p	pb
nb	nb	n	n	z	z
n	n	n	z	z	z
z	n	z	z	z	p
p	z	z	z	p	p
pb	z	z	p	p	pb

When forming Table 2, this way can be following:

Some numbers that proportional to degree of membership functions, are given to the membership functions. For example we can say "-2" for "nb" membership, "-1" for "n" membership, "0" for "z" membership, "1" for "p" membership and "2" for "pb" membership. When calculating the average of two membership function, the numbers are added and then divide by two. If the result is not a whole number, the decimal part can be dropped. The obtained number represents the output membership functions degree. For example inputs can be "n" and "pb". "n" is equal to "-1" and "pb" is eual to "2". The sum of "-1" and "2" is "1". If "1" is divided by 2, the result is "0.5". And decimal part is dropped. So the result is "0". "0" represents "z" membership function. So we can say if the inputs are "n" and "pb", the output is "z", as shown in Table 2.

The table must be translated to linguistic language for solving the problem by using mathematical formulas. The **Rule$_i$**'s and **F$_i$** 's are calculated with the same method in the first method of designing fuzzy lowpass filter. And it continues such as following;

F_i's for 'nby': F_1

F_i's for 'ny': F_2, F_3, F_6, F_7, F_{11}

F_i's for 'zy': F_4, F_5, F_8, F_9, F_{10}, F_{12}, F_{13}, F_{14}, F_{16}, F_{17}, F_{18}, F_{21}, F_{22};

F_i's for 'py': F_{15}, F_{19}, F_{20}, F_{23}, F_{24}

F_i's for 'pby': F_{25}

As the first merhod F_i's represents the membership grades that effect or fire the membership functions. Therefore;

R1 for Mnby = F_1

R2 for Mny = max (F_2, F_3, F_6, F_7, F_{11})

R3 for Mzy = max (F_4, F_5, F_8, F_9, F_{10}, F_{12}, F_{13}, F_{14}, F_{16}, F_{17}, F_{18}, F_{21}, F_{22})

R4 for Mpy = max(F_{15}, F_{19}, F_{20}, F_{23}, F_{24})

R5 for Mpby = F_{25}

After the defuzzification we will have the result with these coefficients.

4.4 Defuzzification

Defuzzification methods are used for transform fuzzy outputs to crisp outputs. There are many methods for defuzzification. In this fuzzy system, we use 'Center Average Method' for defuzzification. The general definition of Center Average Method is expressed as

$$u^* = \frac{\sum u_i \mu_u(u_i)}{\sum \mu_u(u_i)} = \frac{\sum_{i=1}^{l} u_i \cdot \max(\mu_{KDU^{(k)}}(u_i))}{\sum_{i=1}^{l} \max(\mu_{KDU^{(k)}}(u_i))} \tag{6}$$

In this expression the maximum values show the membership grades of Ri rule. u_i is the top value of the output membership function are obtained from Mamdani Method. Therefore we have a static formula with using these values as shown below:

$$u^* = \frac{u_1 R_1 + u_2 R_2 + u_3 R_3 + u_4 R_4 + u_5 R_5}{R_1 + R_2 + R_3 + R_4 + R_5} \tag{7}$$

The obtained value that is found from the formula is only for one signal input. If our input signals are between [0, 2π], we have to do these operations 360 times in a vicious circle. So we will have 360 various crisp values.

5 Simulation Results

The signal that is y=sin(3x), is taken for simulating the fuzzy lowpass filter methods. And a uniformly random noisy is added to input signal. The amplitude of input signal is between -1 and 1. So we can say the "a" value can be 1, and the "da" value can be 1, too. The membership function equations can be written as using these values of "a" and "da". These equations was shown in equation (1), (2), (3), (4) and (5). When the operations that are told in the difference method are applied to the noisy signal, the result in Fig. 5 is obtained. The operations must be done 360 times. Because these operations are conclute two inputs. But these signal has 360 inputs.

When the average method is applied to the noisy signal, we can take the result shown in Fig. 6. In this method the value "a" is 1, too. There are not "da" parameter, because "da" represents the difference and in this method there are no difference. There are the values of input signals. And the output is the average of these inputs by using fuzzy logic.
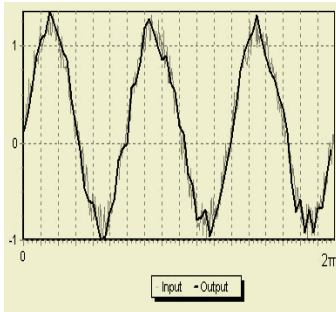
Fig. 5. Result of difference method
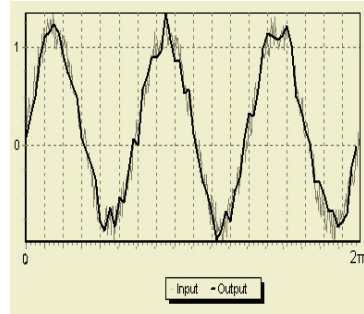
Fig. 6. Result of average method

6   Conclusions

In this paper, two results have been found by two ways but they are close to each other. When this study is compared with a normal lowpass filter that is not done by fuzzy logic, the result have not been gave us the base function exactly because of the membership functions are used. However, more clear results can be found by changing the membership functions. For instance; the cutoff frequency of the lowpass filter varies as a result of changing the parameter 'a' in this application.

References

1. Zimmermann, H.J., "*Fuzzy set theory and its applications*", Kluwer Academic Publishers, Boston, 1991.
2. Mizumoto, M., Fukami, S., Tanaka, K., "Some methods of fuzzy reasoning", *Advances in Fuzzy Set Theory and Applications*, Amsterdam, North-Holland, 1979, pp. 117-136.
3. Kerre, E.E.,Nachtegael, M., "*Fuzzy Techniques in Image Processing*", Physica-Verlag, 2000
4. Wang, L.X.,"*A Course In Fuzzy Systems And Control*", Prentice-Hall International, 1997.
5. Smith, S.W., "*The Scientist and Engineer's Guide to Digital Signal Processing*", California Technical Publishing San Diego, California,1999
6. Zadeh, L.A., Fu, K.S., Tanaka, K., Shimura, M., "*Fuzzy sets and their applications to cognitive and decision processes*", Academic Press, New York., 1975.
7. Mendel, J.M., "Uncertainty, Fuzzy Logic and Signal Processing", *Signal Processing*, 1999, pp. 913-933.
8. Simon, D., "Kalman Filtering for Fuzzy Discrete Time Dynamic Systems," *Applied Soft Computing*, vol. 3, no. 3, May 2003, pp. 191-207.
9. Zhang, H.C., Huang, S.H., "A fuzzy approach to process plan selection", *I. J.of Prod. Res.*, vol. 32, no. 6, 1994, pp. 1265-1279.
10. Kayran, A.H., "*Digital Signal Processing*", İTÜ Press, Istanbul, 1990
11. Yeşil, E., "*Introduction to Fuzzy Logic*", Bumat Press, Istanbul, pp. 7-8, Şubat 2001
12. Chiu,S.,Chand,S., "Adaptive Traffic Signal Control Using Fuzzy Logic", Second IEEE Int. Conference on Fuzzy Systems, Vol.1, 1371-1376,California,1993

Fuzzy Logic Based Intelligent Tool for Databases

Sevinc Ilhan and Nevcihan Duru

Department of Computer Engineering, Kocaeli University, Izmit Turkey
{silhan,nduru}@kou.edu.tr

Abstract. In this paper, a software tool for enabling fuzzy query from a classical database is introduced. By using this tool, some fields (attribute) of a database table can be fuzzified and a supplementary database, which includes fuzzy values, is formed. Developed software tool is applied in a sample database including some fields about the students for the evaluation of scholarship application. It is concluded that, the fuzzy query method is more flexible and the results of such query are more predictive.

1 Introduction

In this study, a database, including the information about the students, has been investigated by using the developed tool, to determine to whom to give the scholarship. Some of the students in Turkey have difficulties in attending the university because of the financial situation of their families. The need of the scholarship for the children of the families especially in the Eastern part of Turkey, having a lot of members, from brothers and sisters point of view, is greater than the other parts of the country. With the help of this system, students will be financially supported.

In our study, an intelligent software has been developed for fuzzy query, which makes it possible to query on a classical database. The developed software fuzzifies the fields, which are required to be fuzzified, by connecting to the classical database and creates a supplementary database. This database includes fuzzy data and values, which are related with the created membership functions. In the recent years, fuzzy query studies have been increasing in the literature [1], [2], [3], [4], [5], [6]. Most of the fuzzy query tools, which have developed beforehand (in the beginning), are not independent from the SQL' s structure form. This study is quite different from the others with this feature. In the developed intelligent query tool, the query sentences are written with linguistic variables, as in the daily speaking language. So, for the user, it is easy to write query sentences. Like Tahani's [5] query tool, it allows the fuzzy and the crisp query to be found in the same query sentence. It has similarities with the other fuzzy query tools, but also it has differences. In Section-2, insufficiency of the classical query is examined by an example. In Section-3, the developed software is explained on the basis of the example given in Section-2.

2 Problems Associate to Classical Query

Classically, the query is a process of retrieving some data according to some criteria, which are set by the user. In practice, the most used query language is SQL since it became a standard for relational databases [7]. SQL has certain rules and commands for the query. The results, which confirm the selection criteria, are retrieved from the

database. Classically, the criteria include logical and mathematical statements. In some of the queries, required results may not be obtained since some of the data do not confirm the criteria by practically insignificant differences.

In Table 1, a database, called as "Scholarship", which includes some information about the students, is shown. Suppose that a query for the evaluation of the necessity of the scholarship is made. In order to determine a student's necessity to the scholarship, his or her family's salary (FSALARY in database) and the number of the brother or sister in their family (BROTHER_SISTER in database) in the experiment are considered. Assume that the lecturer wants to evaluate the students according to these criteria. The SQL query for the particular assessment problem may be in the form of the following.

Table 1. Scholarship database

Data in Table 'SCHOLARSHIP'

NUMBER	NAME	FSALARY	BROTHER_SISTER	POCKETMONEY	HOUSERENT
201010	Mustafa	250	10	110	70
201011	Serpil	500	2	200	120
201012	Oktay	320	7	60	40
201013	Muhsin	335	2	100	50
201014	Gozde	100	4	80	40
201015	Silen	350	4	50	40
201016	Zeynel	150	9	100	75
201017	Riza	200	3	200	90
201018	Muharrem	590	5	150	100
201019	Ahmet	375	5	0	75
201020	Sevda	330	5	200	120
201021	Emine	300	8	80	40
201022	Aydin	420	14	150	100
201023	Murat	350	2	0	60

SELECT NUMBER, NAME, FSALARY, BROTHER_SISTER
FROM SCHOLARSIP   WHERE FSALARY<=300 AND BROTHER_SISTER>=5

The selected students for the scholarship in this case will be as follows:

NUMBER NAME   FSALARY   BROTHER_SISTER

201010   Mustafa   250   10
201016   Zeynel   150   9
201021   Emine   300   8

If the table is inspected carefully, then it can be noticed that, some of the records, which are very close to the selection criteria, were not included in the results table. The record Nr-201012 is not in the list since his BROTHER_SISTER number is much than 5 although his FSALARY value is 320. The record Nr-201014 is not in the list, and so on. Some sort of solution can be found for this small database. Expanding the range and changing the logical operators is a way of enriching the amount of data retrieved from database. Suppose that, query criteria is changed as, FSALARY<=350 AND BROTHER_SISTER>=4. The result would be as follows:

NUMBER NAME   FSALARY   BROTHER_SISTER
----------------- ---------------------------- -----------------------------------
201010   Mustafa   250   10
201012   Oktay   320   7
201014   Gozde   100   4
201015   Silen   350   4
201016   Zeynel   150   9
201020   Sevda   330   5
201021   Emine   300   8

In this case, although the number of students has been increased due to the change of the boundaries, Record Nr-201022 would be out of the list although his BROTHER_SISTER number is 14; and also Record Nr-201019 would be out of the list. In a database with huge data and large number of fields, it is impossible to inspect the effect of the criteria variations over the results. The solution of the problem is making the query flexible and intelligent. Fuzzy logic fits very well to the solution of this problem with its fuzzy and humanlike nature [8].

3   Fuzzy Query Software Tool

Introduced software tool has been developed in Borland DELPHI 6.0 and MS SQL Server 7.0 is used as Database Management System (DMS). The tool is made up of two modules: the module which fuzzifies the attributes and the module in which the fuzzy query is implemented. The block diagram of the introduced system is shown in Fig.1.
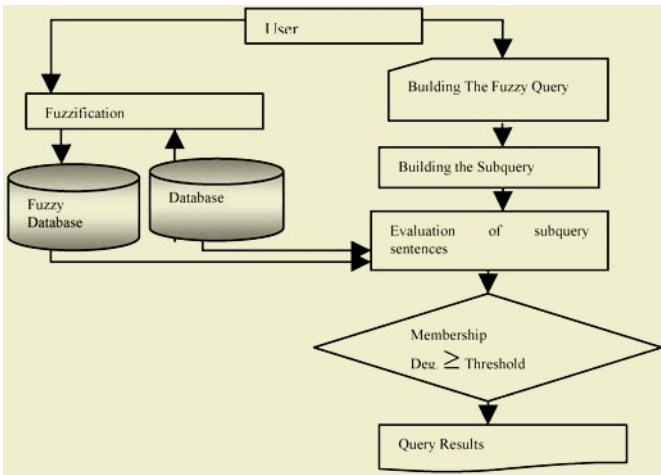
Fig. 1. The block diagram of the introduced system

3.1   The Fuzzyfying Module

As it is known, a fuzzy set is a set, which contains elements having varying degrees of membership. Elements of a fuzzy set are mapped into a universe of "membership values" using a function-theoretic form. Fuzzy sets are denoted by a set symbol. As an example, let "A" denote a fuzzy set. Function μ maps the elements of fuzzy set "A" to real numbered value on the interval 0 to 1. If an element in the universe, say x, is a member of the fuzzy set "A", then this mapping is given as:

$$\mu_A(x) \in [0,1] \tag{1}$$

$$A = (x, \mu_A(x)/x \in X) \tag{2}$$

$$A = \sum \mu_A(x_i)\Big/x_i = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \ldots\ldots\ldots + \mu_A(x_n)/x_n \tag{3}$$

In the definition of the membership functions, above function is used. The membership function is defined by member and its membership degree [9].

In order to fuzzify a database field, the module should connect to the database, and should define the membership degrees of these fields. When the main program activates module, module connects to MS SQL server and reaches to the fields of an existing database. The tables are listed on the screen and the user is asked for making a choice of field(s), which is used in the query. By using the interface screen, which is shown in Fig. 2, minimum and maximum values of fuzzy set are defined.
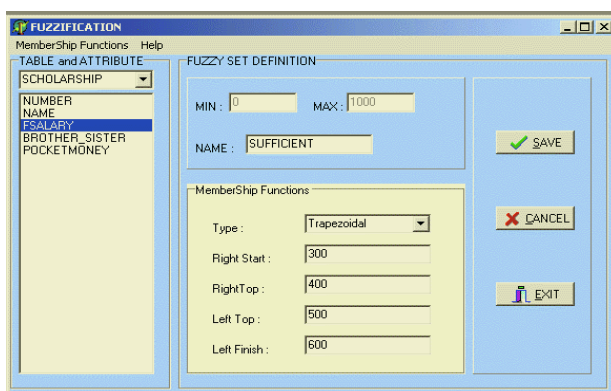
Fig. 2. Fuzzification Interface

Another interface is used for the definition of the type and the name of a membership function and its parameters**.** Types may be triangular, trapezoidal and exponential or any other known form. In the developed tool, triangular and the trapezoidal membership functions are selected because of their ease of usage. The names of membership functions are preferably chosen as linguistic variables. The name of classical database, its fuzzified fields, limits of fuzzy sets, name of membership functions and their types are all located in a supplementary table (called as Fuzzy_Set _Table), which is defined in the database called Fuzzy_Knowledge_Database.

3.2   The Fuzzy Querying Module

This module makes query, which is set by fuzzy query sentence. In general, this sentence has the form of:

ATTRIBUTE FUZZY MODIFICATOR **MEMBERSHIP FUNCTION** NOT **AND/(OR)**   **ATTRIBUTE MEMBERSHIP FUNCTION**.

In Fig.3, the main interface screen of the module is shown where the query sentence "FSALARY INSUFFICIENT OR BROTHER_SISTER TOOMANY" and its results are displayed. In order to implement this query, firstly, the "Salary" field is fetched from the database and its membership degrees in the "Insufficient" membership function are computed which are then inserted in supplementary database. Same process is applied for the second or other parts of fuzzy query sentence (if any) which are linked by AND/OR statements [10]. The membership degrees and the connectors

(and/or) are stored in membership dynamic array and connector dynamic array. The number of subqueries doesn't affect the result. Because, they are stored in dynamic array and the size of it determined by the user. In all of the subqueries, the fetching degree to the database record is calculated and the results stored in dynamic array.

If the first and the second subqueries are connected with "AND"

$$Result = \min \left(\mu_{SubQuery1}(x), \mu_{SubQuery2}(x) \right) \tag{4}$$

If the first and the second subqueries are connected with "OR"

$$Result = \max \left(\mu_{SubQuery1}(x), \mu_{SubQuery2}(x) \right) \tag{5}$$

When calculating the total membership degree of the fuzzy query the membership degrees of the subqueries are used. In this calculation "AND" operation has priority. From the beginning of the fuzzy query, first of all the "AND" operation has been done, then the "OR" operation with the new calculated results.
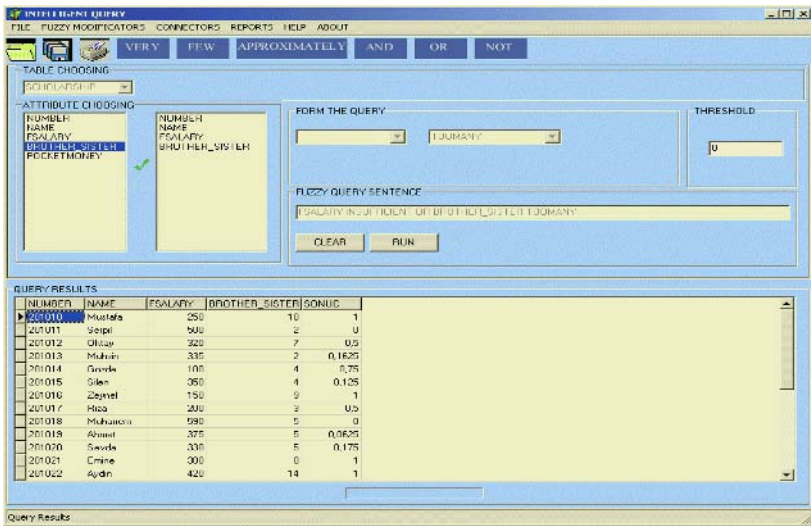
Fig. 3. Fuzzy Query Interface

If the query sentence contains both the crisp and the fuzzy query, the same processes are done. The crisp subquery's matching degree is 1, if the query is true and 0, if the query is not true.

In our query model, querying is based on a representation language that allows the natural language sentences and the use of modifiers or quantifiers like "Very", "Few" "Approximately". Modifier functions, from [0,1] to [0,1], can be applied to fuzzy set membership functions in order to model the effect of linguistic hedges such as "Very", "Not", "Much", "More or Less" etc [11], [12]. $\alpha - Cut$ set of the fuzzy set "A" is shown as A_α and defined as $A_\alpha = \{x \in X | \mu_A(x) \geq \alpha\}$ [12]. In the above-developed software, the result set changes according to the threshold value ($\alpha - Cut$) which is entered to the system by the user.

If the fuzzy query is like "A1 m1 MF1 AND/OR A2 m2 MF2", "A1" and "A2" are the related attribute values, "m1" and "m2" are the linguistic modifiers, and MFs are the fuzzy sets, then the rules of the fuzzy query are shown in 6th and 7th equations.

$$R = \min(\mu_{m1}(\mu_{F1}(x), \mu_{m2}(\mu_{F2}(y)))), \text{ if the connector is AND} \tag{6}$$

$$R = \max(\mu_{m1}(\mu_{F1}(x), \mu_{m2}(\mu_{F2}(y)))), \text{ if the connector is OR} \tag{7}$$

The meaning associated with the linguistic variable, "FSALARY VERY INSUFFICIENT", Fuzzy subset whose membership function is the square of the membership function of the meaning of "insufficient" characteristic hedge operation. The meaning associated with the linguistic variable, "FSALARY NOT VERY HIGH", Fuzzy subset whose membership function is [1-the membership function of the meaning associated with "FSalary is very High"] – fuzzy not operation on relevant subsets [14].

4   Conclusions

In this study, software, which can make fuzzy query on a classical database, is introduced. In the literature most of the developed tools are related with SQL's structured form. This study is quite different from the others with this feature. In the developed intelligent query tool, the query sentences are written with linguistic variables, as in the daily speaking language. Introduced tool is explained on the basis of an example. It is shown that, the help of this software can significantly enrich the results of a query. The system was applied on about 1500 students' application form for choosing the appropriate 120 students. The results were satisfactory. The right students were chosen with the help of the system.

References

1. Rasmussen, D., Yager, R.R.: Summary SQL-A Fuzzy Tool For Data Mining. Intelligent Data Analysis, Vol.1, Elsevier Science Inc. (1997)
2. Bosc, P., Pivert, O.: SQLf: A Relational Database Language for Fuzzy Querying, IEEE Transactions on Fuzzy Systems, Vol 3. (1997)
3. Cox, E.: FuzzySQL-A Tool for Finding The Truth-The Power of Approximate Database Queries, PC AI-Intelligent Applications, Volume 14, (Jan/Feb 2000)
4. Eminov, M.: Querying a Database by Fuzzification of Attribute Values, http://idari.cu.edu.tr/sempozyum/bil46.htm
5. Zadeh, L.A., Kacprzyk, J.: Fuzzy Logic For The Management of Uncertainty, Library of Congress Cataloging in Publication Data Press, New York, (1992) 645-672.
6. Kacprzyk J., Ziolkowsski A.: Database Queries with Fuzzy Linguistic Quantifiers, IEEE Transactions on Systems, Man. And Cybernetics, Volume SMC-16, 474-479.
7. Elmasri, R., Navathe, S., B.: Fundamentals of database systems. Addison Wesley, (2000) 243
8. Zadeh, L.A.: Fuzzy Sets. Information and Control, 8, pp., (1965) 338-353.
9. Jamshidi, M.: Fuzzy Logic and Control: Software and Hardware Applications. PTR Prentice-Hall, Englewood Cliffs, New Jersey (1993) 16-18
10. Takahashi, Y.: Fuzzy Database Query Languages and Their Relational Completeness Theorem. IEEE Transactions on Knowledge and Data Engineering, Vol. 5. (1993) 123

11. Wang Li-Xin: Adaptive Fuzzy Systems and Control: Design and Stability Analysis. PTR Prentice-Hall, Englewood Cliffs, New Jersey (1994) 9-14
12. Bosc, P., Prade H.: An Introduction To the Fuzzy Set Possibility Theory-Based Treatment of Soft Queries and Uncertain or Imprecise Databases. Uncertainty Management in Information Systems: From Needs to Solutions, Kluwer Academic Publ., (1997) 285-324
13. Elmas, Ç.: Fuzzy Logic Control Mechanism. Seçkin, Sıhhiye, Ankara (2003) 63
    Bulanık Mantık Denetleyiciler. Seçkin, Sıhhiye, Ankara (2003) 63
14. Nath, A. K., Lee, T. T.: On The Design of a Classifier With Linguistic Variables As Inputs, Fuzzy Models for Pattern Recognition: Methods That Search for Stuctures in Data. By James C. Bezdek and Sankar K. Pal., New York (1992). 242-244.

Modelling from Knowledge Versus Modelling from Rules Using UML

Anne Håkansson

Department of Information Science, Computer Science
Uppsala University
Box 513, SE-751 20 Uppsala, Sweden
Tel: +46 18 471 10 73, Fax: +46 18 55 44 21
Anne.Hakansson@dis.uu.se

Abstract. Modelling support for knowledge acquisition is a tool for modelling domain knowledge. However, during the implementation of the knowledge new knowledge is created. Event though this knowledge is found in the knowledge base, the model usually is not updated with the new knowledge and do, therefore, not contain all the knowledge in the system. This paper describes how different graphical models support the complex knowledge acquisition process of handling domain knowledge and how these models can be extended by modelling knowledge from rules in a knowledge base including probability. Thus, the models are designed from domain knowledge to create production rules but the models are also extended with new generated knowledge, i.e., generated rules. The paper also describes how different models can support the domain expert to grasp this new generated knowledge and to understand the uncertainty calculated from rules during consultation. To this objective, graphic representation and visualisation is used as modelling support through the use of diagrams of Unified Modelling Language (UML), which is used for modelling production rules. Presenting rules in a static model can make the contents more comprehensible and in a dynamic model can make the uncertainty more evident.

1   Introduction

Knowledge acquisition is the process of capturing the domain knowledge and transferring it from a source to a knowledge system [4]. The knowledge acquisition comprises eliciting, modelling and encoding domain knowledge. Eliciting means to acquire the knowledge from a domain expert. Modelling stands for structuring the knowledge into some knowledge representation form. Encoding involves to transfer the modelled knowledge and implement it into the system. To ensure the usefulness of the system, the system is tested for usability by a verification and validation test of the knowledge base [6], [5], [12].

This process is complex, due to a number of problems, e.g., difficulties with formalising all kinds of knowledge, extracting the knowledge to be inserted into a system, and developing a knowledge base with a specific knowledge representation for the knowledge [16], [2]. Tools for analysing, designing and implementing

knowledge can minimize some of these problems. These tools are used for capturing and modelling the domain where the model constitutes a layer between the domain knowledge and the knowledge representation in the knowledge base [6]. Usually, these models only view the knowledge from one angle while the expert and the engineer may need different. Hence, modelling often requires several different views of the domain knowledge. Moreover, usually the models only consists of knowledge about the domain and not the knowledge in the entire knowledge base, that is if the knowledge base has been extended with rules generated during the implementation.

The knowledge is structured through graphic models, making it possible to implement it into knowledge representation production rules [8]. When domain knowledge is expressed as rules, a side effect of using them together is that additional knowledge can be generated. This generated knowledge may not be obvious to the domain expert but by presenting it to the expert through modelling in the graphic models, i.e., the same models as for structuring knowledge, make the content more evident. By this, the models are growing and become larger than the original ones, which were built from the domain knowledge. The models also become denser because the rules will be more closely connected since the meta-rules, which are created from original rules, are placed on top of the existing rules. Also the facts will be more closely connected due to these meta-rules. A problem with denser knowledge bases is to present them to the domain expert. Another problem is that the expert can be unaware of how the certainty of the conclusions was drawn from certain knowledge.

This paper presents different graphic models to support the complex process of knowledge acquisition by modelling domain knowledge visually. These models are also used to handle the extended and usually complex knowledge bases that contains closely connected rules. The graphic models are modelled with modified Unified Modelling Language (UML) diagrams to cope with the knowledge engineer, the domain expert and the end user different perspectives [9]. These diagrams are used to insert knowledge in the knowledge base as well as extend with generated knowledge and inform the users about the systems processing [8]. An object diagram of UML is used to present rules from the domain knowledge and new generated rules in the knowledge base. The sequence diagram of UML presents a view of the rules with facts in a static fashion by illustrating the meta-rules with their relationship and the collaboration diagram of UML presents a view of the rules with uncertainty in a dynamic fashion. The modelling has been tested in domains for hydropower development and river regulation, water analysis with plankton and alga, water analysis with water plants in nutritious water, and classification of edible fungus.

2   Related Work

UML has become a commonly used modelling language in several different methodologies used for developing systems, usually object-oriented systems. For example, CommonKADS, which is a methodology for developing knowledge sys-

tems, incorporates parts of UML [16]. Another approach is using UML Profiles for knowledge modelling, which is an extension of UML from the design of the knowledge modelling on an eXecutable Meta-modelling Language (XMF) framework [1]. UML can also be used for modelling of other types of systems, such as frame and constraint based systems [7]. In our approach, we work with UML for knowledge-based systems that use the representation production rules [8], [9].

3   Modelling Knowledge

The quality of the knowledge base and usability of the system can be preserved by letting the domain expert be involved in the process of knowledge acquisition but also in other phases until the testing phase. The quality and usability is also dependent of the end users consideration why these users would be involved in the modelling process and until the system has been delivered. Thus, the modelling should be user-centred and involve the different users in the development, i.e., both expert and end users [9].

When the domain knowledge is structured through models to fit the syntax of rules, it is to say that the domain knowledge is filtered through the models to production rules in the knowledge base. These rules are present from the beginning of the development and, therefore, called initial rules. From these initial rules new rules are generated through inferring and derivation, i.e., generated rules. All these rules have to be filtered into new models to correspond to the complete knowledge base. Filtering back and to rules can be presented as an hourglass, se Figure 1.
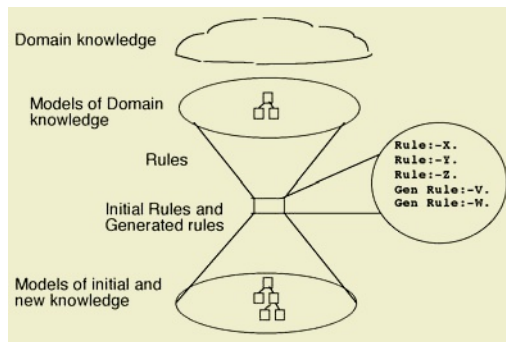
Fig. 1. Filtering knowledge to and from knowledge base.

At the top of the figure, a model of the domain knowledge is created, thus, the domain knowledge is filtered into a graphic model. From this model, rules are created and inserted into the knowledge base. New knowledge is generated through the creation of meta-rules, inferring and deriving rules. This new knowledge is represented as facts and rules. A combination of initial rules and generated rules

constitute larger and denser knowledge base than the model was representing from the beginning. Therefore, the model has to be extended containing old and new facts and rules. This improved model will be larger than the original one since it contains both the domain knowledge and the generated knowledge.

The generated knowledge, i.e., meta-rules, inferred and derived rules, is utilised in the interpretation to draw conclusions. However, the knowledge may not necessary be of good quality. Moreover, this knowledge may not be realised or even unknown to the domain expert. Once the expert is aware of the generated knowledge and finds it to be of good quality, the generated knowledge is confirmed and can be used to find even more knowledge.

Meta-rules are rules that are combining ground rules with relationships or other meta-rules. Thereby, the meta-rules tightly connect the other rules in the knowledge base, which becomes more coherent and more closely connected. Moreover, since the rules use facts, these facts will also be more closely connected with the use of meta-rules.

Uncertainty is a part of the domain knowledge and refers to the degree of plausibility of the statements. This is often used when conclusions are drawn since the outcome of a reasoning strategy is usually not true or false it is something in between, e.g., unlikely, probably, rather likely or possible. The domain expert may be unaware of using uncertainty in their problem solving and will, therefore, have problems with applying the probability on their own knowledge. However, uncertainty can be applied to rules through the use of factors of probability to give the individual rule a specific relevance in a context. These factors are used in a conclusion to calculate the final certainty factor. Unfortunately, some domain experts have difficulties with quantifying subjective and qualitative knowledge. Hence, the factors of probability should be clearly presented to make the probability more graspable and this can be assisted through graphic modelling.

4   Rules from Domain Knowledge

Modelling domain knowledge is the process of structuring knowledge before implementing it in the knowledge base. This includes structuring domain knowledge in the form of graphic models and then in the shape of knowledge representation.

During the development, models are extended with the production rules. These rules constitute a network where the connections (relationships) between rules are important. Since the knowledge base tends to grow quickly, the models have to cope with the complexity of the rules and their relationships. To this objective, we use object diagram in UML [3], [13] as a graphic representation tool for presenting the rules, se Figure 2.

The figure illustrates three different templates for rules connected to each other via relationships. Each rule template is marked with an `r:` followed by a `Rule Name`, which is to give a belonging to a class. The purpose of utilising classes is that classes can be represented in a semantic classification diagram, which can support grasping the detailed and more general rules. This diagram

r: Rule Name	r: Rule Name	r: Rule Name
Rule Number	Rule Number	Rule Number
Rule Object Name	Rule Object Name	Rule Object Name
Conclusion Object	Conclusion Object	Conclusion Object
Probability Factors	Probability Factors	Probability Factors
Facts (Question Object Name)	Facts (Question Object Name)	Facts (Question Object Name)
Rule Object Name (and/or/not)	Rule Object Name (and/or/not)	Rule Object Name (and/or/not)

Fig. 2. Graphic presentation of rules in a knowledge base.

can be represented as a visual library including concepts and information in a hierarchical manner [11].

The `Rule Number` is a unique identification of a specific rule and the `Rule Object Name` is the name of rule or rules. This object name can be used by several similar rules to create a rule set. Thus, all the rules that handle the same subject area are gathered to constitute the rule set. The `Conclusion Object` is the name of the conclusion that can be reached when this rule is satisfied (or fired). `Probability Factors` is the number of certainty ranging from 1 to -1. `Facts` (or Question Object Name) are terms pre-defined in a database or answers to questions that have been answered by the users during a consultation (or session) with the system. `Rule Object Name` is the name of other rules included in this particular rule, i.e., connection to other rules in the knowledge base. These rules are enumerated with **and**, **or** or **not** between the included rules denoting the conditions for the other rules contribution in the consultation. That is to make clear wheather both rules have to be fired (**and**) or either one (**or**) or not at all (**not**). All slots in the templates are filled with information about a rule corresponding to knowledge about the domain.

Through these graphic models, the knowledge is filtered into rules in the knowledge base. Syntactically, each rule is implemented in the following structure:

```
rule(Rule Name, Rule Number, Rule Object Name, Conclusion Object,
Probability Factors):- Facts, Rules.
```

5   Generated Knowledge

During the knowledge acquisition, the domain expert usually extracts the knowledge pieces, i.e., facts needed to draw conclusions, which are used in the rule. This results in a knowledge base with rules only containing facts. However, when the knowledge engineer develops the knowledge base, abstracted rules usually are discovered which constitute the new generated knowledge. The abstracted rules are used to manage other rules accomplished by combining rules through relationships. Thereby, these rules force coherency on the other rules in the knowledge base and, hence, they are closely connected. Moreover since the rules use facts, these facts will also be more closely connected by these abstracted rules.

5.1   Meta-rules

Meta-knowledge is represented as meta-rules. This knowledge is used to decide which rules are fired instead of following a conflict reasoning strategy [12]. Thus, a meta-rule determines a strategy for using task-specific rules in the system [15] by directing the problem solving and also determines how to best solve a problem [6]. This is knowledge about the overall behaviour of the system [15] and not directly related to the domain problem [14].

Rules and meta-rules are presented in a static diagram similar to the sequence diagram of UML, see Figure 3. The diagram is static in the sense that it does not change during consultation. Hence, the diagram is a representation, or snapshot, of the actual contents of the knowledge base without considering external values used by the system to draw conclusions.
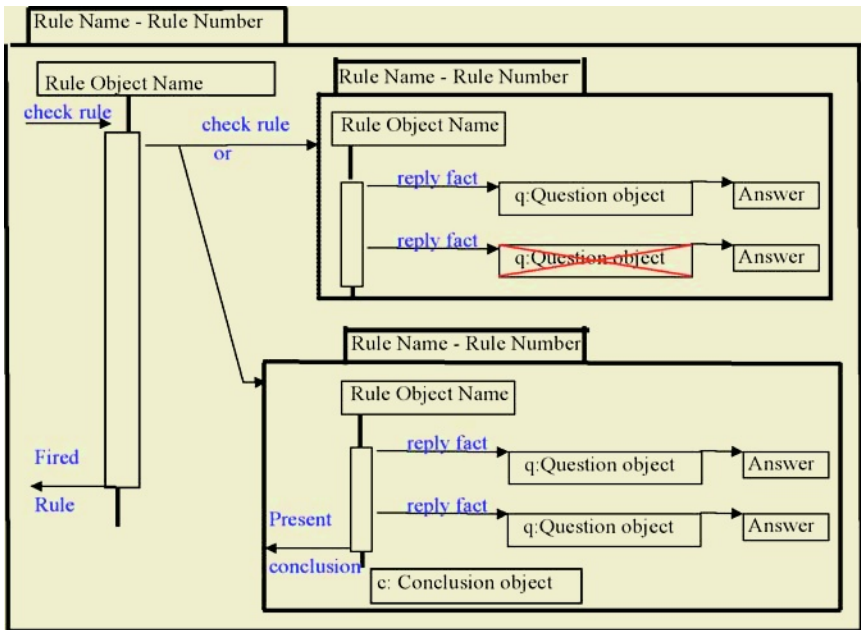
Fig. 3. Presenting meta-rules in a static manner.

This diagram presents a view of the rules in a chronological order of firing the rules. At the top of Figure, the rule is presented with the `Rule Name` and the `Rule Number`. Within that rule, the `Rule Object Name` is specified together with a lifeline. Everything specified along the lifeline corresponds to connections to other rules in the knowledge base or to predefined or user-given facts. In this case, the lifeline invokes two other rules with the arrow `Check Rule` together with an `or`-branch and terminates with the arrow `fired rule`. Within each of the invoked rules, the `reply fact` indicates a predefined fact is fetched or a question

is asked. `q:Question Object` specifies a question where `q` stands for question and the `Question object` is an association to a specific question. The answer is the to the specific question. It the fact is pre-defined, the `Question object` is related to the fact and the answer is the value of it. Putting a cross against denotes that the answer could not under any circumstance be true. In the other rule, the `c` in `c:Conclusion Object` denotes that a conclusion is presented when the rule has been fired. `Conclusion Object` denotes the conclusion to be presented to the user.

From an experience of an assignment of modelling rules as performed in the spring 2005, involving 70 students at Uppsala University, Sweden, we found that new knowledge actually is generated. It turned out that modelling knowledge in domains such as water analysis with plankton and alga, water analysis with water plants in nutritious water, and classification of edible fungus did generate abstracted rules.

5.2   Inferred and Derived Rules

Other types of generated knowledge are inferred rules and derived rules. Inferred rules are derived from the initial rules. It is the process of linking new information from old [14]. For example, if A implies B and B implies C, an inferred rule can link these together by conclude that A implies C. In the inferred rules, the antecedent of the initial rules appears in the body of the inferred rule. Inferring rules can be applied on rules in a knowledge base. Derived rules are derived from other rules. These rules are similar to inferred rules since it is the process of linking new information. However, a derived rule can be regarded as a short cut for several lines of primitive rules.

6   Modelling from Initial and Generated Rules

Modelling the entire knowledge base implies not only modelling from the domain knowledge but also modelling with the generated knowledge. Therefore, the models need to be updated with this new knowledge from the knowledge base. Thus, the models contain both the original domain knowledge and the generated knowledge. Consequently, these models will be larger than the original ones.

The domain expert is, quite likely, not aware of this generated knowledge or has not realised that this knowledge is used in the problem solving. When this generated knowledge becomes obvious, the domain expert needs to check its quality for consistency and completeness. Therefore, the new generated rules have to be modelled into the diagrams to be checked by the domain expert. Thus, the entire knowledge base has to be remodelled where the rules from the domain knowledge and the generated rules are modelled into the diagram, see Figure 4.

The Figure 4 illustrates a remodelling of the model presented in Figure 2, in an object diagram. The three rules at the top are the same rules as in Figure 2.
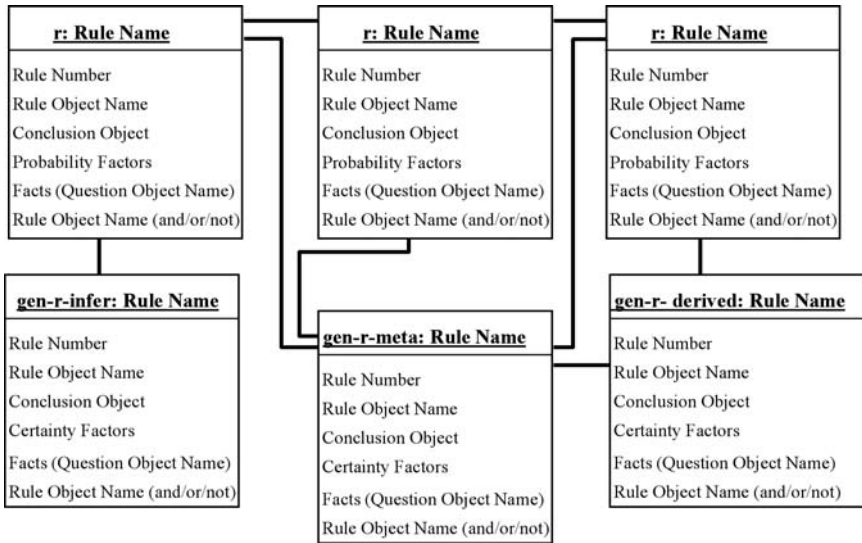
Fig. 4. Graphic presentation of initial and generated rules.

From one of the three rules, three new rules have been generated. These rules are tagged with `gen-r:` and include, meta-rules, inferred rules and derived rules. Except for the tag, the content of the rule is the same as mentioned earlier. One can argue that the slot `Facts` is unnecessary since the generated rules only should contain other rules. However, the system must provide the facility of inserting facts into generated rules needed when concretise a situation to be handled by the system.

7   Uncertainty Management

Most real problems where knowledge systems are used, usually do not provide us with clear-cut knowledge. Instead, the information is incomplete, inconsistent, unmeasurable but also uncertain. The definition of uncertainty is the lack of exact knowledge that enables us to reach a reliable conclusion. The lack of the knowledge can be unsuitable for solving problems but the domain expert can still cope with this situation and make right decisions. To some extent, the systems also have the ability to handle the uncertainty and can, therefore, draw valid conclusions [15] by using methods such as certainty factors or bayesian network.

   Probability theory concerns the existence of a specifiable likelihood that some event occurs, which can be expressed in terms of numbers. This probability is represented as values of likelihood in the rules, mentioned earlier. These values need to be presented together with the conclusion to be comprehensible by the expert who measures the different conclusions probability. With graphic representation of likelihood values together with rules and conclusions, the probability of conclusions can become more evident.

Probability theory, rules and conclusions can be presented as a dynamic presentation in a diagram similar to the collaboration diagram of UML. The diagram is dynamic in the sense that it changes constantly during the execution of the system. The contents of the diagram depend on the inputs inserted by the users during a session and the diagram presents the parts relevant to that specific consultation, see Figure 5.
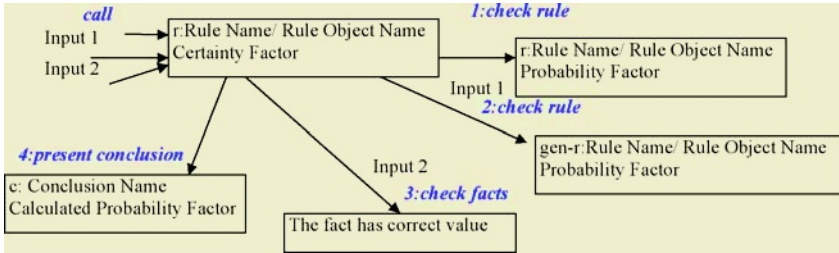
Fig. 5. Graphic presentation of uncertainty in a dynamic diagram.

To the upper left in the Figure, a rule is called with the input, `Input 1` and `Input 2`. The input is two knowledge pieces, facts or rules, calling a particular rule. This rule calls other facts or rules to check if the fact is true or whether the rules can be fired or not. In each rule, respectively, the `probability factor` is specified. Also in the conclusion, the probability factor is presented. This conclusion is presented to the end users with a value calculated from all the factors involved in a consultation, i.e., all values of factors from fired rules are gathered and computed.

8   Conclusions

In this paper we have presented modelling of rules from domain knowledge, i.e., filter rules from domain knowledge trough modelling. We have also presented modelling of generated rules, i.e., meta-rules, initial rules and generated rules. These models make it possible to actually see the connection between rules. It is also possible to comprehend the factors of probability calculated for a conclusion since the uncertainty is illustrated in the models. The modelling of rules has been tested on several different domains. However, modelling of generated rules remains to be investigated. Moreover, the applicability of modelling of generated rules in a knowledge base for reengineering of knowledge will be investigated.

References

1. Abdullah M. S., Benest, I., Evans, A. & Kimble, C.: Modelling Knowledge-Based Systems Using UML Profiles. *(ISDA 2004) IEEE 4th International Conference on Intelligent Systems Design and Application.* Budapest, Hungary, August 26 - 28. ISBN9637154299 (2004) 49-54

2. Alonso, F., Fuertes, J.L., Martinez, L. & Montes, C.: An incremental solution for developing knowledge-based software: its application to an expert system for isokinetics interpretation. Expert Systems with Applications **18** (2000) 165-184

3. Booch, G., Rumbaugh, J., & Jacobson, I.: *The Unified Modeling Language User Guide* . Addison Wesley Longman, Inc.(1999)

4. Chapman, C.: Introducing Knowledge Based Engineering & Applications from Industry (military-aerospace-automotive). *(ES2002) The Twenty-first SGES International Conference on Knowledge Based Systems and Applied Artificial Intelligence.* Peterhouse College, Cambridge, UK December 10th-12th (2002) 49-54

5. Darlington, K.: *The Essence of Expert Systems* . Prentice Hall, England.(2000)

6. Durkin, J.: *Expert System Design and Development* . Prentice Hall International Editions. Macmillian Publishing Company, New Jersey.(1994)

7. Helenius, E.: *UML and Knowledge Engineering of Frame Based Knowledge Systems. (UML och systemutveckling av framebaserade kunskapssystem)* . Master Thesis, Department of information science, Computer Science Division, Uppsala University. (2001)

8. Håkansson A.: UML as an approach to Modelling Knowledge in Rule-based Systems. *(ES2001) The Twenty-first SGES International Conference on Knowledge Based Systems and Applied Artificial Intelligence.* Peterhouse College, Cambridge, UK December 10th-12th (2001)

9. Håkansson, A.: *Graphic Representation and Visualisation as Modelling Support for the Knowledge Acqusition Process* . Ph D thesis, Uppsala University. ISBN 91-506-1727-3 (2003)

10. Håkansson, A.:Visual Conceptualisation for Knowledge Acquisition in Knowledge Based Systems. In: Frans Coenen (ed.): Expert Update (SGAI) Specialist Group on Artificial Intelligence, **7** No. 1, ISSN 1465-4091 (2004)

11. Håkansson, A.: An Expert System for the Environment Impact Assessment Method. Research Report 2004:1. Department of Information Science, Division of Computer Science, University of Uppsala, Sweden. ISSN 1403-7572. (2004)

12. Jackson, P.:*Introduction to Expert Systems* . Addison-Wesley Longman Limited. ISBN 0-201-87686-8.(1999)

13. Jacobson, I. Rumbaugh, J. & Booch, G.: *The Unified Modeling Language User Guide* . Addison-Wesley, USA (1998)

14. Merritt, D.: *Building Expert Systems in Prolog.* Springer-Verlag, New York Inc. ISBN 0-387-97016-9. (1989)

15. Negnevitsky, M.: *Artificial Intelligence A Guide to Intelligent Systems.* Addison-Wesley, Pearson Education ISBN 0201-71159-1. (2002)

16. Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., Van de Velde, W., & Wielinga B.: *Knowledge Engineering and Management the CommonKADS Methodology.* The MIT Press, Cambridge, Massachusetts, London, England. (2001)

Meeting the Need for Knowledge Management in Schools with Knowledge-Based Systems – A Case Study

Anneli Edman

Department of Information Science, Division of Computer Science
Uppsala University, Box 513, 751 20 Uppsala, Sweden
`Anneli.Edman@dis.uu.se`

Abstract. Usually the aim with knowledge management is to increase the profit, through capturing, storing, sharing, and utilising knowledge in an innovative way. Can knowledge management be of use in organisations as schools, where profit is of no interest? This article discusses the need for knowledge management from the perspectives of students, teachers and parents. In a case study master students in computer science developed knowledge-based systems for schools. These systems can be classified as support for (1) students' learning, (2) students and teachers regarding well being, (3) teachers' decision making. The conclusion is that intelligent systems, such as knowledge-based systems, could be used in schools, facilitating the knowledge management within the organisation.

1 Introduction

An important discovery of our time is that knowledge is the key to business and corporate success [11]. As a matter of fact, managers have realised that the only sustainable source of competitive advantage is the organisation's knowledge [12]. In a "knowledge-based society" it is vital to capture, store, share, and utilise knowledge in an innovative way, preferably in *intelligent systems* (IS), thereby creating new knowledge. *Knowledge management* (KM) is the discipline that deals with this process. Usually, when speaking about KM the aim is to increase companies' profit. There are several organisations, though, where profit is of no interest, in Sweden e.g. schools. Is there a need for KM in such an organisation and can it be successfully applied?

During 2004 there was a case study performed towards KM for school organisations. The aim was to investigate the possibility to use intelligent systems in form of *knowledge-based systems* (KBS) as support in schools. The case study was carried out within the master course Knowledge Management in the Computer Science program at Uppsala University. Uppsala municipality collaborated in the case study. During the course the students developed KBS for schools. The co-operation with the schools was deemed as a success and, therefore, it continued this year. This article will present the case study and an analysis of the result.

First, there is a discussion about the need for KM and KBS in the school organisation. This is followed by a presentation of the case study. Then the different types of intelligent systems developed will be described, of which two are further developed to real systems. Conclusion and further work end the article.

2   The Need for Knowledge Management in Schools

Knowledge management can be defined as management of organisational knowledge for creating business value and generating a competitive advantage [12]. Is there really a need for KM systems within the schools where it is difficult to speak about business values? After repeated discussions with teachers, parents, and representatives from Uppsala municipality[1] and the VLM Institute[2] the answer is undoubtedly yes. Let us study the arguments from the viewpoints of teachers, students, and parents.

The demands on *teachers* have, at least in Sweden, increased. They are, besides tutoring, responsible for a lot of documentation and taking care of the students' social well being. The documentation comprises, e.g., the students' absence from school, goals for the students, results at tests and grading. The well being may concern problems with drugs, eating disturbance, depression, and harassment. In some of these cases the teacher is imposed to deal with problems by contacting parents or the social authorities. Regarding the tutoring some students may need extra tutoring or special ways of teaching because they may have, e.g., dyslexia, the Asperger's syndrome, or damp. At the same time, as the teachers' responsibilities are increasing, the staff supporting the teachers is decreasing. The teachers could really benefit from a KM system that relieves the pressure regarding the administration and offers decision support in difficult situations in form of IS.

The *students* have more and more experience with interactive environments, e.g. through interactive websites. Integrating school activities into this environment has the potential of adding value to, and increasing the efficiency of the students' intellectual development and academic success. Services in a KM system for students could mean anything from getting information like schedules, goals for the education, criteria for the grading, results, interactive teacher support to actual learning environments. A learning environment may offer material, from e.g. libraries and museums, which the students can investigate and use to create their own learning material. The availability of interactive learning environments may support students with different cognitive abilities and different learning styles, thereby improving their learning and maybe also their willingness to study.

The *parents* expect that information regarding their children's education should be available through Internet or school intranets. They are interested in being informed about their children's presence, behaviour and results, in being involved in the schoolwork and able to take initiative in the communication with school. But, maybe even more important is that demands on parental involvement have been increased. The education in Sweden is since 1994 goal directed. The students are supposed to work, according to their prerequisites, towards personal knowledge goals. The teacher should be more like a supervisor and a mentor. It may be problematic for the teachers to adapt to this new role, to find relevant learning material, and especially to find time to handle this at the same time as the number of students increases. Thus,

---

[1] www.uppsala.se

[2] VLM is a Swedish abbreviation for Virtual Learning Environments, www.vlm.se. Eleven municipalities close to Uppsala own the insitute.

the parents have to participate in the education, which could be facilitated by a KM system.

Through a KM system for schools there is a possibility to get a more frequent, closer and natural *co-operation* between parents, students and teachers. The students can get access to e-learning environments, which is especially important when the education is goal oriented. The availability of extended interactive support during home work hours may reduce differences in academic results between students of different social backgrounds as well as reducing the impact of gender related issues. Moreover, the teachers' administrative workload may decrease and thereby increase their time for tutoring.

3 The Case Study

The case study was performed as a part of the course Knowledge Management, which is given once a year at Dep. of Information Science, Uppsala University, and comprises 10 weeks of full-time studies. 2004 the students numbered 76. The goal of the course is that the students should get an understanding of explicit KM within companies and organisations. The course content is focused on different types of IT-systems that can facilitate KM, not only how to work with KM within an organisation. This is in line with Awad and Ghaziri [1] who states that KM involves people, organisational processes, and technology. *Knowledge engineering*, utilised for developing KBS, can be seen as a part of KM [9]. Therefore, the students have the possibility to learn how to develop such systems.

In the course 2004 we collaborated with the VLM Institute and the municipality of Uppsala. The aim was to investigate if KBS could be used in the school organisation, facilitating KM within the organisation.

The students developed KBS principally for upper secondary schools. Before the course started a school teacher made an investigation, initiated by the VLM Institute, whether KBS are relevant for the school environment in Sweden or not. The pilot study showed that there is both a need and an interest in getting KBS supporting the activities in schools [7]. In the study several possible tasks appropriate for schools were proposed. Therefore, the municipality of Uppsala in collaboration with VLM provided us with experts within different areas related to schools, so that the students in the course could develop the requested KBS.

The development comprised three weeks of interviewing, implementing, testing, documenting, and presenting the work. The students interviewed the experts, in average, three times. The knowledge engineering methodology [9] was used to develop and document the prototype systems. The tools used for the development of the KBS were Match[3] built upon decision tables and Klöver [5], a rule-based shell implemented in Prolog. Finally, the experts tested the prototypes.

The work was discussed continually by students and teachers at project meetings during the course. In the end the students demonstrated their implemented KBS and these were evaluated and approved by other students, the teachers, and representatives from the VLM Institute and Uppsala municipality.

---

[3] www.knowledge-values.com

4   Knowledge Engineering for Schools

The knowledge engineering process involves identifying, capturing, representing and encoding knowledge followed by testing and evaluating the KBS [9]. As already mentioned, one goal with knowledge management is to create new knowledge. Nonaka and Takeushi describe knowledge creation and transformation as a knowledge spiral [10]. Creation of organisational knowledge develops through a continuous and dynamic interaction between tacit and explicit knowledge. This is in line with knowledge engineering and KM where both tacit and explicit knowledge is involved, and the emphasis is in capturing tacit knowledge.

Among the proposed 30 tasks in the pilot study [7], 16 were chosen and implemented in the course. I classify the developed KBS into three groups; support for students, for students and teachers, and for teachers, see Table 1. The first group, related to the students' learning, is a common direction when implementing KBS for schools. Wenger [13], e.g., has claimed that KBS offer an ideal basis on which to implement tutorial programs. Within our department we have experiences in developing such systems [3]. The other two groups are more rare. They are referring to the students' and teachers' psychological and physiological health but also to diagnosing students' abilities to be able to support them in their education.

It may be interesting to categorise the types of systems developed in alignment to types of problems solved by KBS [2]. We can find control, diagnosis, interpretation, and prescription[4]. Note, that in some of the cases it is difficult to do the classification. Control refers to governing the behaviour of an object to meet specified requirements. The systems in this group are number (5) and (11). In a diagnosis the system tries to infer malfunctions of an object based on observations. In the case study number (1), (9), (10), (13), and (15) belong to this group. An interpretation deals with inferring a description of a situation from, often noisy, data. Number (2), (6), (8), (12), and (16) can be seen as interpretations. Finally, when performing a prescription, recommended solutions to some kind of malfunction is generated. Here we find number (3), (4), (7), and (14).

Two of the prototypes have been further developed to *real systems*; the investigation of Asperger's syndrome and the classification of drug usage. The work is documented in two master theses in computer science. The work with the latter system will be briefly described.

The task for the drug system changed after discussions with the expert and the teachers. The reason was that it is more important to get support to determine whether a student is using drugs at all than to decide which drug the person may be using [8]. The problem for the teacher is to decide if he/she should contact the parents or the social authorities after being suspicious that a student may be using drugs. In some cases the law states that the teacher is obliged to contact the authorities. This is problematic, though, because several of the symptoms of drug usage are the same as ordinary teenage behaviour, and one would of course not like to take drastic measures without a very good reason [6].

The change of the system did not stop there. The user interface and the explanations in the system were further developed. Explanations are vital in KBS, especially

---

[4] The others are design, instruction, monitoring, planning, prediction, selection, and simulation

Table 1. The implemented knowledge-based systems categorised in different groups

Support for	Subject
Students	(1) Diagnosing chest pain (for nursing students)
	(2) Classification of senile diseases (for nursing students)
	(3) Generating advice for customers having stomach pains (for pharmacological students)
	(4) Generating advice when dying hair (for hairdresser students)
	(5) Practise for salesman (for business students)
	(6) Classification of mushrooms (for biology students)
Students & teachers	(7) Generating information about diet, exercise and weight
	(8) Estimating stress
	(9) Diagnosis of depression
	(10) Diagnosis and possible treatment of venereal diseases
Teachers	(11) Estimating the risk for harassment and generating guidance
	(12) Classification of drug usage
	(13) Diagnosis of dyslexia
	(14) Support for teachers' tutoring children with reading and writing problems
	(15) Investigating Asperger's syndrome
	(16) Evaluating the risk for heart diseases based on the life style

if the system should be a support when making decisions or facilitate learning. Important is to present the domain context, not only presenting the rules in the system [4]. In the drug system the rules are not explained at all but the system generates different texts, in form of canned text, dependent on the inferred conclusion. It is also possible to get information regarding symptoms that are related to different drugs. The system is now *integrated in an intranet for schools*, Rexnet, administered by TietoEnator[5]. If students, teachers, and parents are counted the intranet is available to 70.000 users in the regions of Uppsala and Sandviken.

5   Conclusions and Further Work

Awad and Ghaziri [1] states that "creating technology to help users derive tacit knowledge from explicit knowledge is an important goal of knowledge management". We have contributed with such technology by showing that KBS can be utilised in schools' for KM, both for teachers and students.

The KBS implemented in the case study can be seen as knowledge modules in a KM system. What the systems have in common is that they *provide explicit knowledge based on experts' implicit knowledge*. This knowledge is shared with the users and it can be applied, e.g. in learning environments or when making decisions. Furthermore, the knowledge provided may be integrated with the personal knowledge, i.e. going from explicit to implicit, which is in accordance with Nonaka's and Takeuchi's knowledge spiral [10].

When analysing the project the involved groups were very satisfied. The *students* appreciated working with real tasks developing prototypes that could be further de-

---

[5] www.rexnet.se

veloped, e.g. as knowledge modules in a school intranet. Moreover, they thought it was very valuable to meet and interview the experts, and also to participate in the concluding discussion with people from VLM and the municipality. In the *teacher group* we have followed students developing KBS during several years. For the first time we had the opportunity to develop required systems within one organisation. This has enabled us to get a deeper understanding of crucial variables within KM and KBS. We have also appreciated this collaboration with VLM and Uppsala municipality, thereby fulfilling the 'third task', i.e. the university's task to co-operate with the society. The representatives from the *VLM Institute* and *Uppsala municipality* had the opinion that their investment was very valuable and they appreciated the possibility to co-operate with the university. They got involved in a project where they had the possibility to test the KBS technique and see that it worked out well. VLM even offered the students to further develop some of the *prototypes to real systems*. Furthermore, they realised the potential and benefits of KBS and also the need for dealing with KM within their organisation. Thus, there is an affirmative answer to the question whether KM can be of use in organisations such as schools, where profit is of no interest.

Experiences from this project will be utilised in a regional EU project, OFFe (e-services in a public environment), where the course teachers will participate.

The systems implemented, based on knowledge engineering, are prototypes and should be seen as forming a case study before a more complete KM system is developed. Such a system has of course to be planned together with the people involved. Thus, in further work a more thorough investigation of the needs within the school organisation has to be performed, meeting the demands from students, parents, teachers and staff in the organisation. Furthermore, it is very important to test the knowledge-based modules carefully, both with experts and with the forthcoming users. The KM system could be based on an intranet such as Rexnet. Rexnet can, as a school portal, offer several of the services required in a KM system for schools. The drug system has shown that it is possible to integrate knowledge-based modules in Rexnet, thereby building an advanced KM system.

References

1. Awad, E. & Ghaziri, H.: Knowledge management. Pearson & Prentice Hall (2004)
2. Durkin, J.: Expert systems, design and development. Prentice Hall (1994)
3. Edman, A.: Combining knowledge systems and hypermedia for user co-operation and learning. PhD thesis, Computer Science, Uppsala University, Uppsala, Sweden (2001)
4. Edman, A.: Generating context-based explanations. Proceeding of ES2002, The twenty-second SGAI International Conference on Knowledge-Based Systems and Applied Artificial Intelligence. Cambridge, England (2002)
5. Edman, A., Håkansson, A. & Sundling, L.: Investeringsrådgivningssystemet Klöver (A consultative system for investments, Klöver). Dep. of Information Science, Uppsala University, Sweden (2000)
6. Edman, A. & Pettersson, A.: Digital rådgivare ett stöd för lärare (Digital advisor as support for teachers). Datorn i utbildningen (The computer in the education) 7:2004 32-34
7. Fredriksson, B.: Digitala rådgivare i skolan – kunskapssystem som ny tillämpning av IT i skolan (Digital advisors in the school – knowledge-based systems as a new application of IT in the school). Case study, www.vlm.se, VLM Institute (2004)

8. Jeppesen, S. & Lundén, A. (2005) SUNE – Ett beslutstödssystem för lärare om narkoti-kamissbruk (SUNE – A descision support system for teachers about drug abuse). Thesis for the Degree of Master, Dep. of Information Science, Uppsala University, Sweden.
9. Liebowitz, J. Knowledge management, learning from knowledge engineering. CRC Press LCC (2001)
10. Nonaka, I. & Takeuchi, H.: The knowledge-creating company. Oxford University Press (1995)
11. Nonaka, I. & Teece, D.: Managing industrial knowledge, SAGE Publications (2001)
12. Tiwana A.: The knowledge management toolkit, practical techniques for building a knowledge management system. Prentice Hall PTR (2000)
13. Wenger, E.: Artificial intelligence and tutoring systems. Morgan Kaufman (1987)

Aspects of Consideration When Designing Educational Knowledge Based Hypermedia Systems

Narin Mayiwar

Department of Information Science, Division of Computer Science, Uppsala University
Narin.mayiwar@dis.uu.se

Abstract. In the last decades, there has been extensive use of hypermedia systems in education as a tool for enhancing learning and teaching strategies. But the result of researches shows that little research effort has been directed towards making the presentation and navigation of a hypermedia system more knowledge-based, i.e. driven by knowledge about the information contents [2]. One approach to achieve this is to extend the hypermedia system with techniques from knowledge-based systems. In a hypermedia system chunks of information, stored in nodes, are presented in a non-linear way, similar to the way the human brain process information. Therefore hypermedia systems can be seen as naturally augmenting human abilities to access and process information. Development of a knowledge based hypermedia system for learning is not just about designing a few screens and specifying their relationship. It is also about pedagogical support for learning. Therefore, in this paper some aspects that are important to consider when designing knowledge based hypermedia systems for learning have been addressed.

1 Introduction

Systems containing knowledge can be utilized when tutoring both children and adults. There are a wide variety of these kinds of systems. One type is *Knowledge-based Systems*. This term includes all types of systems built on some kind of domain knowledge, independent of implementation [12]. These systems simulate human reasoning and judge capabilities by accepting knowledge from external sources and accessing stored knowledge through the reasoning process to solve problems [1].

In the last decades, hypermedia systems have also been increasingly used in learning environments. According to Arents & Bogaerts [2] present day hypermedia systems are powerful tools for multimodal (multiple access to data or ways of interacting with a computer) information storage and consultation. But, until now little research effort has been directed towards making the presentation and navigation of a hypermedia system more knowledge-based, i.e. driven by knowledge about the information contents. An important observation, in this respect, is that hypermedia systems have not been designed as educational tools. Indeed, the design has focused on information retrieval rather than learning [23] and learning is not the same as retrieving information [11]. One approach to achieve this is to extend the hypermedia system with techniques from knowledge-based systems. Therefore in this paper some aspects that are important to consider for creating effective and useful knowledge based hypermedia systems have been addressed.

The paper is structured as follows. First the paper gives a brief definition of learning and teaching approaches. Then a definition of knowledge based hypermedia system and the benefits of using it are introduced. Later on, some aspects of consideration when designing knowledge based hypermedia systems for learning are addressed. Concluding remarks and further work will complete the paper.

2   Learning and Teaching Approaches

The constructivist approach to learning dominates learning theory today [18]. Constructivists view knowledge as something that a learner actively constructs in a knowledge-building process, piece by piece.

Entwistle [8] has defined four orientations to learning: meaning, reproducing, achieving and holistic. A combination of these four orientations together with external factors, such as the need to pass examinations or the interest for the subject, can lead to learning strategies which categorized certain approaches to study, from deep to ground levels of thinking [3]. There are two main strategies for learning: ground and deep. In a ground approach to learning the students focus on memorizing set of facts, reproducing parts of the content and thereby developing an atomic view. The deep approach to learning takes place when students focus on significant issues in a particular topic and reflect on what they have read, relating their own previous knowledge to the new knowledge they have obtained. The students look for the overall meaning of the material and thereby develop a holistic view, which is desirable. In traditional classrooms teachers focus on getting the students remembering as much as possible, while the new pedagogy helps students focus more on knowing what to know, where to find and how to store knowledge [13].

3   Knowledge Based Hypermedia System

According to Edman [5] hypermedia systems and knowledge systems can be seen as a complement to each other. Hypermedia is a system that presents chunks of information, stored in nodes, in a non-linear way. And the term knowledge based system includes all types of systems built on some kind of domain knowledge, independent of implementation. These systems simulate human reasoning and judge capabilities by accepting knowledge from external sources and accessing stored knowledge through the reasoning process to solve problems [1]. Knowledge technology offers methodologies for program and system development for complex domains. Therefore, it has been a natural step to apply these techniques to hypermedia [6], [24] resulting in what may be termed *knowledge-based* or *intelligent hypermedia*.

Edman and Hamfelt argue that a hypermedia system can be called intelligent or knowledge based if knowledge-based computer reasoning sets up its links. Knowledge based hypermedia systems must be able to dynamically compose pieces of informal knowledge to present to the user. It should also be capable of guiding the user to static informal knowledge [5], [6].

The advantages with using knowledge based hypermedia in a learning environment are related to the information that can be reproduced in such systems and to the different ways to attain this information. The technology facilitates the storage of large amounts of information. This information may well represent different views of a

domain [4]. For learning purpose it is very important that the users can study critically and build their own opinion. A textbook describes just the opinion of the author but in a hypermedia system different view can be offered [19]. Moreover, it is easy for the user to find different kinds of associations by tracing the links. Navigation, on the other hand, can support the search for specific information.

Several aspects need to be taken into account when designing knowledge based hypermedia systems for learning purpose. Some will be mentioned in the following section.

4   Some Aspects of Consideration

Development of a knowledge based hypermedia system for learning involves several pedagogical aspects. It is not just about designing a few screens and specifying their relationship. It is also about pedagogical support for learning. Therefore, in this section some aspects that are important to consider when designing knowledge based hypermedia systems for learning have been addressed.

Individualization
When designing knowledge based hypermedia systems for learning, it is important to consider individual differences. Individual differences appear to have a big impact on human-computer interaction. Human beings learn all the time, but we learn in a variety of ways. One person can have a good memory for what he/she has heard, while another one needs to see an image or read a text to remember. The reason may be that humans have different kinds of intelligences [9] and have developed different learning strategies. By offering students alternative ways such as an individualized knowledge based hypermedia system we may help them to achieve deep learning [7], [13].

Different Representationforms
Some important issues when developing a system supporting learning are that different types of media, e.g., text, pictures, animations, and digitized speech, can be used in such a system. This will support people with different learning styles, to learn in accordance with their intelligences, as mentioned above.

Transparency and Organization of Content
Knowledge based hypermedia system's transparency could be discussed in terms of the *usability* of the system [5]. She argues that the usability of a knowledge based hypermedia system can be seen as a combination of the usability of the underlying system engine and the usability of the contents and the structure of the information in the system, as well as by how the contents and the structure fit together [19]. There have been many discussions about the difficulty of navigating in a knowledge based hypermedia system and the *lost in hyperspace* problem. Therefore it is important to consider design features that will support users in their knowledge building when creating knowledge based hypermedia systems.

In knowledge based hypermedia system the user is involved in the problem solving process in co-operation with the system. It is therefore very important that the system is transparent when implementing knowledge systems for learning. Learning and understanding are closely linked together and according to Schank [22] "explanation is critical to the understanding process".

   If the knowledge presented in the knowledge base is not transparent, the user has to pay more effort finding the appropriate knowledge instead of concentrating on solving the problem.

Learning Support

According to Nunes and Fowell [20] learning is "a process of constructing knowledge and developing reflexive awareness, where the individual is an active processor of information. They argue that learning occurs through interaction with rich learning environments, and results from engaging in authentic activities, and by social interaction and negotiation". Experiences from, e.g., using the knowledge-based system Analyze More [5] showed that a cooperation within and between groups of students took place in the classroom. Collaborative discussions when working with the system can encourage a dialogue beneficial to effective learning, according to constructivists approach to learning.

   The most essential components of an effective knowledge based hypermedia for learning, are well-defined goals. It has been argued that it is important that the system architecture reproduce a coherent domain theory. Through the explanations of the context the knowledge can be displayed in different ways as various overviews of the domain. Another important aspect is that the links between the chunks of information are not static, which means that the explanations can be adjust to the current user. Furthermore, the system should handle mixed-initiative dialogues; i.e., both the user and the system can control the interaction, since the access to the system's knowledge can be facilitated. As mentioned before, in order to learn better the users must be active, in their learning. This can be achieved by co-operation between user and the system co-operate in the reasoning [5].

Visualization of Knowledge

"Visualization is a tool to enhance productivity and quality" [16]. As mentioned before, in knowledge based hypermedia system knowledge is stored in a knowledge base, which can be enormous. Furthermore, there are no restrictions on how users navigate through the information, they can become disoriented while navigating, or with other words get lost in hyperspace. There are technologies like Overview diagrams or site maps, which can be used to visualize the structure of the knowledge in the knowledge base. They help users in their orientation and navigation in hypermedia documents.

   Different views of a system can also support different users [10]. The users will have the opportunity to choose the best suitable view of the system. By view handling it is possible to show static content and dynamic problem solving which gives the possibility to transfer reasoning via the system from one user to another one.

Human Memory Capacity

According to cognitive psychology the human memory has limitations in the capacity to store and process information. Miller's [17] theoretical review of a "magical number seven, plus or minus two" shows the limitation in short-term memory (STM) storage capacity. Therefore it is significant to consider this aspect when designing knowledge based hypermedia systems for learning, i.e. by chunking information, which can lead to an increase in the short-term memory capacity. Interference often causes disturbance in short-term memory retention. This accounts for the desire to complete the tasks held in short term memory as soon as possible.

Meaningful Interface

How much new information we human beings can store in the memory depends on its meaningfulness. In knowledge based hypermedia system the information presented to the user i.e. the name of links and icons must have a meaning for the user so she/he can easily recognize the content and the search path [15]. Experiences from, e.g., using the knowledge-based system Analyse More showed that the main interaction with the system has been constrained to one central node [7]. Emphasis has been given to the design of an interactive map, intended to visualize and reinforce the conceptual structure, connections and concordance among important objects of the domain.

Feedback

There is no doubt that feedback is a fundamental factor in the learning process. Feedback can be defined as information about what happened, the result or effect of our actions. As mentioned before people learn in different ways therefore differentiated feedback should be used with a combination of animation, text, colors and graphics, to appeal to different cognitive areas and maintain the users' motivation [7].

5   Conclusions and Further Work

Knowledge based hypermedia system can offer support for individualized learning. In a traditional learning environment the teacher cannot support all forms of intelligence, which means that some users' needs, will probably not be satisfied. To give the users the opportunity to learn in accordance with their abilities, the education provided has to be individualized. Therefore, the main objective of this paper has been to highlight the importance of using knowledge based hypermedia systems in education. But as mentioned before, creating knowledge based hypermedia system for learning is not just about designing a few screens and specifying their relationship. It is also about pedagogical support for learning. We have to consider some pedagogical aspects in order to develop effective and useful knowledge based hypermedia systems for learning. Some of them have been mentioned in this paper.

As a further work the identified aspects will lay the foundations for a set of design guidelines for designing suitable knowledge based hypermedia systems for learning.

To conclude, knowledge based hypermedia systems have many important contributions to make for education, but only when we have a good understanding of its technical and pedagogical aspects.

References

1. Anderson, R.G.: Information & Knowledge Based Systems. An Introduction. Prentice Hall International, Great Britain (1992)
2. Arents, H.C, Bogaerts W.F.L.:
   http://www.comp.glam.ac.uk/~NRHM/h-volume3/ha-1991-7.htm
3. Capel, S., Leask, M. & Turner, T. (1999) *Learning* to Teach in the Secondary School. London: Routledge.
4. Conklin, J.: Hypertext: An introduction and survey, Computer, **20**, (1987)
5. Edman, A.: Combining Knowledge Systems and Hypermedia for Use Co-operation and Learning. Ph thesis, Computer Science, Uppsala University, Uppsala, Sweden (2001)

6. Edman A. & Hamfelt A. (1999) A system architecture for knowledge based hypermedia. International Journal of Human-Computer Studies, **51**, pp. 1007-1036.

7. Edman, A., Mayiwar, N.: A Knowledge-Based Hypermedia Architecture Supporting Different Intelligences and Learning Styles. Proceedings of the eleventh PEG2003 Conference. Powerful ICT for Teaching and Learning, St. Petersburg, Russia (2003)

8. Entwistle, N.J.:Styles of Learning and Teaching: An Integrated Outline of Educational Psychology. Wiley: Chichester (1981)

9. Gardner, H.:Intelligence reframed. Multiple intelligences. New York: Basic Books (1999)

10. Hakansson, A.: *Graphic Representation and Visualisation as Modelling Support for the Knowledge Acquisition Process.* Uppsala. ISBN 91-506-1727-3 (2003)

11. Hammond, N.: Learning with hypertext: problems, principles and prospects. In C. McKnight, A. Dillon & J. Richardson, Eds. Hypertext, a psychological perspective, pp. 51-69. Ellis Horwood (1993)

12. Hayes-Roth, F., Waterman, D., Lenat, D.: Building Expert Systems, Addison-Wesley (1983)

13. Lovless, A., Ellis, V.: ICT, Pedagogy and the Curriculum: London: Subject to Change, RoutledgeFalmer (2001)

14. Mayiwar, N.: Hypermedia som ett pedagogiskt verktyg (Hypermedia as a pedagogical tool). Thesis for the Degree of Master of Science in Computer Science, No. 38/01, Division of Computer and Systems Science, Department of Information Science, Uppsala University (2001)

15. Mayiwar, N., Hakansson, A.:Considering Different Intelligences and Learning styles when Transferring Knowledge from Expert to End User. Proceedings of the eleventh KES2004 Conference. (2004)

16. Mazaika, Robert J.: "A Visualization Tool for Analysts", Aerospace & Defense Science (1990)

17. Miller G. A, The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information, *The Psychological Review, vol. 63, pp. 81-97* (1956)

18. Morphew, V.N.: Web-Based Instructional Learning/[Edited by] Mehdi Khosrow-Pour. USA: IRM Press (2002)

19. Nielsen, J.:Multimedia and hypertext, the Internet and beyond. AP Professionals, Academic Press (1995)

20. Nunes, J. & Fowell, S.P. Hypermedia as an experimental learning tool: a theoretical model. Information Research New, 6 (4), 1996.

21. Pohl, M. & Purgathofer, P.:Hypertext authoring and visualization. Int. Journal of Human-Computer Studies, **53**, (2001).

22. Schank, R.C.: *Explanation patterns, understanding mechanically and creatively*. Lawrence Erlbaum Associates (1986)

23. Whalley, P.: An alternative rhetoric for hypertext.. In C. McKnight, A. Dillon & J. Richardson, Eds. *Hypertext, a psychological perspective*, pp. 7-17. Ellis Horwood (1993)

24. Woodhead, N.: *Hypertext and hypermedia, theory and applications*. Sigma Press & Addison-Wesley (1991)

Semantic Tags: Evaluating the Functioning of Rules in a Knowledge Based System

Torsten Palm

Department of Information Science, Computer Science Division, Uppsala University
Box 513, SE-751 20, Uppsala, Sweden
Torsten.Palm@dis.uu.se

Abstract. During the development of a software system, many errors, both physical and logical, sneak into the source code. However, the system's ability to work properly depends on the ability to find these errors and get rid of them. To this objective, a controller is needed to find and eliminate both types of errors. In this paper, we suggest an approach to a semantic controller, including a syntactic check, by applying tags to production rules. The syntactic check of the controller takes care of the physical errors, and the semantic check takes care of the logical ones.

1 Introduction

Quality assurance and security issues are important when a knowledge system is developed. This increases the feasibility that the 'right' system is developed since the system satisfies the guidelines for high quality assurance. Quality assurance includes completeness and consistency. Consistency concerns the development and updating of the knowledge base, where the consistency checking of the knowledge base can help the knowledge engineer to assure the system's reliability. This reliability is especially important in critical applications where failure can mean disaster, for example in a system for plants.

Completeness checking tests whether a system is able to find conclusions, or give answers, to all reasonable situations in its domain of expertise (Polat & Guvenir, 1993). A system is complete when everything that can be derived from the data, really (actually) is derived. The knowledge base is prepared to reach conclusions for all possible situations that could arise within its domain. Completeness checking involves finding logical cases that have not been considered, e.g., missing rules. One example of a completeness problem is the dead-end rule since the dead-end rule has a premise that can never be reached.

The quality assurance problem includes incompleteness and inconsistencies in the knowledge base, which can arise because human experts are not prepared to provide all the knowledge needed in one complete and consistent piece (Polat & Guvenir, 1993). In this case, the knowledge base must be modified accordingly to achieve a correct functioning (Lopez et al, 1990). This quality assurance process can be simplified by the development of automatic verification or validation tools (Mengshoel, 1993).

However, the quality assurance, including consistency and completeness, only takes care of syntactic problems with the source code, namely physical errors. But a powerful controller also has to take care of the semantic meaning of the source code

to avoid logical errors. To this objective, it is interesting to develop a combined syntactic and semantic controller. This controller has to check the knowledge base for physical errors as well as for logical errors. The syntactic part of the controller handles the physical errors and the semantic part handles the logical errors.

The difference between a debugger and a controller is that the debugger is a syntax controller, whereas the semantic controller handles concepts or so called semantic tags. This paper focuses on the semantic tags.

2   Errors in the Knowledge Base

The production of large-scale, rule-based knowledge systems implies that a large number of rules are implemented into a knowledge base. As new ideas about improvements of the rule-based knowledge base are presented, new rules are simultaneously added, and existing rules are modified. Thus, a rule-based knowledge based system must be regarded as a system that continuously changes. In addition to the adding of new rules and the modification of older ones, similar activities are going on, namely, the introduction of new relationships and the modification of existing relations. A relationship is a connection between two rules or a reference to another rule where one rule is pointing to another.

When we make changes in the source code of ordinary software, we expect that our efforts will improve or minimise errors or so-called bugs. However, the consequence of the ambition to improve and correct may cause unexpected results, and new types of errors can be created and placed at unpredictable locations in the source code.

The same type of problem may arise in a rule-based system. This is the motive for finding methods to investigate the rules and the relationships in a rule-based system.

3   The Controller

It is not unusual that a rule-based system consists of hundreds or even thousands of rules. Roughly, these rules can be categorised into different types: rules that only make use of user-given facts, rules that only operate on other rules, and rules that only handle conclusions.

With such a great number of rules and all kinds of relationships, it is a big problem to check that all rules in the knowledge base have been fired or utilised during the interpretation of the rules. For this reason, it is valuable to know whether certain rules have been called by another rule or not. If it obvious that some rules have never been fired at all, it may be interesting to find out the reason for it. In addition, the contents of the rules need to be investigated.

Most likely, no other rules are referring to these unused rules. Therefore, these rules are not involved in the consultation process. Thus, these rules are isolated from the rest of the rules. Another reason can be that these rules are not using facts supplied by the database or given by the end user. If this is the case, special action must be taken to find them. An action is required that looks for rules that are unexploited by other rules. This action also has to check whether these rules utilise facts or not. If the rules do not employ facts or rules, they will never be fired during a consultation.

To maintain a rule-based system, a type of supervisor or a controller is needed. The controller can be activated constantly or for a short period of time. Controllers should

be modelled in such a way that their actions can be performed automatically whether they are used constantly or only for a while. For this reason, the automatic controller will be a program that examines the knowledge base regularly. After accomplishing the interpretation of rules, the information about the found errors is collected and presented in a report, which is handed over to the user. The contents of the report gain in readability if some parts are presented visually.

Every time a rule is fired, an announcement notifies this action (vem blir "notified" om "this action"? Ex. "… an announcement notifies this action to the user."). Such an announcement is applied inside the rule to inform whether this particular rule has been activated or not. Depending on the type of rule, information on the type will also be reported.

Information on the type of rules that another rule has called is stored inside the rule itself. The rule will also produce a message that is sent to the controller. This message contains information on the rule that has become activated. The controller will also be notified of how many times a rule has been employed and of the usage of relationships, i.e., relations referring to other rules, when this has been established. Information on the rule will be kept in the temporary knowledge base. The temporary base will contain all the rules that have been used, but the contents of these rules are extended compared with the rules in the knowledge base. Consequently, the more the knowledge base is used, the more information the controller will be able to access.

There are at least two kinds of information pieces that the controller can store: information on the syntactic construction of the rules and information that will be assigned to the knowledge itself. The first category of information is easily detected by the syntactic controller. The second type of information concerns errors that will be handled by the semantic controller.

Consider what happens if a statement is completely erroneous. The relationship that relates one rule containing the erroneous statement to another rule is incorrect. If we continue to follow the relationships starting from these defective rules, we will construct a chain of information that will be incorrect.

We have to decide what kind of data that has to be picked up from a session. It depends on the kind of results we expect and how these results are structured and visualised. The more effort that is spent on the construction and structure of the knowledge base, the fewer the semantic errors that will be produced. Therefore, it is important to put considerable effort into the preparation of these structures.

3.1   Requirements on the Controller

As is the case with traditional programme construction, the semantic construction of rule-based systems (i.e., the individual rules and rules with relationships) and the logic construction of the rule system have to be performed in accordance with methods that are already established on the market. However, when it comes to the construction of knowledge-based systems consisting of rules and relationships, the number of accessible methods for modelling is lower.

We need to find out what kind of requirements we have on the correctness of the semantics applied in a knowledge-based system. Assume that a ruled-based system consists of rules and relationships between the rules. The structure of the rule-based system can be compared to a net or a tree, and a search in this tree structure usually

starts from the top and moves to the bottom, as well as from the left to the right, of the tree.

The first and most primitive semantic checking of rules is similar to the syntactic checking. The most advanced semantic control of the knowledge base is difficult to implement as a debugger. It is, however, possible to a certain degree. In this case, the semantic controller has to utilise information picked up from the most advanced automatic semantic checking.

To make the semantic analysis as effective as possible, the control has to be performed step by step, and the smallest possible part of the analysis has to be made on a simple and rudimentary level. Therefore, it is important to decrease the complexity of the semantic assessment, especially if the knowledge base contains a small set of rules or is non-complex, i.e., it contains (a) few relationships between rules. Consequently, it is of essence that the control can enhance (kräver objekt eller "that the control is gradually enhanced as the…") gradually as the complexity of the knowledge base increases.

To minimise the complexity of the controller's work, the easiest semantic assessment will be carried out first. Then, if needed, more profound assessments will take place. If the result from the first control turns up empty-handed, it is probably of no use to continue to the more advanced level of the quality control.

3.2 Semantic Tags

Since we want to perform checking of quality control, we introduce the idea of setting tags inside the individual rules. The main task of the tags is to collect data. But the tags do not only collect quality data of the semantic content inside rule, they also perform certain assessments, i.e., they infer to take certain actions from the inside of the rule.

Our intension is to let all rules be equipped with individual functions for collecting and evaluating the knowledge base by using tags. This implies that there will be a large amount of basic data with which to perform a central control of evaluation. For this purpose, a central function for the collecting of rules is required. This will be part of the semantic controller.

A supervision system will work on a low and more concrete level and on a high and more abstract level. This gives an opportunity to create groups from the lower parts and to process the information stored in the groups. The result of the evaluation of the lower groups will be transferred to a higher level in order, finally, to reach the highest part.

Thus, the semantic evaluation is carried out by gathering results from the individual tags of the rules. Accordingly, it will be the gathering of the results from the different groups of rules that will finally evaluate the complete knowledge base of the system.

3.3 An Example of Using Semantic Tags

The individual rule is equipped with a semantic tag. It consists of utilities for recording data in semantic tags and in a semantic controller. Recorded data in semantic tags will be interpreted by a local semantic controller and be presented to the user.

A semantic tag is defined as information chunks (pieces of information?) in which all events that take place are collected and stored in a rule. Examples of events are:

− Number of times the rule has been activated;
− Other rules that have called the rule that is under consideration (in question?);
− Number of times the inference mechanism has succeeded in interpreting the rule;
− Number of times the rule has called facts (or questions);
− Number of times the rule has used other rules;
− Number of times the rule has an inference of its own.

Some events are the same for all rules, which is why there is no reason to store them in local semantic tags. It is better to store them as a global semantic tag at the highest level of the group of rules. An example of a global semantic tag is:

− Number of times the rule-based system has been used.

The structure of a rule can influence what type of information that is stored in a semantic tag. The structure also influences what kind of information that will be evaluated.

The following example shows some rules from a knowledge base in a ruled-based system that handles an assessment of the environmental impact of a hydropower and dam site building on the surrounding environment (Hakansson, 2004). The rule number 57 is a common rule in the knowledge base:

```
rule(57,ca_sedimenthighsoil_rule,ca_sedimenthighsoil_text,0)
:-
        reply ( sedimenttransport_object, 'High' ) ,
            ( reply ( soilerosion_object, 'Small' ) ;
            reply ( soilerosion_object, 'Moderate' ) ) .
```

The `ca_sedimenthighsoil_rule` is a geological categorisation, and `ca_sedimenthighsoil_text` is the rule object that is used by the semantic controller to group rules. The `ca_sedimenthighsoil_text` is the conclusion, which is presented to the end user.

This rule uses three(two?) different facts—transport of sediment and soil erosion. The value of the fact "transport of sediment" is high, and the value of "soil erosion" is either small or moderate.

The information that is added to the rule concerns the number of calls to the rule, the rule/rules calling a particular rule and also the number of times the rule has been fired. This means that it is possible to gather the following facts, which can be placed into a semantic tag of the rule:

− Number of times the rule has been activated;
− Other rules that have called the particular rule under consideration (in question?);
− Number of times the inference mechanism of the rule has done (accomplished?) a piece of work.

The above is explained in the following example:

```
temporary_rule( 57, ca_sedimenthighsoil_rule,
ca_sedimenthighsoil_text, Number, Number_of_calling, Fired, Num-
ber_of_fired, Calling_rule, 0) :-
        Number_of_calling is Number +1 ,
        assert ( Calling_rule ),
```

```
reply (sedimenttransport_object, 'High' ),
( reply ( soilerosion_object, 'Small' ) ;
reply ( soilerosion_object, 'Moderate' ) ) .
Number_of_fired is Fired +1.
```

Firstly, the temporary rule increases the number of times the rules have been activated. Secondly, it asserts the rules that have called the current rule. Thirdly, the rule checks the facts of sediment transport and soil erosion. If these facts are found to be true, the number of times the inference mechanism has completed its work with this rule is increased.

The order of the tags is important since each time the rule is invoked it will increase the number of callings and with all the rules that have invoked this particular rule (tror denna mening behöver omformuleras lite för att bli tydlig). But the number of times this rule has been fired depends on the contents of the rule. If the content is fulfilled, the times of firing will increase. Otherwise, it will continue to have the same value as it had before the invocation.

4   Conclusion

This paper describes 'semantic tags' and the way they can be used as indicators to record a selected number of special events. The purpose of the events is to serve as information on semantic status of individual rules, but they will also be a part of collected information that will be used by an overall semantic controller. The events can be processed by a semantic controller, which means that the quality of the status of the semantic (semantic vad?) will increase.

The function of the interpreter of local and global semantic controllers is not dealt with here. However, it will be (is?) possible to catch the remaining eligible events by increasing the rule with the function of a semantic tag, which is desirable. From the example of the rule, it is clear that it is possible to pick up desirable facts about events for a semantic tag without too much violation (of vad?). Transparency is important in order to present the information the users. The information can be emphasised through a visual presentation in order to increase the transparency during interpreting the rules.

In our further work, we will report on the treatment of the events of semantic tags by a semantic controller. Furthermore, a thorough perusal of a complete rule-based system will be presented.

References

1. Hakansson (2004): An Expert System for the Environment Impact Assessment Method. Research Report 2004:1. Department of Information Science, Division of Computer Science, University of Uppsala, Sweden, 2004. ISSN 1403-7572
2. Halle von, Barbara (2002): Business Rules Applied. Wiley
3. Polat F. & Guvenir H.A (1993), UVT: A Unification-Based tool for Knowledge Base Verification. *IEEE Expert,* June 1993, No. 3.
4. Lopez B., Meseguer P., Plaza E. (1990), Knowledge Based Systems Validation: A State of the Art. *AI Communications*, June 1990, Vol 3, No. 2, pp 55-71.
5. Mengshoel O. (1993), Knowledge Validation: Principle and Practice. *IEEE Expert*, June 1993, No. 3, pp 62 - 68.

Knowledge Management for Robot Activities in a Real World Context.
A Case for Task Pattern Analysis (TAPAS)

Lars Oestreicher

Department of Information Science, Computer Science Unit, Uppsala University
Box 513, SE-751 20 Uppsala, Sweden
`Lars.Oestreicher@dis.uu.se`

Abstract. Knowledge management is used for handling knowledge and activities in real world context. However, knowledge management can also be used for the modelling of robot activities in a real world context. Although robot technology is still under development, Intelligent Service Robots are slowly becoming a reality. The programming of these robots to support a closer interaction with the users is a complex problem including the difficulty of correct modelling of the tasks of the robots. The need for close cooperation between user and robot adds extra complexity, compared to standard models for task modelling. In this paper we outline a representation scheme, Task Pattern Analysis (TAPAS), that is directed towards a notion of activities as rule-based behaviours. In task pattern analysis, the tasks are described in terms of Task Patterns, consisting of a frame rule, and Task Fragments, specifying and constraining the applicability of the rules.

1 Introduction

There is currently ongoing research on (semi-) intelligent service robots (ISR), which will serve as a type of butler or servant in the house. The needed technology is under development, both in terms of anthropomorphic robots [1] and machinelike self-propelling agents already sold on the market, such as the autonomous vacuum cleaner [2]. Regardless of their physical shape, these ISR:s are facing interesting challenges from the perspective of knowledge representation. One serious consideration is known as the frame problem [3]. In short, the frame problem concerns how to handle defaults in a changing environment. Any agent that is moving in a dynamic world, either needs to keep an account of what is changing in the world, or continuously reason about the current state of the world. The robot has to consider possible changes in the goals that are posed by the dynamics of the environment. Since the ISR type of robot is moving around physically, it needs to be especially aware of the world in order to avoid new obstacles. Such obstacles could be things that have been temporarily placed in new places, e.g. a chair that has been placed in a corridor.

Furthermore, a robot needs to be able to act upon the knowledge it has of the world. To achieve this in a general way is naturally recognized as a difficult

task. In this article we discuss one possibility to model robot tasks in a dynamic context in a restricted manner, using TAsk Pattern AnalysiS (TAPAS). The TAPAS method is based on the notion of work tasks (for a robot) as constructed from a well-founded set of activities.

2   Intelligent Service Robots

The application area for the TAPAS method is a specialized domain of robotics, namely ISR for people with special needs. The basic assumption is that intelligent machines could be a useful support in homes or workplaces. In this area there are several important factors that influence the design of the robot activities. The robots need to work in cooperation with the intended user using advanced modes of interaction, such as natural speech or multimodal interfaces [4]. The modes of operating the robot have to take the communicative capabilities of the user into consideration, as these vary depending on e.g., whether the user is paralysed or immobilised [5].

Apart from the close cooperation with the user there are other considerations on the interaction design for the robot. The requirements on the functionality for the robot are highly individual, depending both on the kind of disability the user displays, as well as the personal preferences with the specific user. One such fundamental consideration, established from an interview study made by the author, is that the robot in many cases will not be requested as a performer of tasks, but rather as a catalyst or facilitator. Thus the robot should be enabling the user to perform activities in contrast to doing things for the user autonomously [6]. This perspective of facilitating provides certain crucial constrains on the description methods. One such constraint is the need for the robot to be able to ask the user for advice at certain points or in the case of problems appearing when performing the task.

In their role as task supporting agents, robots will need precise knowledge about the situational context. This is relatively easy for aspects such as navigation and location identification. However, the robot also needs knowledge about objects and properties of objects in order to perform certain tasks. Suppose, for example, that we ask the robot to pick up a wallet from the floor. In order to do so it needs to know what a wallet is, whether it is fragile or not, whether it contains something that can be spilled or not. If we, on the other hand, ask the robot to fetch a blue bowl in the living room, the robot needs to know that it should bring the bowl including its content, rather than emptying the contents of pop corn on the table before getting the bowl. Thus, task modelling for service robots incorporates a need for a useful representation of common sense knowledge concerning the tasks that need to be accomplished.

The robot itself, through its physical design, defines a number of minimal activities or *basic tasks*. Examples of such basic tasks are "Go to a location", "Grasp object", "Locate object", etc. Currently robots are, for example, not sufficiently advanced to locate an object on a cluttered desktop but much research is done on image recognition, gesture recognition, etc. Thus, it will be feasible

to assume that future robots will have at least a limited working functionality of this kind. How to represent this kind of behaviour will, therefore, be more and more important. It is desirable that the instruction of the robot is simple to handle for the final user, once the initial programming of the robot has been finished.

Several task description methods have been investigated for this area, such as Hierarchical Task Analysis (HTA) [7] and Extended Task Action Grammar (ETAG) [8, 9]. The purpose has been to find common properties of task descriptions for robots. For most cases these descriptions:

- tend to describe shallow structures
- tend to generate numerous similar cases
- displayed a similarity in structure between different tasks
- do not manage to describe situational aspects of the tasks

From these attempts it turned out that the traditional methods used for task analysis are not optimal for the modelling, which is necessary for the problem domain, and that the specification needs to be more extensive in order to facilitate a useful modelling of the contexts. In previous research it has been shown that tasks using artifacts can be modelled in a formal description, such as a pure logic programming language [10, 11]. With these logic descriptions, electronic mail tasks can be represented in a way that allows for simulation of the tasks. However, these descriptions are still insufficient to describe tasks for intelligent service robots. Therefore, a new description method has been developed, aiming at minimizing the problems described above. The method is based on traditional methods but combines the basic ideas for task modelling with a slightly different perspective on task analysis.

3   The Task Pattern Analysis Method, TAPAS

Work tasks in general, and specifically in situations where robots are involved, tend to consist of a set of activities that display common properties, making it possible to generalize the descriptions. The general descriptions furthermore display common features (which facilitate a description as patterns) as well as differences (which make it interesting to describe them as patterns rather than fixed rule structures). The common features are expressed in terms of patterns, and the differences are expressed by filling in variable parts of the patterns, or by specifying general constraints.

Many previous methods for task analysis and modelling have relied on the existence of these similarities, trying, e.g., to catch them in hierarchical structures. However, in many cases the similarities occur on a lower level of abstraction than allowed for by the modelling method. Moreover, usually these similarities can not be described in such a way that the advantage of finding these similarities can be utilized in full. Using the more or less standard method of HTA [7, 12] tasks of saving files can be merged into one single hierarchical task structure, which is a tree composed of goals and subgoals. Such a description may fit many

different cases. The HTA *structure* does not directly address how to describe such things as file names, or catalogues for the saving of the file. This means that the HTA structures can be similar for a large number of tasks, but also that a certain structure can be applied to different cases that are not represented by the meaning of that particular HTA description. The use of plans does not help out much since the plans only describe the order of the goals. Different tasks may still be described by the same general plans.

In order to provide a solution to this problem, Task Pattern Analysis (TAPAS) was invented. The basic construct of a task pattern is a rule, e.g. as in pattern 1. The task is described as a pattern (first line) and a basic action structure (lines two to six), complemented by qualifiers (variables), so called task fragments, that are used to describe some conditions on the applicability of the pattern (lines seven and eight). Two of the task fragments, `<agent>` and `<place1>`, are defined in this pattern as the robot and the robots location. Saying ... `<agent> is_one_of [robot, John]` ... would make this pattern applicable for both the robot and the human actor "john". By allowing only the robot to be the `<agent>` we restrict the pattern to the robot.

Pattern 1 *Sample pattern expressing the action of getting an object from a place in a room.*

```
<agent> gets <object> from <location2> at <room1> is
     [ <agent> goes from <room1> to <room2>,
       <agent> goes from <location1> to
                          <location2> in <room2>,
       <agent> grasps <object> at <location2>,
       <agent> goes from <room2> to <room1>  ]
                     where <room1> = <agent>.current_room()
                     and <location1> is <room2>.entry_point()
                     and <agent> is_one_of [robot].
```

In the description there are several predefined building blocks. A special entity is the use of function references such as

```
                    <room1>.entry_point()
```

These entities are special functions that are tightly connected to the robot's functionality. In this case, the function refers to the robot's entry point, i.e. the doorway of the room. These low level functions are programmed in the same way as the basic patterns.

The most basic activities are not divisible, i.e. they are *physical robot behaviours*. They can be described as *basic patterns*, which are not expanded (cf. patterns 2 and 3 below) but may still use constraining task fragments. In pattern 2 a basic pattern is displayed, constrained only in that the robot cannot go from a place to the same place in one action. Such constrains may be expressed alternatively as meta rules, but most of the time it is clearer to express them explicitly in the rules.

Pattern 2 *A basic pattern matched to a specific behaviour programmed into the robot.*

```
<agent> goes from <room1> to <room2>
   where <room1> is_one_of [kitchen, study]
   and <room2> is_one_of [kitchen,study]
   and <room1> =/= <room2>
```

Note that in pattern 2 there is no further definition of the task, and there are only constraints regulating the applicability of the rule. This is an indication that the rule is matching a specific behaviour in the robot. Rules of this kind are typical *basic patterns*.

In this example, the task fragment `<agent>` has not been constrained. However, since this pattern is used in pattern 1 and the task fragment `<agent>` is constrained in the latter, the constraint is applied to this sub rule as well. The same holds for the constraints on `<room1>` and `<room2>` in pattern 2, which are automatically applied to pattern 1. In this way, sub or super patterns may constrain each other. Also if there are alternative sub patterns that match for certain super patterns, constraints will direct the continued instantiation of the sub pattern.

Pattern 3 *A second example of a basic pattern. This pattern shows the movement of the robot within a room.*

```
<agent> goes from <location1> to <location2> in <room>
   where <room> is_one_of [kitchen, study]
   and <location1> is_within <room>
   and <location2> is_within <room>
   and <location1> =/= <location2>
```

Pattern 3 is used to describe how a robot goes from a location in a room to another. The pattern is structurally similar to the pattern displayed in 2 but displays a different behaviour, in that it specifies a movement within a room rather than a movement between rooms.

Pattern 4 *The fourth pattern displays the constraints on the task needed for the robot to grasp an object on a specified location.*

```
<agent> grasps <object> on <location> is
   [ <agent> locates <object> at <place>
     <agent> moves forceps to <coordinates>
     <agent> closes forceps
     <agent> retracts forceps ]
   where <location> is_one_of [table, work desk,
                                 chest of drawers]
   and <agent> is_at <location>
   and <coordinates> is <place>.getCoordinates(<object>)
   and graspable(<object>) is true
```

Pattern 4 displays the necessary activities used to pick up an object from a place. In this example, we display the use of functions to get low level information (`getCoordinates()`), in this case, to get the physical coordinates of the object to lift . Three of the sub patterns for this pattern involve the physical behaviours of moving the forceps for the manipulation of the objects. The last line of this pattern uses a check for a certain property of the object to be grasped.

The TAPAS examples displayed in this paper are purposely small and relatively simple. For a real application the pattern systems are larger and more complex. As a result the hierarchical structures may become relatively deep and difficult to overlook in a manual process. Furthermore, patterns can also express meta activities, such as asking the user to define objects which have not yet been defined. This is possible by letting also task patterns be used to fill in as task fragments. Thus the task pattern analysis method can be used to describe relatively complex tasks for the Intelligent Service Robot. The backside of this is that the description process is difficult to handle without computerized tools.

4   Task Pattern Sets

Task patterns are gathered into task pattern sets collecting those patterns that are applicable for a robot in a certain situation. The task pattern sets are dynamically changing as the task is performed. Initially, a task pattern set may contain only default actions, such as a pattern for what to do, if a time out occurs during a task, or a pattern that tells the robot to wait for instructions. Such patterns can be kept in the set continuously, also during the performance of a certain task.

This is exemplified in figure 1, where the agent has 6 different patterns to choose from. When pattern TP3 is chosen and executed, the context changes, so that task patterns TP1, TP2, TP3 and TP5 are not available any more, for example due to context-based constraints, whereas the new task patterns TP8 and TP9 have been advanced for the new context. Task pattern TP6 is still available, and is not affected by the execution of TP3. In this case TP6 may be a default pattern, with a low priority, which can be used to get help in case the robot gets stuck during the execution of a task. Such a Task Pattern could, e.g., describe the activity needed to ask the user for new instructions. It could also be the "time out" pattern that would allow the robot to return to the user in case it cannot proceed with the task within a specified period of time.

When a first instruction is given to the robot, e.g., "Fetch book from table in study", the instruction is translated into an first pattern (such as pattern 1, but with some of the fragments instantiated), which is added to the current task pattern set. Each pattern in the set is assigned a priority. The default patterns, which are supposed to be used as emergency actions or idle activities, are given low priorities. This means that they will take precedence only in case the main task pattern or any of its sub patterns cannot be concluded. When the task is executed by the robot, the pattern is instantiated whenever possible, and then replaced by its component sub patterns. The sub patterns are evaluated in their turn, using the constraints as explained above to guide the order of evaluation.
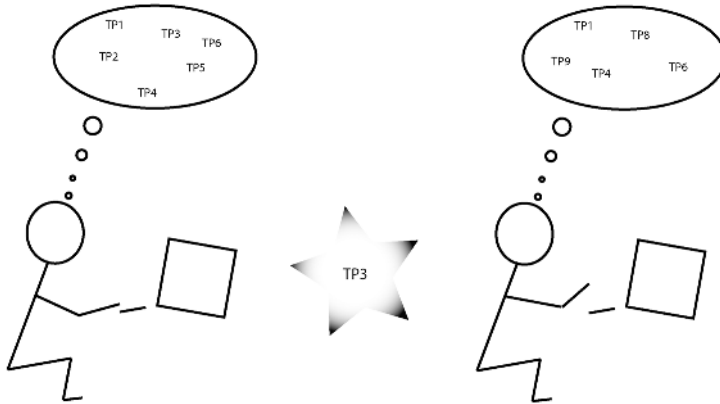
Fig. 1. Task patterns are collected into dynamic sets, which change as the task progresses. In this figure the task pattern set changes as task pattern TP3 is executed. (figure by the author)

Models according to task pattern analysis as defined in this paper have been developed by hand for a few sample applications. However, it turns out that the automation of this process is a necessity, since task pattern analysis to a large extent depends on the management of a large number of possible patterns. A prototype editor for task patterns has been developed as a Masters Thesis project [13].

5   Concluding Discussion

Task pattern analysis (TAPAS) is a conceptually simple but still powerful method for the description of grounded tasks. By defining tasks in terms of patterns, the large variety of small task structures describing similar tasks for the robot can be accounted for. The method is aimed at the description of tasks that are well founded, such as task descriptions for intelligent service robots where the robot basic behaviours constitute the lowest level of description. This property also makes the method suitable for describing tasks involving command based software, e.g., as a macro description language. However, task pattern analysis may be less useful for describing non-computerised tasks (such as traditional task analysis methods), since these are not necessarily be well-founded.

In terms of human-robot interaction, task pattern analysis may be interesting since it allows for the description of meta tasks, such as asking the user to define an unknown object. The possibility to handle meta rules in the TAPAS method has shown to be an advantage when it has been compared to most previous methods used for task analysis.

References

1. Honda: Asimo home page. http://asimo.honda.com/ (2005)
2. Electrolux: Trilobite vacuum cleaner web-site. http://trilobite.electrolux.com/ (2005)
3. McCarthy, J., Hayes, P.: Some philosophical problems from the standpoint of artificial intelligence. In Meltzer, B., Michie, D., eds.: Machine Intelligence. Volume 4. Edinburgh University Press, Edinburgh, Scotland (1969) 463–502
4. Green, A., Severinson Eklundh, K.: Task-oriented dialogue for cero: a user-centered approach. In: Proceedings of Ro-Man'01 (10th IEEE International Workshop on Robot and Human Communication), Bordeaux-Paris (2001) 146–151
5. Green, A., Hüttenrauch, H., Norman, M., Oestreicher, L., Severinson Eklundh, K.: User-centered design of intelligent service robots. In Kamejima, K., ed.: ROMAN, Osaka, Japan (2000)
6. Hüttenrauch, H., Green, A., Severinson Eklundh, K., Oestreicher, L., Norman, M.: Involving users in the design of a mobile office robot. IEEE Transactions on Systems, Man and Cybernetics **C** (2003)
7. Annett, J., Duncan, K.: Task analysis and training design. Occupational Psychology **41** (1967) 211 – 221
8. Tauber, M.: Top-down design of human computer interfaces. In Chang, S.K., Ichikawa, T., Legomenides, P., eds.: Visual Languages. Plenum Press, New York, New York (1986) 393 – 430
9. Tauber, M.: On mental models and the user interface. In Green, T., van der Veer, G., Hoc, J.M., Murray, D., eds.: Working with Computers Theory versus Outcome. Computer and People Series. Academic Press, London, England (1988)
10. Oestreicher, L.: Executable Formal Descriptions for Software Design: Logic and Logic Programming as a design Tool. Ph.L. thesis, Uppsala University (1991)
11. Oestreicher, L.: Logic descriptions in rapid prototyping of applications. In Diaper, D., Gilmore, D., Cockton, G., Shackel, B., eds.: Human-Computer Interaction, Cambridge, U.K., North-Holland (1990)
12. Shepherd, A.: Analysis and training in information technology tasks. In Diaper, D., ed.: Task Analysis for Human-Computer Interaction. Ellis Horwood, Chichester, England (1989) 15 – 54
13. Johansson, S., Mottaghi, E.: TAPAS - A Support System for Task Pattern Analysis. M.Sc. thesis, Uppsala University (2005)

Temporal Knowledge Representation
and Reasoning Model for Temporally Rich Domains

Slobodan Ribarić

Faculty of Electrical Engineering and Computing, University of Zagreb
Unska 3, 10000 Zagreb, Croatia
`slobodan.ribaric@fer.hr`

Abstract. A formal model of knowledge representation and reasoning for - temporally rich domains is given in the paper. The model is based on the high-level Petri Nets, called the Petri Nets with Time Tokens (PNTT). The PNTT is the main building block of the knowledge representation scheme KRPNTT. The scheme integrates theory of the PNTT and Allen's interval-based temporal logic. Reasoning process in the proposed scheme is *a time dependent event-driven process*. It includes the ability to reason about time relations among actions and/or states, as well as planning in the temporally rich domains.

1 Introduction

In the last two decades, the problem of knowledge representation and reasoning has been an important issue in the fields of artificial intelligence (natural language under-standing, planning, data mining, casual explanation), and the fields of time-critical systems (robotic assembly systems, flexible manufacturing systems, traffic control systems). Common characteristics for all the above applications are that they include concurrent actions which take time, the simultaneous occurrence of many actions at once, as well as external events caused by natural forces. Domains having these fea-tures are *temporally rich domains* [1]. Different formal models are proposed as at-tempts of an effective solution of the problem crisp or fuzzy temporal knowledge representation and reasoning: situation calculus [2], a time specialist [3], time maps [4], a formal logic of plans [1], the interval-based temporal logic [5], Dechter's metric constrains [6], timed, time and temporal Petri Nets [7], [8], [9] fuzzy time Petri Nets [10], and a point-interval logic [11].

In the paper, a formal model of knowledge representation and reasoning for tem-porally rich domains is given. The model is based on the high-level Petri Nets, called the Petri Nets with Time Tokens (PNTT). The PNTT uses p-timed Petri Net model and a new concept of time tokens.

2 Petri Nets with Time Tokens

The Petri Net with Time Tokens is a 9-tuple: PNTT = (P, T, I, O, τ, M, υ, δ, Ω), where P, T, I and O are components of a generalized Petri Net (PN) defined as fol-lows [12]: P is a finite set of places P = {p_1, p_2, … , p_n}, n $\geq$ 0; T is a finite set of transitions T = {t_1, t_2, …, t_m}, m $\geq$ 0; I is an input function I: T $\rightarrow$ P$^\infty$, a mapping from

transitions into bags of places; O is an output function $O: T \rightarrow P^\infty$, a mapping from transitions into bags of places, $P \cap T = \varnothing$.

A function $\tau: P \rightarrow (Q^* \cup \infty)$ is mapping from a set of places P into a set of time delays. A $\tau(p_i) \in (Q^* \cup \infty)$ corresponds to the time duration of some action or state (Q^* is the set of nonnegative rational numbers). The value $\tau(p_i)$ may be considered as time delay of tokens at the place p_i. All $\tau(p_i)$, $i = 1, 2, ..., n$, where n is cardinality of a set P, form a set of time delays $TD = \{\tau(p_1), \tau(p_2), ... , \tau(p_n)\}$.

A set $M = \{m_1^0, m_1^1, ... , m_1^j, m_2^0, m_2^1, ..., m_2^s, m_3^0, m_3^1 ..., m_3^l, ..., m_r^0, ... , m_r^p\}$, $1 \leq r \leq \infty$, is a set of time tokens. A time token m_i^{k+1} is a successor of a time token m_i^k, $k = 0, 1, 2, ... $.

The function $\upsilon: M \rightarrow \prod_{1 \leq i < \infty} (P \times TD)_i$ where $\prod$ denotes Cartesian product, P is a set of places, TD is a set of time delays, and $\times$ denotes Cartesian product, is mapping called *a time track*. A $\upsilon(m_i^k)$, $m_i^k \in M$, determines the sequence of places and time delays, which is assigned to a time token m_i^k as a consequence of the execution of the PNTT: It determines the time track for m_i^k which is formed on the basis of visiting places and time delays of all time tokens – ancestors of m_i^k.

A function $\delta: M \rightarrow (Q^* \cup \infty)$, is a mapping called *a time accumulation.* It determines the current amount (accumulated time) of detaining time for the time token $m_i^k \in M$. An initial time accumulation δ_0 can be assigned to a time token that is initially distributed in the PNTT. It is interpreted as an initial extension of the time interval of the corresponding place. By means of functions υ and δ, the structure of a time token is defined as n-tuple: $(\delta_0, (p_j, \tau(p_j)), (p_k, \tau(p_k)), ... , (p_l, \tau(p_l)), \tau_c)$, where τ_c is a current accumulated time. This information (a time track and a current accumulated time) will be used for temporal reasoning.

An injective function $\Omega: P \rightarrow \wp(M)$, with the restriction that $\Omega(p_i) \cap \Omega(p_j) = \varnothing$, for $i \neq j$, is called marking of the PNTT. A $\wp(M)$ denotes the power set of M. By Ω_0 we denote initial marking, that is, the initial distribution of time tokens in places of the PNTT.

Generally, tokens give to marked Petri Nets (PN) dynamical properties and they are used to define *the execution* of the PN [12]. In the PNTT, a transition t_i is enabled if each of its input places has at least as many tokens in it as arcs from the place to the transition *and* if time of detainment of these tokens in places has elapsed. Such time tokens are *movable time tokens.* Firing of an enabled transition in the PNTT is performed *automatically* and *immediately* after the transition is enabled. The number of tokens at input and output places of the fired transition is changed in accordance with the basic definition for the original PN [12]. The firing of the enabled transition in the PNTT removes time tokens – ancestors from its input places and simultaneously generates time tokens – successors in its output places.

The PNTT can be represented by a bipartite directed multigraph. The *circles* represent places while *bars* represent transitions. Directed arcs connected the places and the transitions are defined by means of an input function I, while arcs directed from transitions to places are defined by an output function O. Multiple input places and multiple output places are represented by multiple arcs. The time tokens are represented by dots (•) in the places.

3   Temporal Knowledge Scheme

The temporal knowledge scheme KRPTT is defined as follows:

KRPTT = (PNTT, TLM, α, β, F), where the PNTT is a Petri Net with Time Tokens, while TLM is a temporal logic module based on Allen's time-interval logic. A function α: P $\rightarrow$ (D $\cup$ C$_p$) is a bijective function from a set of places P to a union of a set of actions and /or states D and set of control states C$_p$. A function β: T$\rightarrow$ ($\Sigma \cup$ C$_t$) is mapping from a set of transitions T to a union of a set of events Σ and control events C$_t$. Both mappings α and β give semantic meaning to a model. A set F is a set of flags. In general, a flag $f_i \in$ F has structure (p_i, p_j, tr, p_l, p_k, ..., p_c), where p_i and p_j denote elements from P as places potentially with time tokens that have to be tested by the TML, according to the temporal relation tr. The evaluation of the tr is based on information that is carried by the time tokens from the places p_i and p_j. The temporal relation tr is from a set of thirteen possible Allen's time-interval relations (Table 1) defined in [5]. In flags, tr may be compound by using the logical connectives $\vee$ and $\wedge$. The p_k, ..., p_c, where $\alpha(p_k) \in$ C$_p$, ...$\alpha(p_c) \in$ C$_p$, in the flag f_i specify the places in which the TLM sets tokens depending on the result of evaluation of tr. These tokens, called *control tokens*, are treated as time tokens without time history. A degenerative type of a flag(s) f_{Gi} = (p_h, - , - , -) is(are) used to denote goal state(s) of the system.

Table 1. Allen's temporal relations [5]

Relation	Symbol	Symbol for inverse
X before Y	<	>
X equal Y	=	=
X meets Y	m	mi
X overlaps Y	o	oi
X during Y	d	di
X starts Y	s	si
X finishes Y	f	fi

A temporal logic module TLM is a component of KRPTT scheme, which is capable to inference about temporal relations between time intervals. The evaluation of tr is activated when time tokens are available and movable at the places specified by the flag. At that moment copies of the time tokens are transferred to the TLM. The TLM is data driven and the transfer activates evaluation of the temporal relation. Immediately after evaluation, if the relation tr is satisfied, the control tokens are transferred to places p_k, ..., p_c designated as control states.

3.1   Reasoning

The reasoning process can be described as follows: The input is the initial marking of the KRPTT representation of the world. It determines current activities (states) and temporal relationships among the activities (states). As time proceeds, the enabled transitions are automatically fired and time tokens are distributed through the graph of the KRPTT. Depending on their path through the graph and time duration of the activities (states), each time token carries a history of the execution of the graph. The firing sequences are additionally controlled by the TLM in such a way that it tests and

evaluates temporal relations between activities (states), which correspond to the places specified by flags. According to the results of the temporal relationship evaluation, the control tokens are put at the places specified by the flags. These tokens have an influence on the result of the firing sequence in the graph by means of enabling transitions t_i, i =1,2,…. The combination of the time tokens associated with activities (states) and control tokens, both present at the same time can be interpreted as time-dependent *if-then* rule implementation: If there are enough time and control tokens at corresponding input places the transition is fired, i.e., the time dependent rule is activated and action (or conclusion) is generated. It is obviously that the above process is driven by time-dependent events. If some places in the graph are denoted as goal states, the scheme can conclude if these time dependent goals may be achieved and detect sequence(s) of activities (states) which leads to these goals. By varying the initial marking of the scheme it can be used for planning in temporally rich domains. The result of all the sequences of firing of enabled transitions, starting with the initial marking is represented by a reachability tree [12] having additional time information or by a firing diagram (Fig. 3).

4   Example

In this Section we give an example of using the knowledge representation scheme KRPTT for modelling and planning activities in a dynamic laboratory scene. Fig. 1 shows initial positions of three agents (*robot 1* and *robot 2* are equipped with sonars, and *robot arm*). Two movable robots share a common goal: one of them has to reach the charger. This goal has to be achieved in the shortest possible time. It seems that the *robot 1* is nearer to the charger, but due to an obstacle (*obstacle 1*; Fig. 1), its path is longer. The *robot 2* is the candidate for achieving the goal in the shortest possible time, but there is *a robot arm,* which can drop a brick on the path of the *robot 2*, making its mission impossible.
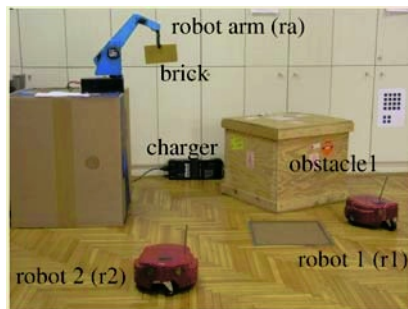
Fig. 1. A dynamic laboratory scene

The knowledge about the scene is modelled by using the KRPTT scheme as follows (Fig. 2): P = {p_1, p_2, ... , p_{20}}, T = {t_1, t_2, ... , t_{11}}, I: T $\rightarrow$ P$^\infty$, O: T $\rightarrow$ P$^\infty$, (for our example a bag P$^\infty$ degenerates into a set P).The input and output functions can be concluded from Fig. 2. The time delays $\tau(p_i) = \tau_i$, as well as, functions α and β are denoted at each place and transition, respectively. The initial marking Ω_0, initial time accumulations of time tokens, and distribution of flags are depicted in Fig. 2.

r1 is approaching charger; $\tau_1 = 0.1$

p_1
$m_1^0; \delta_{0m1} = 0.0$

r1 detects obstacle1

r1 goes around obstacle1; $\tau_2 = 0.125$ p_2

$p_6; \tau_6 = 0$ t_2

ctrl. state

r1 finishes going
$p_3; \tau_3 = 0.2$ around
ctrl. state

ctrl. event

t_3 r1 reaches charger

t_4 $p_5; \tau_5 = 0$

$p_7; \tau_7 = \infty$ f_{G1}

r1 has stopped   r1 has reached charger

$p_4; \tau_4 = \infty$

$\alpha(p_3) = $"r1 continues approaching charger"

$p_8; \tau_8 = 0.2$ r2 is moving by
$m_2^0; \delta_{0m2} = 0.2$ const velocity
r2 is entering into
$p_{10}; \tau_{10} = 0$ t_5 coll. zone
p_9 f_1-f_3

ctrl. state   r2 goes out of coll.
t_6 zone
$p_{11}; \tau_{11} = 0.1$ r2 is out of coll.
zone

t_7 r2 reaches charger
f_{G2}

$p_{12}; \tau_{12} = \infty$ r2 has reached
charger

$\alpha(p_9) = $"r2 is in collision zone"; $\tau_9 = 0.1$

$\delta_{0m3} = 0.45$
m_3^0 p_{13} ra holds brick; $\tau_{13} = 0.0$

$p_{15}; \tau_{15} = 0.0$ t_8 ra releases brick
ctrl. state p_{14} f_1-f_3
brick is in air; $\tau_{14} = 0.05$

t_9 t_{10} brick is falling t_{11}
ctrl. event $\tau_{16} = 0.0$ ctrl. event

p_{18} p_{16} p_{19} p_{20} p_{17}
ctrl. state
r2 chrases into   Brick is on floor   r2 has stopped in front
brick; $\tau_{18} = \infty$ $\tau_{19} = \infty$ brick; $\tau_{20} = \infty$

Fig. 2. The knowledge about the scene modelled by using the KRPTT

A set of flags is: $F = \{f_1, f_2, f_3, f_{G1}, f_{G2}\}$, where $f_1 = (p_9, p_{14}, <, p_6, p_{10})$, where "<" denotes temporal relation "before" (Table 1), $f_2 = (p_9, p_{14}, >, p_5, p_7)$, $f_3 = (p_9, p_{14}, tr, p_5, p_{17})$, where *tr* stands for mi ∨ m ∨ oi ∨ i ∨ = ∨ di ∨ d ∨ si ∨ s ∨ fi ∨ f (Table 1). The goal states of the system are denoted by f_{G1} and f_{G2}.

The knowledge base designed by the KRPTT scheme can generate answers for different inquiries from it. Let us suppose that we put to the knowledge base the following inquiry: "What will occur if the *robot 1* starts its moving at the time $t_0 = 0$ (the action starts at the beginning of the world), and the *robot 2* starts its moving after 0.2 time units, and after 0.45 time units the *robot arm* releases the brick?"

The answer generation, i.e. the reasoning process, starts at $t_0 = 0$ with the initial marking and the initial extensions of time intervals. The reasoning process is based on the execution of the KRPTT, which is driven by means of the TLM in such a way that it tests and evaluates temporal relations between the activities, which are specified by flags f_1, f_2 and f_3. The control tokens generated by the TLM control the execution of the KRPTT and determine sequences of activities. Fig. 3. shows a firing dia-

gram generated for our example. From the firing diagram it is obvious that the modelled system has reached two states assigned to places p_4 and p_{18}. These states correspond to the following states: "*r 1 has reached charger*" and "*r 2 crashes into brick*", respectively. The structure of the time token $m_1{}^3$ is (0.0, (p_1, 0.1), (p_2, 0.125), (p_3, 0.2), (p_4, ∞), τ_c - ∞ = 0.425), which means that the *robot 1* has reached the charger at t = 0.425.
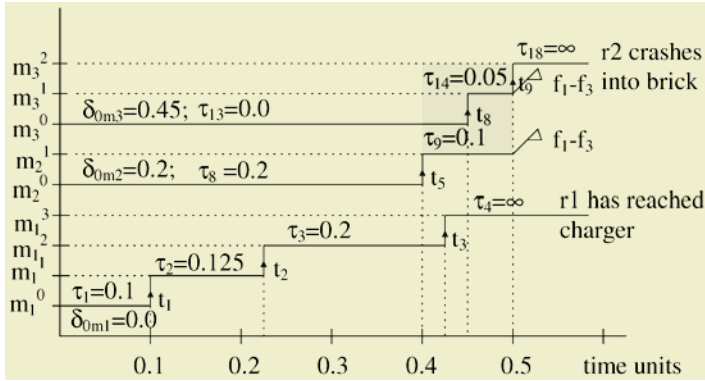
Fig. 3. The firing diagram

5   Conclusion

The proposed temporal knowledge representation scheme KRPTT is used to design the knowledge bases of dynamical systems with more concurrent and simultaneous activities, as well as, external events. The model effectively integrates temporal information and metric, information of the system's states, interval-based temporal logic, action specification and ability of planning. The major advantages of the scheme are unified representation of different temporal information, well-defined methods for the analysis of the model, and simple modelling of different time relations among actions by changing the initial marking and time values assigned to time tokens. The model also offers graphical representation and gives a straightforward view of relations among the objects, actions and states.

The future research work is directed to object-oriented implementation of a simulator of the models based on the KRPTT scheme.

References

1. Pelevan, R., Allen, J. F.: A Formal Logic Plans in Temporally Rich Domains, Proceedings of the IEEE, Vol. 74. No. 10. (1986) 1364-1382
2. McCharty, J., Hayes, P. J.: Some Philosophical Problems from the Standpoint of Artificial Intelligence, in Meltzer, B., Michie, D., (eds.), Machine Intelligence 4, Edinburgh University Press, Edinburgh (1969)
3. Kahn, K., Gorry, G. A.: Mechanizing Temporal Knowledge, Artificial   Intelligence 9 (1977) 87-108.
4. Dean, T. L., McDermott, D., V.: Temporal Data Base Management, Artificial Intelligence 32 (1987) 1-55

5. Allen, J., F.: Maintaining Knowledge about Temporal Intervals, Communications of the ACM, Vo. 26, No. 11 (1983) 832-843

6. Decther R., Meiri, I., Pearl J.: Temporal Constraint Network, Artificial Intelligence 49, (1991) 61-69

7. Merlin, P.: A Methodology for the Design and Implementation of Communication Protocols, IEEE Transactions on the Communications, Vol. 24, No. 6, (1976), 614-621

8. Suzuki, I., Lu, H.: Temporal Petri Nets and Their Application to Modeling and Analysis of a Handshake Daisy Chain Arbiter, IEEE Transactions on Computers, Vol. 38, No. 5, (1989) 696-704

9. Meimei, G., Xiaoguang, H., Zhiming, W.: Linear logic as a tool for presentation and temporal reasoning of time Petri nets, Proceedings of the American Control Conference, Chicago, (2000), 3177-3181

10. Murata, T., Suzuki, T., Shatz, S. M.: Fuzzy-Timing High-Level Petri Nets (FTHNs) for Time-Critical Systems, in Cardoso, J., Camargo, H. (eds.), Fuzziness in Petri Nets, A Springer-Verlag, Heidelberg (1999)

11. Zaidi, A. K., Rizvi, K. H., Hussain, S. S.: On Spatial Modeling of Discrete Event Systems Using Point-interval Logic, Proceedings of the Second Int. Conf. on Machine Learning and Cybernetics, Xian,(2003) 1699-1704..

12. Peterson, J. L.: Petri Net theory and Modeling of Systems, Prentice-Hall, Englewood Cliffs, (1981)

Reciprocal Logic: Logics for Specifying, Verifying, and Reasoning About Reciprocal Relationships

Jingde Cheng

Department of Information and Computer Sciences, Saitama University
Saitama, 338-8570, Japan
cheng@aise.ics.saitama-u.ac.jp

Abstract. To specify, verify, and reason about various reciprocal relationships in a human society and/or a cyber space, we need a right fundamental logic system to provide us with a criterion of logical validity of reasoning as well as a representation and specification language. This paper proposes a new family of conservative extensions of relevant logic, named "reciprocal logic," for specifying, verifying, and reasoning about reciprocal relationships. The paper shows that various reciprocal logics can be obtained by introducing predicates and related axioms about reciprocal relationships into strong relevant logics and spatial-temporal relevant logics. A case study is focused on trust relationships.

1 Introduction

In human society and/or a cyber space, there are many reciprocal relationships that must concern two parties, such as parent-child relationship, relative relationship, friendship, adjacent relationship, high and low relationship, cooperative relationship, complementary relationship, adverse relationship, dependent relationship, trust relationship, trade relationship, buying and selling relationship, and so on. As a result, many reciprocal relationships appear in various disciplines including Artificial Intelligence, Cryptography, Economics, Game Theory, Information Security Engineering, Knowledge Engineering, Linguistics, Logic, Multi-agent Systems, Philosophy, Psychology, and Software Engineering.

To specify, verify, and reason about various reciprocal relationships in a human society and/or a cyber space, we need a right fundamental logic system to provide us with a criterion of logical validity of reasoning as well as a representation and specification language. The question, "Which is the right logic?" invites the immediate counter-question "Right for what?" Only if we certainly know what we need, we can make a good choice. It is obvious that different applications may require different characteristics of logic.

The present author considers that we should consider the following essential requirements for the fundamental logic. First, as a general logical criterion for the validity of reasoning as well as proving, the logic must be able to underlie relevant reasoning as well as truth-preserving reasoning in the sense of conditional, i.e., for any reasoning based on the logic to be valid, if its premises are true in the sense of conditional, then its conclusion must be relevant to the premises and true in the sense of

conditional. Second, the logic must be able to underlie ampliative reasoning in the sense that the truth of conclusion of the reasoning should be recognized after the completion of the reasoning process but not be invoked in deciding the truth of premises of the reasoning. From the viewpoint to regard reasoning as the process of drawing new conclusions from given premises, any meaningful reasoning must be ampliative but not circular and/or tautological. Third, the logic must be able to underlie paracomplete reasoning and paraconsistent reasoning. In particular, the so-called principle of Explosion that everything follows from a contradiction cannot be accepted by the logic as a valid principle. In general, our knowledge about various reciprocal relationships may be incomplete or even inconsistent in many ways, i.e., it gives us no evidence for deciding the truth of either a proposition or its negation, or even it directly or indirectly includes some contradictions. Therefore, reasoning with incomplete and/or inconsistent knowledge is the rule rather than the exception in our everyday lives and almost all scientific disciplines. Finally, because reciprocal relationships themselves may change over space and time, the right fundamental logic system must be able to underlie spatial reasoning, or temporal reasoning, or both.

Classical mathematical logic (**CML** for short) cannot satisfy any of the above essential requirements because of the following facts: a reasoning based on **CML** is not necessarily relevant; the classical truth-preserving property of a reasoning based on **CML** is meaningless in the sense of conditional; a reasoning based on **CML** must be circular and/or tautological but not ampliative; reasoning under inconsistency is impossible within the framework of **CML** [1, 2, 4, 5]. The above facts are also true to those classical conservative extensions or non-classical alternatives of **CML** including temporal (classical) logics [3, 11, 12] and spatial (classical) logics [8-10] where the classical account of validity is adopted as the logical validity criterion and the notion of conditional is directly or indirectly represented by the material implication. **CML** does not underlie spatial reasoning and temporal reasoning explicitly.

Traditional relevant (or relevance) logics ware constructed during the 1950s in order to find a mathematically satisfactory way of grasping the elusive notion of relevance of antecedent to consequent in conditionals, and to obtain a notion of implication which is free from the so-called 'paradoxes' of material and strict implication [1, 2]. Some major traditional relevant logic systems are 'system **E** of entailment', 'system **R** of relevant implication', and 'system **T** of ticket entailment'. A major characteristic of the relevant logics is that they have a primitive intensional connective to represent the notion of (relevant) conditional and their logical theorems include no implicational paradoxes. The underlying principle of the relevant logics is the relevance principle, i.e., for any entailment provable in **E**, **R**, or **T**, its antecedent and consequent must share a propositional variable. Variable-sharing is a formal notion designed to reflect the idea that there be a meaning-connection between the antecedent and consequent of an entailment. It is this relevance principle that excludes those implicational paradoxes from logical axioms or theorems of relevant logics. Also, since the notion of entailment is represented in the relevant logics by a primitive intensional connective but not an extensional truth-function, a reasoning based on the relevant logics is ampliative but not circular and/or tautological. Moreover, because the relevant logics reject the principle of Explosion, they can certainly underlie paraconsistent reasoning. However, traditional relevant logics still include conjunction-implicational paradoxes and disjunction-implicational paradoxes [4, 5]. As a result,

they only can guarantee the relevance between the premises of a valid argument and its conclusion and the validity of its conclusion in a sense of weak relevance. Relevant logics do not underlie spatial reasoning and temporal reasoning explicitly.

Thus, no existing logic can satisfy all essential requirements for the fundamental logic. This paper proposes a new family of conservative extensions of relevant logic, named "***reciprocal logic***," for specifying, verifying, and reasoning about reciprocal relationships. The paper shows that various reciprocal logics can be obtained by introducing predicates and related axioms about reciprocal relationships into strong relevant logics and spatial-temporal relevant logics. A case study is focused on trust relationships.

2 Strong Relevant Logics and Spatio-temporal Relevant Logics

In order to establish a satisfactory logic calculus of conditional to underlie relevant reasoning, the present author has proposed ***strong relevant*** (or ***relevance***) ***logics*** [4]. The logics require that the premises of an argument represented by a conditional include no unnecessary and needless conjuncts and the conclusion of that argument includes no unnecessary and needless disjuncts. As a modification of traditional relevant logics, strong relevant logics reject all conjunction-implicational paradoxes and disjunction-implicational paradoxes in traditional relevant logics. What underlies the strong relevant logics is the strong relevance principle: If A is a theorem of strong relevant logics, then every propositional variable in A occurs at least once as an antecedent part and at least once as a consequent part. In the framework of strong relevant logics, if a reasoning and/or argument is valid, then both the relevance between its premises and its conclusion and the validity of its conclusion in the sense of conditional can be guaranteed in a certain sense of strong relevance.

The logical connectives, axiom schemata, and inference rules of strong relevant logics are as follows:

Primitive Logical Connectives: { $\Rightarrow$ (entailment), $\neg$ (negation), $\wedge$ (extensional conjunction) }

Defined Logical Connectives: { $\otimes$ (intensional conjunction, $A \otimes B =_{df} \neg(A \Rightarrow \neg B)$), $\oplus$ (intensional disjunction, $A \oplus B =_{df} \neg A \Rightarrow B$), $\Leftrightarrow$ (intensional equivalence, $A \Leftrightarrow B =_{df} (A \Rightarrow B) \otimes (B \Rightarrow A)$), $\vee$ (extensional disjunction, $A \vee B =_{df} \neg(\neg A \wedge \neg B)$), $\rightarrow$ (material implication, $A \rightarrow B =_{df} \neg(A \wedge \neg B)$ or $A \rightarrow B =_{df} \neg A \vee B$), $\leftrightarrow$ (extensional equivalence, $A \leftrightarrow B =_{df} (A \rightarrow B) \wedge (B \rightarrow A)$) }

Axiom Schemata: E1: $A \Rightarrow A$, E2: $(A \Rightarrow B) \Rightarrow ((C \Rightarrow A) \Rightarrow (C \Rightarrow B))$, E2′: $(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C))$, E3: $(A \Rightarrow (A \Rightarrow B)) \Rightarrow (A \Rightarrow B)$, E3′: $(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$, E3″: $(A \Rightarrow B) \Rightarrow ((A \Rightarrow (B \Rightarrow C)) \Rightarrow (A \Rightarrow C))$, E4: $(A \Rightarrow ((B \Rightarrow C) \Rightarrow D)) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow D))$, E4′: $(A \Rightarrow B) \Rightarrow (((A \Rightarrow B) \Rightarrow C) \Rightarrow C)$, E4″: $((A \Rightarrow A) \Rightarrow B) \Rightarrow B$, E4‴: $(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (((A \Rightarrow C) \Rightarrow D) \Rightarrow D))$, E5: $(A \Rightarrow (B \Rightarrow C)) \Rightarrow (B \Rightarrow (A \Rightarrow C))$, E5′: $A \Rightarrow ((A \Rightarrow B) \Rightarrow B)$, N1: $(A \Rightarrow (\neg A)) \Rightarrow (\neg A)$, N2: $(A \Rightarrow (\neg B)) \Rightarrow (B \Rightarrow (\neg A))$, N3: $(\neg(\neg A)) \Rightarrow A$, C1: $(A \wedge B) \Rightarrow A$, C2: $(A \wedge B) \Rightarrow B$, C3: $((A \Rightarrow B) \wedge (A \Rightarrow C)) \Rightarrow (A \Rightarrow (B \wedge C))$, C4: $(\boldsymbol{L}A \wedge \boldsymbol{L}B) \Rightarrow \boldsymbol{L}(A \wedge B)$, where $\boldsymbol{L}A =_{df} (A \Rightarrow A) \Rightarrow A$, D1: $A \Rightarrow (A \vee B)$, D2: $B \Rightarrow (A \vee B)$,

D3: $((A\Rightarrow C)\wedge(B\Rightarrow C))\Rightarrow((A\vee B)\Rightarrow C)$, DCD: $(A\wedge(B\vee C))\Rightarrow((A\wedge B)\vee C)$, C5: $(A\wedge A)\Rightarrow A$, C6: $(A\wedge B)\Rightarrow(B\wedge A)$, C7: $((A\Rightarrow B)\wedge(B\Rightarrow C))\Rightarrow(A\Rightarrow C)$, C8: $(A\wedge(A\Rightarrow B))\Rightarrow B$, C9: $\neg(A\wedge\neg A)$, C10: $A\Rightarrow(B\Rightarrow(A\wedge B))$, IQ1: $\forall x(A\Rightarrow B)\Rightarrow(\forall xA\Rightarrow\forall xB)$, IQ2: $(\forall xA\wedge\forall xB)$ $\Rightarrow\forall x(A\wedge B)$, IQ3: $\forall xA\Rightarrow A[t/x]$ (if x may appear free in A and t is free for x in A, i.e., free variables of t do not occur bound in A), IQ4: $\forall x(A\Rightarrow B)\Rightarrow(A\Rightarrow\forall xB)$ (if x does not occur free in A), IQ5: $\forall x_1 \ldots \forall x_n(((A\Rightarrow A)\Rightarrow B)\Rightarrow B)$ $(n\geq 0)$

Inference Rules: $\Rightarrow$E: from A and $A\Rightarrow B$ to infer B (Modus Ponens), $\wedge$I: from A and B to infer $A\wedge B$ (Adjunction), $\forall$I: if A is an axiom, so is $\forall xA$ (Generalization of axioms)

Various relevant logic systems are defined as follows, where we use 'A | B' to denote any choice of one from two axiom schemata A and B: $\mathbf{T}_\Rightarrow =_{df}$ {E1, E2, E2′, E3 | E3″} + $\Rightarrow$E, $\mathbf{E}_\Rightarrow =_{df}$ {E1, E2 | E2′, E3 | E3′, E4 | E4′} + $\Rightarrow$E, $\mathbf{E}_\Rightarrow =_{df}$ {E2′, E3, E4″} + $\Rightarrow$E, $\mathbf{E}_\Rightarrow =_{df}$ {E1, E3, E4‴} + $\Rightarrow$E, $\mathbf{R}_\Rightarrow =_{df}$ {E1, E2 | E2′, E3 | E3′, E5 | E5′} + $\Rightarrow$E, $\mathbf{T}_{\Rightarrow,\neg} =_{df} \mathbf{T}_\Rightarrow$ + {N1, N2, N3}, $\mathbf{E}_{\Rightarrow,\neg} =_{df} \mathbf{E}_\Rightarrow$ + {N1, N2, N3}, $\mathbf{R}_{\Rightarrow,\neg} =_{df} \mathbf{R}_\Rightarrow$ + {N2, N3}, $\mathbf{T} =_{df} \mathbf{T}_{\Rightarrow,\neg}$ + {C1~C3, D1~D3, DCD} + $\wedge$I, $\mathbf{E} =_{df} \mathbf{E}_{\Rightarrow,\neg}$ + {C1~C4, D1~D3, DCD} + $\wedge$I, $\mathbf{R} =_{df} \mathbf{R}_{\Rightarrow,\neg}$ + {C1~C3, D1~D3, DCD} + $\wedge$I, $\mathbf{Tc} =_{df} \mathbf{T}_{\Rightarrow,\neg}$ + {C3, C5~C10}, $\mathbf{Ec} =_{df} \mathbf{E}_{\Rightarrow,\neg}$ + {C3~C10}, $\mathbf{Rc} =_{df} \mathbf{R}_{\Rightarrow,\neg}$ + {C3, C5~C10}, $\mathbf{TQ} =_{df} \mathbf{T}$ + {IQ1~IQ5} + $\forall$I, $\mathbf{EQ} =_{df} \mathbf{E}$ + {IQ1~IQ5} + $\forall$I, $\mathbf{RQ} =_{df} \mathbf{R}$ + {IQ1~IQ5} + $\forall$I, $\mathbf{TcQ} =_{df} \mathbf{Tc}$ + {IQ1~IQ5} + $\forall$I, $\mathbf{EcQ} =_{df} \mathbf{Ec}$ + {IQ1~IQ5} + $\forall$I, $\mathbf{RcQ} =_{df} \mathbf{Rc}$ + {IQ1~IQ5} + $\forall$I. Here, $\mathbf{T}_\Rightarrow$, $\mathbf{E}_\Rightarrow$, and $\mathbf{R}_\Rightarrow$ are the purely implicational fragments of $\mathbf{T}$, $\mathbf{E}$, and $\mathbf{R}$, respectively, and the relationship between $\mathbf{E}_\Rightarrow$ and $\mathbf{R}_\Rightarrow$ is known as $\mathbf{R}_\Rightarrow = \mathbf{E}_\Rightarrow + A\Rightarrow LA$; $\mathbf{T}_{\Rightarrow,\neg}$, $\mathbf{E}_{\Rightarrow,\neg}$, and $\mathbf{R}_{\Rightarrow,\neg}$ are the implication-negation fragments of $\mathbf{T}$, $\mathbf{E}$, and $\mathbf{R}$, respectively; $\mathbf{Tc}$, $\mathbf{Ec}$, $\mathbf{Rc}$, $\mathbf{TcQ}$, $\mathbf{EcQ}$, and $\mathbf{RcQ}$ are strong relevant logics.

However, both traditional relevant logics and strong relevant logics do not underlie spatial reasoning and temporal reasoning explicitly. In order to specify, verify, and reason about spatio-temporal knowledge, we have proposed *spatio-temporal relevant logics* [7], which are obtained by introducing region connection predicates and axiom schemata of RCC [8-10], point position predicates and axiom schemata, and point adjacency predicates and axiom schemata into *temporal relevant logics* [6]. Below we present a modification of spatio-temporal relevant logics in [7].

Let $\{r_1, r_2, r_3, \ldots\}$ be a countably infinite set of individual variables, called *region variables*. Atomic formulas of the form $C(r_1, r_2)$ are read as 'region r_1 connects with region r_2.' Let $\{p_1, p_2, p_3, \ldots\}$ be a countably infinite set of individual variables, called *point variables*. Atomic formulas of the form $I(p_1, r_1)$ are read as 'point p_1 is included in region r_1.' Atomic formulas of the form $Id(p_1, p_2)$ are read as 'point p_1 is identical with p_2.' Atomic formulas of the form $Arc(p_1, p_2)$ are read as 'points p_1, p_2 are adjacent such that there is an arc from point p_1 to point p_2, or more simply, points p_1 is adjacent to point p_2.' Note that an arc has a direction. Atomic formulas of the form $Path(p_1, p_2)$ are read as 'there is a directed path from point p_1 to point p_2.' Here, $C(r_1, r_2)$, $I(p_1, r_1)$, $Id(p_1, p_2)$, $Arc(p_1, p_2)$, and $Path(p_1, p_2)$ are primitive binary predicates to represent relationships between geometric objects. Note that here we use a many-sorted language.

Temporal Operators: { G (future-tense always or henceforth operator, GA means 'it will always be the case in the future from now that A'), H (past-tense always opera-

tor, *HA* means 'it has always been the case in the past up to now that *A*'), *F* (future-tense sometime or eventually operator, *FA* means 'it will be the case at least once in the future from now that *A*'), *P* (past-tense sometime operator, *PA* means 'it has been the case at least once in the past up to now that *A*') } Note that these temporal operators are not independent and can be defined as follows: $GA =_{df} \neg F \neg A$, $HA =_{df} \neg P \neg A$, $FA =_{df} \neg G \neg A$, $PA =_{df} \neg H \neg A$.

Primitive Binary Predicate: { *C* (connection, $C(r_1, r_2)$ means 'r_1 connects with r_2'), *I* (inclusion, $I(p_1, r_1)$ means 'p_1 is included in r_1'), *Id* (the same point, $Id(p_1, p_2)$ means 'point p_1 is identical with p_2'), *Arc* (arc, $Arc(p_1, p_2)$ means 'p_1 is adjacent to p_2'), *Path* (path, $Path(p_1, p_2)$ means 'there is a directed path from p_1 to p_2') }

Defined Binary Predicates: $DC(r_1, r_2) =_{df} \neg C(r_1, r_2)$ ($DC(r_1, r_2)$ means 'r_1 is disconnected from r_2'), $Pa(r_1, r_2) =_{df} \forall r_3(C(r_3, r_1) \Rightarrow C(r_3, r_2))$ ($Pa(r_1, r_2)$ means 'r_1 is a part of r_2'), $PrPa(r_1, r_2) =_{df} Pa(r_1, r_2) \wedge (\neg Pa(r_2, r_1))$ ($PrPa(r_1, r_2)$ means 'r_1 is a proper part of r_2'), $EQ(r_1, r_2) =_{df} Pa(r_1, r_2) \wedge Pa(r_2, r_1)$ ($EQ(r_1, r_2)$ means 'r_1 is identical with r_2'), $O(r_1, r_2) =_{df} \exists r_3(Pa(r_3, r_1) \wedge Pa(r_3, r_2))$ ($O(r_1, r_2)$ means 'r_1 overlaps r_2'), $DR(r_1, r_2) =_{df} \neg O(r_1, r_2)$ ($DR(r_1, r_2)$ means 'r_1 is discrete from r_2'), $PaO(r_1, r_2) =_{df} O(r_1, r_2) \wedge (\neg Pa(r_1, r_2)) \wedge (\neg Pa(r_2, r_1))$ ($PaO(r_1, r_2)$ means 'r_1 partially overlaps r_2'), $EC(r_1, r_2) =_{df} C(r_1, r_2) \wedge (\neg O(r_1, r_2))$ ($EC(r_1, r_2)$ means 'r_1 is externally connected to r_2'), $TPrPa(r_1, r_2) =_{df} PrPa(r_1, r_2) \wedge \exists r_3(EC(r_3, r_1) \wedge EC(r_3, r_2))$ ($TPrPa(r_1, r_2)$ means 'r_1 is a tangential proper part of r_2'), $NTPrPa(r_1, r_2) =_{df} PrPa(r_1, r_2) \wedge (\neg \exists r_3(EC(r_3, r_1) \wedge EC(r_3, r_2)))$ ($NTPrPa(r_1, r_2)$ means 'r_1 is a nontangential proper part of r_2').

Axiom Schemata: T1: $G(A \Rightarrow B) \Rightarrow (GA \Rightarrow GB)$, T2: $H(A \Rightarrow B) \Rightarrow (HA \Rightarrow HB)$, T3: $A \Rightarrow G(PA)$, T4: $A \Rightarrow H(FA)$, T5: $GA \Rightarrow G(GA)$, T6: $(FA \wedge FB) \Rightarrow F(A \wedge FB) \vee F(A \wedge B) \vee F(FA \wedge B)$, T7: $(PA \wedge PB) \Rightarrow P(A \wedge PB) \vee P(A \wedge B) \vee P(PA \wedge B)$, T8: $GA \Rightarrow FA$, T9: $HA \Rightarrow PA$, T10: $FA \Rightarrow F(FA)$, T11: $(A \wedge HA) \Rightarrow F(HA)$, T12: $(A \wedge GA) \Rightarrow P(GA)$, RCC1: $\forall r_1 \forall r_2(C(r_1, r_2) \Rightarrow C(r_2, r_1))$, RCC2: $\forall r_1(C(r_1, r_1))$, PRCC1: $\forall p_1 \forall r_1 \forall r_2((I(p_1, r_1) \wedge DC(r_1, r_2)) \Rightarrow \neg I(p_1, r_2))$, PRCC2: $\forall p_1 \forall r_1 \forall r_2((I(p_1, r_1) \wedge Pa(r_1, r_2)) \Rightarrow I(p_1, r_2))$, PRCC3: $\forall r_1 \forall r_2(O(r_1, r_2) \Rightarrow \exists p_1(I(p_1, r_1) \wedge I(p_1, r_2)))$, PRCC4: $\forall r_1 \forall r_2(PaO(r_1, r_2) \Rightarrow \exists p_1(I(p_1, r_1) \wedge I(p_1, r_2)) \wedge \exists p_2(I(p_2, r_1) \wedge \neg I(p_2, r_2)) \wedge \exists p_3(\neg I(p_3, r_1) \wedge I(p_3, r_2)))$, PRCC5: $\forall r_1 \forall r_2(EC(r_1, r_2) \Rightarrow \exists p_1(I(p_1, r_1) \wedge I(p_1, r_2) \wedge \forall p_2(\neg Id(p_2, p_1) \Rightarrow \neg I(p_2, r_1) \wedge \neg I(p_2, r_2)))$, PRCC6: $\forall p_1 \forall r_1 \forall r_2((I(p_1, r_1) \wedge TPrPa(r_1, r_2)) \Rightarrow I(p_1, r_2))$, PRCC7: $\forall p_1 \forall r_1 \forall r_2((I(p_1, r_1) \wedge NTPrPa(r_1, r_2)) \Rightarrow I(p_1, r_2))$, APC1: $\forall p_1 \forall p_2(Arc(p_1, p_2) \Rightarrow Path(p_1, p_2))$, APC2: $\forall p_1 \forall p_2 \forall p_3((Path(p_1, p_2) \wedge Path(p_2, p_3)) \Rightarrow Path(p_1, p_3))$.

Inference Rules: TG: from *A* to infer *GA* and *HA* (Temporal Generalization)

The minimal or weakest propositional temporal relevant logics are defined as follows: $\mathbf{T_0Tc} =_{df} \mathbf{Tc} + \{T1 \sim T4\} + TG$, $\mathbf{T_0Ec} =_{df} \mathbf{Ec} + \{T1 \sim T4\} + TG$, $\mathbf{T_0Rc} =_{df} \mathbf{Rc} + \{T1 \sim T4\} + TG$. Note that the minimal or weakest temporal classical logic $\mathbf{K_t}$ = all axiom schemata for **CML** + $\rightarrow$E + $\{T1 \sim T4\}$ + TG. Other characteristic axioms such as T5~T12 that correspond to various assumptions about time can be added to $\mathbf{T_0Tc}$, $\mathbf{T_0Ec}$, and $\mathbf{T_0Rc}$ respectively to obtain various propositional temporal relevant logics. Various predicate temporal relevant logics then can be obtained by adding axiom schemata IQ1~IQ5 and inference rule $\forall$I into the propositional temporal relevant logics. For examples, minimal or weakest predicate temporal relevant logics are as

follows: $T_0TcQ =_{df} T_0Tc + \{IQ1\sim IQ5\} + \forall I$, $T_0EcQ =_{df} T_0Ec + \{IQ1\sim IQ5\} + \forall I$, $T_0RcQ =_{df} T_0Rc + \{IQ1\sim IQ5\} + \forall I$.

We can obtain some *spatial relevant logics* by adding region connection, point position, and point adjacency axiom schemata into the various predicate strong relevant logics. For examples: $STcQ =_{df} TcQ + \{RCC1, RCC2, PRCC1\sim PRCC7, APC1, APC2\}$, $SEcQ =_{df} EcQ + \{RCC1, RCC2, PRCC1\sim PRCC7, APC1, APC2\}$, $SRcQ =_{df} RcQ + \{RCC1, RCC2, PRCC1\sim PRCC7, APC1, APC2\}$.

Finally, we can obtain various spatio-temporal relevant logics by adding region connection, point position, and point adjacency axiom schemata into the various predicate temporal relevant logics. For examples: $ST_0TcQ =_{df} T_0TcQ + \{RCC1, RCC2, PRCC1\sim PRCC7, APC1, APC2\}$, $ST_0EcQ =_{df} T_0EcQ + \{RCC1, RCC2, PRCC1\sim PRCC7, APC1, APC2\}$, $ST_0RcQ =_{df} T_0RcQ + \{RCC1, RCC2, PRCC1\sim PRCC7, APC1, APC2\}$.

3   Reciprocal Logics

Based on strong relevant logics and spatio-temporal relevant logics, we can construct various *reciprocal logics* to underlie specifying, verifying, and reasoning about reciprocal relationships by introducing predicates and related axioms about various reciprocal relationships into strong relevant logics and spatial-temporal relevant logics. Therefore, they are conservative extensions of temporal relevant logics as well as strong relevant logics. As a case study, here we focus our interests on trust relationships. Other reciprocal relationships can be considered and dealt with in the same way.

Various reciprocal relationships may be symmetrical or unsymmetrical, transitive or non-transitive, but in general they are not reflective. Although there are many various definitions on the concept of trust, various trust relationships should have something in common. In general, a trust relationship must concern two parties, say A and B, such that A trusts B to do something, and any trust relationship is not necessarily symmetrical and not necessarily transitive. In many cases, a trust relationship is conditional in the form that A trusts B to do something, if some condition is true. On the other hand, the relationship of trust is just one of many kinds of reciprocal relationships in a human society and/or a cyber space.

Let $\{pe_1, pe_2, pe_3, \ldots\}$ be a countably infinite set of individual variables, called *person variables*. Atomic formulas of the form $TR(pe_1, pe_2)$ are read as 'person pe_1 trusts person pe_2.' Let $\{o_1, o_2, o_3, \ldots\}$ be a countably infinite set of individual variables, called *organization variables*. Atomic formulas of the form $TRpo(pe_1, o_1)$ are read as 'person pe_1 trusts organization o_1.' Atomic formulas of the form $TRop(o_1, pe_1)$ are read as 'organization o_1 trusts person pe_1.' Atomic formulas of the form $TRoo(o_1, o_2)$ are read as 'organization o_1 trusts organization o_2.'

Primitive Predicate: { TR (trust, $TR(pe_1, pe_2)$ means 'pe_1 trusts pe_2'), B (belong to, $B(pe_1, o_1)$ means 'pe_1 belongs to o_1') }

Defined Predicates: $NTR(pe_1, pe_2) =_{df} \neg(TR(pe_1, pe_2))$ ($NTR(pe_1, pe_2)$ means 'pe_1 does not trust pe_2'), $TREO(pe_1, pe_2) =_{df} TR(pe_1, pe_2) \wedge TR(pe_2, pe_1)$, ($TREO(pe_1, pe_2)$

means 'pe_1 and pe_2 trust each other'), $\textbf{\textit{ITR}}(pe_1, pe_2, pe_3) =_{df} \neg(\textbf{\textit{TR}}(pe_1, pe_2) \wedge \textbf{\textit{TR}}(pe_1, pe_3))$, ($\textbf{\textit{ITR}}(pe_1, pe_2, pe_3)$ means 'pe_1 does not trust both pe_2 and pe_3' (incompatibility)), $\textbf{\textit{XTR}}(pe_1, pe_2, pe_3) =_{df} (\textbf{\textit{TR}}(pe_1, pe_2) \vee \textbf{\textit{TR}}(pe_1, pe_3)) \wedge (\textbf{\textit{NTR}}(pe_1, pe_2) \vee \textbf{\textit{NTR}}(pe_1, pe_3))$, ($\textbf{\textit{XTR}}(pe_1, pe_2, pe_3)$ means 'pe_1 trusts either pe_2 or pe_3 but not both' (exclusive disjunction)), $\textbf{\textit{JTR}}(pe_1, pe_2, pe_3) =_{df} \neg(\textbf{\textit{TR}}(pe_1, pe_2) \vee \textbf{\textit{TR}}(pe_1, pe_3))$, ($\textbf{\textit{JTR}}(pe_1, pe_2, pe_3)$ means 'pe_1 trusts neither pe_2 nor pe_3'(joint denial)), $\textbf{\textit{TTR}}(pe_1, pe_2, pe_3) =_{df} (\textbf{\textit{TR}}(pe_1, pe_2) \wedge \textbf{\textit{TR}}(pe_2, pe_3)) \Rightarrow \textbf{\textit{TR}}(pe_1, pe_3)$, ($\textbf{\textit{TTR}}(pe_1, pe_2, pe_3)$ means 'pe_1 trusts pe_3, if pe_1 trusts pe_2 and pe_2 trusts pe_3'), $\textbf{\textit{CTR}}(pe_1, pe_2, pe_3) =_{df} \textbf{\textit{TR}}(pe_1, pe_3) \Rightarrow \textbf{\textit{TR}}(pe_2, pe_3)$, ($\textbf{\textit{CTR}}(pe_1, pe_2, pe_3)$ means 'pe_2 trusts pe_3, if pe_1 trusts pe_3'), $\textbf{\textit{NCTR}}(pe_1, pe_2, pe_3) =_{df} \neg\textbf{\textit{TR}}(pe_1, pe_3) \Rightarrow \textbf{\textit{TR}}(pe_2, pe_3)$, ($\textbf{\textit{NCTR}}(pe_1, pe_2, pe_3)$ means 'pe_2 trusts pe_3, if pe_1 does not trust pe_3'), $\textbf{\textit{CNTR}}(pe_1, pe_2, pe_3) =_{df} \neg\textbf{\textit{TR}}(pe_1, pe_3) \Rightarrow \neg\textbf{\textit{TR}}(pe_2, pe_3)$, ($\textbf{\textit{CNTR}}(pe_1, pe_2, pe_3)$ means 'pe_2 does not trust pe_3, if pe_1 does not trust pe_3'), $\textbf{\textit{TRpo}}(pe_1, o_1) =_{df} \forall pe_2(\textbf{\textit{B}}(pe_2, o_1) \wedge \textbf{\textit{TR}}(pe_1, pe_2))$, ($\textbf{\textit{TRpo}}(pe_1, o_1)$ means 'pe_1 trusts o_1'), $\textbf{\textit{NTRpo}}(pe_1, o_1) =_{df} \forall pe_2(\textbf{\textit{B}}(pe_2, o_1) \wedge \textbf{\textit{NTR}}(pe_1, pe_2))$, ($\textbf{\textit{NTRpo}}(pe_1, o_1)$ means 'pe_1 does not trust o_1', note that $\textbf{\textit{NTRpo}}(pe_1, o_1)$ is not a simple negation of $\textbf{\textit{TRpo}}(pe_1, o_1)$), $\textbf{\textit{TRop}}(o_1, pe_1) =_{df} \forall pe_2(\textbf{\textit{B}}(pe_2, o_1) \wedge \textbf{\textit{TR}}(pe_2, pe_1))$, ($\textbf{\textit{TRop}}(o_1, pe_1)$ means 'o_1 trusts pe_1'), $\textbf{\textit{NTRop}}(o_1, pe_1) =_{df} \forall pe_2(\textbf{\textit{B}}(pe_2, o_1) \wedge \textbf{\textit{NTR}}(pe_2, pe_1))$, ($\textbf{\textit{NTRop}}(o_1, pe_1)$ means 'o_1 does not trust pe_1', note that $\textbf{\textit{NTRop}}(o_1, pe_1)$ is not a simple negation of $\textbf{\textit{TRop}}(o_1, pe_1)$), $\textbf{\textit{TRoo}}(o_1, o_2) =_{df} \forall pe_1 \forall pe_2(\textbf{\textit{B}}(pe_1, o_1) \wedge \textbf{\textit{B}}(pe_2, o_2) \wedge \textbf{\textit{TR}}(pe_1, pe_2))$, ($\textbf{\textit{TRoo}}(o_1, o_2)$ means 'o_1 trusts o_2'), $\textbf{\textit{NTRoo}}(o_1, o_2) =_{df} \forall pe_1 \forall pe_2(\textbf{\textit{B}}(pe_1, o_1) \wedge \textbf{\textit{B}}(pe_2, o_2) \wedge \textbf{\textit{NTR}}(pe_1, pe_2))$, ($\textbf{\textit{NTRoo}}(o_1, o_2)$ means 'o_1 does not trust o_2', note that $\textbf{\textit{NTRoo}}(o_1, o_2)$ is not a simple negation of $\textbf{\textit{TRoo}}(o_1, o_2)$), $\textbf{\textit{TRpoEO}}(pe_1, o_1) =_{df} \textbf{\textit{TRpo}}(pe_1, o_1) \wedge \textbf{\textit{TRop}}(o_1, pe_1)$, ($\textbf{\textit{TRpoEO}}(pe_1, o_1)$ means 'pe_1 and o_1 trust each other'), $\textbf{\textit{TRooEO}}(o_1, o_2) =_{df} \textbf{\textit{TRoo}}(o_1, o_2) \wedge \textbf{\textit{TRoo}}(o_2, o_1)$, ($\textbf{\textit{TRooEO}}(o_1, o_2)$ means 'o_1 and o_2 trust each other')

Axiom Schemata: TR1: $\neg(\forall pe_1 \forall pe_2(\textbf{\textit{TR}}(pe_1, pe_2) \Rightarrow \textbf{\textit{TR}}(pe_2, pe_1)))$, TR2: $\neg(\forall pe_1 \forall o_1(\textbf{\textit{TRpo}}(pe_1, o_1) \Rightarrow \textbf{\textit{TRop}}(o_1, pe_1)))$, TR3: $\neg(\forall pe_1 \forall o_1(\textbf{\textit{TRop}}(o_1, pe_1) \Rightarrow \textbf{\textit{TRpo}}(pe_1, o_1)))$, TR4: $\neg(\forall o_1 \forall o_2(\textbf{\textit{TRoo}}(o_1, o_2) \Rightarrow \textbf{\textit{TRoo}}(o_2, o_1)))$, TR5: $\neg(\forall pe_1 \forall pe_2 \forall pe_3((\textbf{\textit{TR}}(pe_1, pe_2) \wedge \textbf{\textit{TR}}(pe_2, pe_3)) \Rightarrow \textbf{\textit{TR}}(pe_1, pe_3)))$, TR6: $\neg(\forall pe_1 \forall pe_2 \forall o_1((\textbf{\textit{TRpo}}(pe_1, o_1) \wedge \textbf{\textit{TRop}}(o_1, pe_2)) \Rightarrow \textbf{\textit{TR}}(pe_1, pe_2)))$, TR7: $\neg(\forall pe_1 \forall pe_2 \forall o_1((\textbf{\textit{TRop}}(o_1, pe_1) \wedge \textbf{\textit{TR}}(pe_1, pe_2)) \Rightarrow \textbf{\textit{TRop}}(o_1, pe_2)))$, TR8: $\neg(\forall o_1 \forall o_2 \forall o_3((\textbf{\textit{TRoo}}(o_1, o_2) \wedge \textbf{\textit{TRoo}}(o_2, o_3)) \Rightarrow \textbf{\textit{TRoo}}(o_1, o_3)))$

We can obtain various reciprocal logics for specifying, verifying, and reasoning about trust relationships by adding axiom schemata about trust relationships into the various predicate strong relevant logics, predicate temporal relevant logics, spatial relevant logics, or spatial-temporal relevant logics, respectively. For examples: if we do not take the notions of space and time into account but just consider some "static" trust relationships, then we can use the following logics: $\textbf{TrTcQ} =_{df} \textbf{TcQ} + \{TR1\sim TR8\}$, $\textbf{TrEcQ} =_{df} \textbf{EcQ} + \{TR1\sim TR8\}$, $\textbf{TrRcQ} =_{df} \textbf{RcQ} + \{TR1\sim TR8\}$. When we want to specify, verify, and reason about trust relationships themselves may change over space and time, we should use the following logics: $\textbf{TrST}_0\textbf{TcQ} =_{df} \textbf{ST}_0\textbf{TcQ} + \{TR1\sim TR8\}$, $\textbf{TrST}_0\textbf{EcQ} =_{df} \textbf{ST}_0\textbf{EcQ} + \{TR1\sim TR8\}$, $\textbf{TrST}_0\textbf{RcQ} =_{df} \textbf{ST}_0\textbf{RcQ} + \{TR1\sim TR8\}$. While if minimal or weakest temporal relevant logics are not adequate, then those characteristic axioms such as T5~T12 that correspond to various assumptions about time can be added into $\textbf{TrST}_0\textbf{TcQ}$, $\textbf{TrST}_0\textbf{EcQ}$, and $\textbf{TrST}_0\textbf{RcQ}$ respectively to obtain various stronger logics.

4   Concluding Remarks

We have proposed a new family of conservative extensions of relevant logic, named "reciprocal logic," for specifying, verifying, and reasoning about reciprocal relationships. We showed that various reciprocal logics can be obtained by introducing predicates and related axioms about reciprocal relationships into strong relevant logics and spatial-temporal relevant logics. Our case study is focused on trust relationships.

Because the strong relevance between the antecedent and the consequent of a conditional is intrinsically important to representing conditional reciprocal relationships, it is intrinsically important to construct various reciprocal logics based on strong relevant logics rather than traditional relevant logics as well as classical mathematical logic and its various classical conservative extensions.

On the other hand, the first three essential requirements for the fundamental logic mentioned in Section 1 are also essential to any applied logic for representing and reasoning about knowledge in a domain where there may be some incompleteness or inconsistency. Therefore, strong relevant logics can be considered as the universal basis of various applied logics for knowledge representation and reasoning.

References

1. Anderson, A.R., Belnap Jr., N.D.: Entailment: The Logic of Relevance and Necessity, Vol. I. Princeton University Press, Princeton (1975)
2. Anderson, A.R., Belnap Jr., N.D., Dunn, J. M.: Entailment: The Logic of Relevance and Necessity, Vol. II. Princeton University Press, Princeton (1992)
3. Burgess, J. P.: Basic Tense Logic. In: Gabbay, D., Guenthner F. (eds.): Handbook of Philosophical Logic, 2nd edition, Vol. 7. Kluwer Academic, Dordrecht (2002) 1-42
4. Cheng, J.: A Strong Relevant Logic Model of Epistemic Processes in Scientific Discovery. In: Kawaguchi, E., Kangassalo, H., Jaakkola, H., Hamid, I.A. (eds.): Information Modelling and Knowledge Bases XI. IOS Press, Amsterdam Berlin Oxford Tokyo Washington DC (2000) 136-159
5. Cheng, J.: Automated Knowledge Acquisition by Relevant Reasoning Based on Strong Relevant Logic. In: Palade, V., Howlett, R.J., Jain, L.C. (eds.): Knowledge-Based Intelligent Information & Engineering Systems, 7th International Conference, KES 2003, Oxford, UK, September 3-5, 2003, Proceedings, Part I. Lecture Notes in Artificial Intelligence, Vol. 2773, Springer-Verlag (2003) 68-80
6. Cheng, J.: Temporal Relevant Logic as the Logical Basis of Anticipatory Reasoning-Reacting Systems (Invited Paper). In: Daniel M. Dubois (ed.): Computing Anticipatory Systems: CASYS 2003 - Sixth International Conference. AIP Conference Proceedings, Vol. 718. American Institute of Physics, Melville (2004) 362-375
7. Cheng, J.: Spatio-temporal Relevant Logic as the Logical Basis for Specifying, Verifying, and Reasoning about Mobile Multi-agent Systems. In: Wang, S., Tanaka, K., et al. (eds.): Conceptual Modeling for Advanced Application Domains, ER 2004 Workshops Co-MoGIS, CoMWIM, ECDM, CoMoA, DGOV, and eCOMO, Shanghai, China, November 2004, Proceedings. Lecture Notes in Computer Science, Vol. 3289, Springer-Verlag (2004) 470-483
8. Cohn, A.G., Bennett, B., Gooday, J., Gotts, N.M.: RCC: A Calculus for Region based Qualitative Spatial Reasoning. GeoInformatica, Vol. 1 (1997) 275-316

9. Cohn, A.G., Bennett, B., Gooday, J., Gotts, N.M.: "Representing and Reasoning with Qualitative Spatial Relations About Regions. In: Stock, O. (ed.): Spatial and Temporal Reasoning. Kluwer Academic (1997) 97-134

10. Cohn, A.G., Hazarika, S.M.: Qualitative Spatial Representation and Reasoning: An Overview. Fundamenta Informaticae, Vol. 45 (2001) 1-29

11. van Benthem, J.: Temporal Logic. In: Gabbay, D.M., Hogger, C.J., Robinson, J.A. (eds.): Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 4. Oxford University Press, Oxford (1995) 241-350

12. Venema, Y.: Temporal Logic. In: Goble, L. (ed.): The Blackwell Guide to Philosophical Logic, Blackwell, Oxford (2001) 203-223

A New Paradigm Vertical Handoff Algorithm in CDMA-WLAN Integrated Networks

Kyung-Soo Jang[1], Jang-Sub Kim[2], Jae-Sang Cha[3], and Dong-Ryeol Shin[2]

[1] School of Computer Information, Kyungin Women's College
548-4 Gyesan-Dong, Gyeyang-Gu, InCheon, 407-740 Korea
`ksjang@kic.ac.kr`
[2] School of Information and Communication Engineering, Sungkyunkwan University
300 ChunChun-Dong, JangAn-Gu, Suwon, 440-740 Korea
`{jangsub,drshin}@ece.skku.ac.kr`
[3] Dept. of Information and Communication Engineering, Seokyeong University
16-1 Jung-Nung-Dong, SungBuk-Gu, Seoul, 136-704, Korea
`chajs@skuniv.ac.kr`

Abstract. The integration of WLANs and CDMA networks has recently evolved into hot issue. We propose a vertical handoff algorithm between WLAN and CDMA network. A handoff initiation is decided by the received signal strength (RSS). To reduce the unnecessary handoff probability, we also consider a distance. As a mechanism of network selection, we propose a context based network selection algorithms between WLAN to CDMA network, based on wireless channel assignment such as dropping probability, blocking probability, GoS (Grade of Service), the number of handoff attempts. As a decision making criterion, velocity threshold is determined to optimize the system performance. The optimal velocity threshold is adjusted to assign available channels to the mobile stations. The proposed scheme is validated by computer simulation.

1 Introduction

There has been a huge development in wireless communication technologies: mobile technology and WLAN technology. Mobile technologies offer high mobility but with low rates. In contrast, WLAN technologies offer high rates but with low mobility. The integration of mobile technology and WLAN technology, which compensates the coverage, bandwidth, and mobility to each other, achieves the requirements of the increasing user demands. In order to provide a convenient access of both technologies in different environments, interworking [1] and integration [2] of the two networks are regarded as a very important work.

The combination of WLAN and CDMA technology uses the best features of both systems. The key goal of this integration is to develop heterogeneous mobile data network, capable to support ubiquitous data services with very high data rates in hotspots. The effort to develop such heterogeneous networks, especially seamless roaming, is linked with many technical challenges including seamless vertical handoff across WLAN and CDMA technologies, security, common authentication, unified accounting & billing, WLAN sharing, consistent QoS and service provisioning, etc [3].

A handoff mechanism in an overlay CDMA and underlay WLAN should perform well so that the users attached to the CDMA just easily check the availability of the

underlay WLAN. The decision criteria for handoff (or network selection) can be based on the maximum link speed, reliability, power utilization, billing, cost, user preference, mobile speed, and Quality of Service like bandwidth, delay, jitter, and loss rate, etc [4]. For simplicity, we do only consider the mobile speed in this paper. A good handoff algorithm is to be derived in order to minimize unnecessary handoff attempts. An appropriate handoff control is also an important issue in system management for the sake of the benefits above with the overlaid cell structures.

In this paper, we deal with a vertical handoff decision making algorithms based on context information (GoS) and we propose a context based network selection method and handoff initiation and the corresponding mechanism between WLAN to CDMA system, based on wireless channel assignment. As a decision criterion, velocity threshold is determined to optimize the system performance. Combining WLAN/CDMA presents unique dimensioning problem. What is the system performance, given: number of radio channels, voice traffic, and data traffic (queuing delays). The proposed scheme is validated using computer simulation as a voice traffic model.

The rest of the paper is organized as follows. In Section 2, the proposed vertical handoff decision making algorithms are presented, problems are formulated, and core part of algorithmic decision procedure for the optimal velocity threshold for the WLAN and CDMA selection schemes. Section 3 explains the architecture for integrated networks, the mobility model, performance parameters (i.e. new call blocking probability and handoff call dropping probability, and Grade of Service). Simulations are performed in Section 4 to validate the proposed approach. Finally, the summary of the result and the future related research topics are presented in the conclusion section.

2   A Handoff Decision Making Process and Algorithm

A handoff decision making process decides when to invoke a vertical handover operation. A handoff decision making is classified into handoff initiation and network selection in this paper. Generally, the handoff initiation process evaluates user location changes (as users may leave or enter a particular network coverage). The evaluation of user location changes is carried out based on the Received Signal Strength (RSS). To reduce the unnecessary handoff probability, we also use distance. The next, as networks selection method, we also propose a context based network selection process between WLAN to CDMA network, based on wireless channel assignment. We focus on the network selection method which uses context information such as the GoS and the number of handoff attempts. The GoS is a function of the dropping probability and blocking probability. As a network selection parameter, velocity threshold is determined to optimize the system performance. The optimal velocity threshold is adjusted to assign available channels to the mobile stations.

Our proposed handoff decision making algorithm between WLAN and CDMA cellular networks considered velocity threshold related to GoS performance and handoff rates is shown in Fig. 1.

Where X_{CDMA} and X_{WLAN} is predefined signal strength threshold value when the handoff in CDMA network and WLAN, respectively. D_{CDMA} and D_{WLAN} is predefined distance threshold value, respectively. V_T is velocity threshold whether a fast MS or a slow MS.
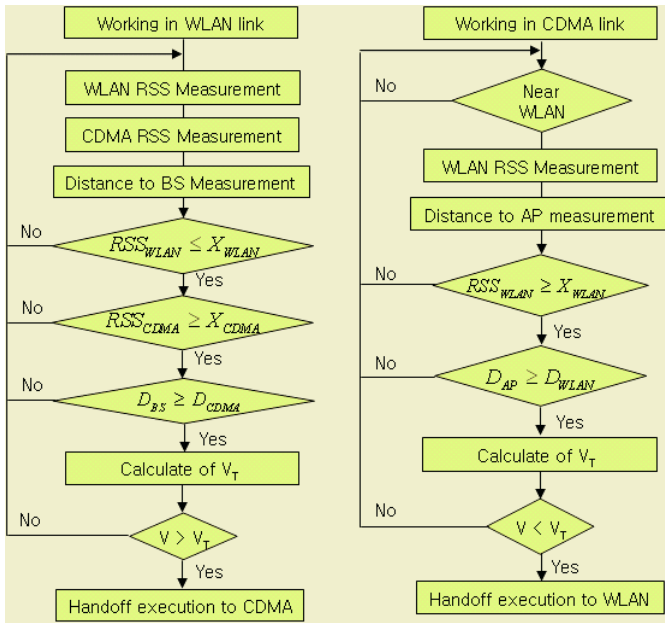
Fig. 1. Handoff decision making process

When the signal from the WLAN access point (AP) is strong, the MS is connected to the WLAN when the MS is larger than velocity threshold (V_T). As the MS moves away from the coverage of the access point, the signal strength falls and the distance between AP and MS far away. The MS then scans the air for other access points. If no other access point is available, or if the signal strengths from available access points are not strong enough, the handoff algorithm uses this information along with other possible information to make a decision on handing off to the CDMA network.

Handoff from the CDMA network to WLAN will occur when the following relation is satisfied,

$$\{[RSS_{WLAN} \geq X_{WLAN}] \, and \, [D_{AP} \leq D_{WLAN}] \, and \, [V \leq V_T]\}$$

On the other hand, when the active MS is using WLAN link, handoff from WLAN to CDMA network will occur when the following relation satisfied,

$$\{[RSS_{WLAN} \leq X_{WLAN}] \, and \, [RSS_{CDMA} \geq X_{CDMA}] \, and \, [D_{AP} \leq D_{WLAN}] \, and \, [V \geq V_T]\}$$

Handoff criteria in this direction should consider the criteria of RSS and the distance on CDMA link and velocity.
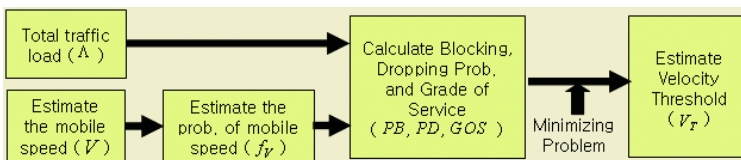
Fig. 2. The estimation of velocity threshold

In a proposed network selection process, the estimation of the velocity threshold (V_T) procedure at carried out the system is shown in Fig. 2. For the estimation of the mobile speed, Global Positioning System (GPS) can provide adequate location information. Using GPS information from the user signal, we can estimate for user's velocity. We develop the network selection algorithm based on an velocity threshold. The problem here is to find V_T improving the GoS with the given traffic parameters. We have to find the velocity threshold satisfied the following equation.

$$\min_{V_T}\{GoS(V)\} \tag{1}$$

The procedure is now concerned with the GoS in Equation (7). The GoS can be written as a function of V_T, and hence finding the optimum value of V_T minimizing the value of GoS and N_h is a typical minimization problem.

3   Performance Measures and Analysis

3.1   System Description

We consider a large geographical area covered by contiguous WLANs. The WLAN constitutes the lower layer of the two-layer hierarchy. All the WLANs are overlaid by a large CDMA system. The overlaying CDMA system forms the upper cell layer. Each CDMA system is allocated c_0 traffic channels, and the number of channels allocated to the WLAN cell-i is c_i, $i = 1, 2, \cdots, N$. All channels are shared among new calls and handoff calls. In our system, mobile stations (MSs) are traversing randomly the coverage area of WLAN and CDMA system. We further assume that an MS does not change its speed during a call.

We present analytical results for the proposed system. As stated, our objective is to focus on simple and tractable mechanism for which analytical results can give an insight into handoff between different networks. According to the velocity threshold, all the mobile users are divided into two groups; slower moving users (λ^S) and fast moving users (λ^F). In order to determine the value, which is one of the main goals of this study, a few assumptions related to mobility characteristics are made in system.

The assumptions we employ in the mobility models are taken from [5].

3.2   The New Call Blocking Probability of WLAN and CDMA System

We denote the blocking probability of calls from CDMA system and WLAN by P_{B0} and P_{B1}, respectively. And the handoff traffic from slow and fast mobiles is denoted as follows. The λ_{h0}^F and λ_{h0}^S is the rate of fast and slow mobile handoff traffic in a CDMA systems, respectively. The λ_{h1}^F and λ_{h1}^S is the rate of fast and slow mobile handoff traffic in a WLAN, respectively.

The aggregate traffic rate into the WLAN due to a slow MS and fast MS is computed as follows:

$$\lambda_1^S = \lambda_{n1}^S + \lambda_{h1}^S$$
$$\lambda_1^F = 1/N \times (\lambda_{n0}^F + \lambda_{h0}^F) P_{B0} + \lambda_{h1}^F \tag{2}$$

The generation rate of the handoff traffic of a slow MS and fast MS in a WLAN is given as follows:

$$\lambda_{h1}^{S} = P_{h1}^{S}(\lambda_{n1}^{S} + \lambda_{h1}^{S})(1 - P_{B1})$$
$$\lambda_{h1}^{F} = P_{h1}^{F}\{1/N \times (\lambda_{n0}^{F} + \lambda_{h0}^{F})P_{B0}(1 - P_{B1}) + \lambda_{h1}^{F}(1 - P_{B1})\}$$

(3)

The parameter ρ is the actual offered load to a WLAN from the new call arrival and the handoff call arrival. Invoking this important property, we can use $\rho_1 = \lambda_1^{S} / \mu_1^{S} + \lambda_1^{F} / \mu_1^{F}$ as the offered load to the WLAN, the Erlang-B formula calculates the blocking probability with the traffic ρ_1 and the number of channels c_1 [6]

$$P_{B1} = B(c_1, \rho_1)$$

(4)

Like as the new call blocking probability of WLAN, we can use $\rho_0 = \lambda_0^{S} / \mu_0^{S} + \lambda_0^{F} / \mu_0^{F}$ as the offered load to CDMA system, and blocking probability can be written as (4).

3.3   The Handoff Call Dropping Probability of WLAN and CDMA System

Slow MSs are supposed to use WLAN channels. However, since handoff to CDMA system is also allowed, the probability of handoff call drop in WLAN can be calculated as follows. Let P_{10} denote the probability that a slow MS fails to be handoffed to a near WLAN. Then the handoff call dropping probability is

$$P_D^{S} \approx P_{10}P_{B0} + P_{10}(1 - P_{B0})P_{F0}^{S}$$

(5)

Here P_{F0}^{S} is the probability that a slow MS handoff to CDMA system fail. The P_{10} is defined in such a way that the i th handoff request is successful but the $(i+1)$ th request is dropped:

$$P_{10} = f_1 + s_1 f_1 + s_1^{2} f_1 + \cdots = f_1 /(1 - s_1)$$

(6)

where $f_1 = P_{h1}P_{B1}$ and $s_1 = P_{h1}(1 - P_{B1})$. f_i describe the probability that handoff fails due to channel shortage and the s_i is the probability of successful handoff.

3.4   The Number of Handoffs and Grade of Service (GoS)

Among many system performance measures, GoS is most widely used. In fact users complain much more for call dropping than for call blocking. It is evaluated using the prespecified weights, PB and PD ,

$$GoS = (1 - \alpha)PB + \alpha PD$$

(7)

where PB and PD represent the blocking and dropping prob. of systems, respectively. The weight α emphasizes the dropping effect with the value of larger than one half.

4   Numerical Examples

The proposed procedure is tested with a number of numerical examples for the over-laid structure. The test system consists of 10 WLANs in the CDMA system. The total traffic $\Lambda = \lambda_0 + n\lambda_1$, where λ_0 and λ_1 are the new call arrival rate for the CDMA system and the WLAN, respectively. The radius of the WLAN and the CDMA system are assumed 300m and 1000m, respectively. The average call duration is $1/\mu = 120$ sec. The number of channels in each CDMA system and WLAN is $c_0 = 30$, $c_1 = 10$ for the total $\Lambda = 60$ Erlang. Assume the traffic mobility distribution is same as [7].
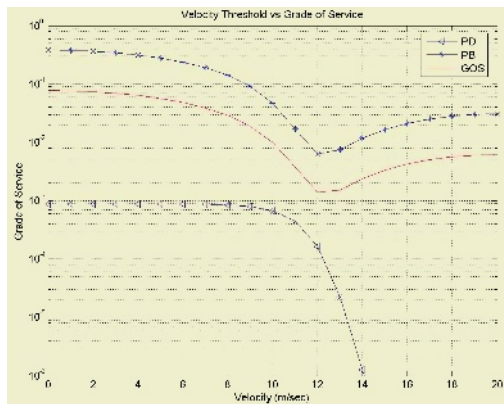
Fig. 3. Grade of Service vs. velocity threshold

We investigate the GoS, which is a function of both the traffic load and mobility distribution. Fig. 3 shows the plot of Equation (7). The vertical arrows in the figure show the range of the possible velocity thresholds at a certain load level. The lowest point in the range corresponds to the maximum allowable and optimal velocity threshold. Optimal V_T is 12m/sec. If the GoS value given by 10^{-2}, the range of opti-mal threshold value (V_T) become a larger than 10m/sec. But the optimal threshold value which have minimum handoff rates become 10 m/sec. The smaller threshold value become the optimal value. Thus, the optimal threshold value becomes 10 m/sec. In this case, more users are serviced in the CDMA system while the WLAN serves the fewer users. As a results, the WLAN will give rise to a higher number of handoff requests for high-mobility users, and the corresponding number of handoff requests of the calls in progress may cause an excessive processing load in the network.

With all the observations in mind, the strategy we proposed has desirable charac-teristics, i.e., finding the optimal velocity threshold value of GoS and the number of handoff rate.

5   Conclusion and Future Work

We have proposed a handoff decision making with handoff initiation and network selection deciding the optimal velocity threshold in order to improve the GoS and

minimize the number of handoff attempts with the given traffic volume in WLAN-CDMA integrated network. The simulation results show the dependency of the system performance upon the velocity threshold, V_T. The velocity threshold has shown to be an important system parameter that the system provider should determine to produce better GoS and lower handoff rate. From the simulation results we were able to validate the procedures determining the optimal V_T in which depends upon GoS as well as the number of handoff attempts. If knowledge based system for traffic and velocity information of mobile station is adopted, unnecessary handoffs can be noticeably reduced and fast handoff is possible.

References

1. K. Ahmavaara, H. Haverinen, and R. Pichna, "Interworking architecture between 3GPP and WLAN systems," IEEE Commun. Mag., vol.41, no.11, pp.74-81, Nov. 2003.
2. "Requirements and architectures for interworking between HIPERLAN/3 and 3rd generation cellular system," Tech. rep. ETSI TR191.957, Aug. 2001.
3. Milind M. Buddhikot, Girish Chandranmenon, etal, Design and Implementation of a WLAN/CDMA2000 Interworking Architecture, IEEE Communications Magazine, November 2003.
4. Balasubramaniam, S., Indulska, J., Vertical Handover Supporting Pervasive Computing in Future Wireless Networks, Computer Communication Journal, Special Issue on 4G/Future Wireless networks. Vol 27/8, pp.708-719., 2003.
5. Kwan L. Yeung and Sanjiv Nanda. "Channel Management in Microcell/Macrocell Cellular Radio Systems." IEEE Transactions on Vehicular Technologies, 45(4):601–612, November 1996.
6. W. Fischer and K. S. Meier-Hellstern. "The Markov Modulated Poisson Process (MMPP) cookbook." Performance Evaluations, 18:149–171, 1992.
7. JangSub Kim, WooGon Chung, HyungJin Choi and JongMin Cheong, Soon Park, "Determining Velocity Threshold for Handoff Control in Hierachically Structured Networks," PIMRC 98, 1998.

Efficient Revocation of Security Capability in Certificateless Public Key Cryptography

Hak Soo Ju[1], Dae Youb Kim[2], Dong Hoon Lee[2],
Jongin Lim[2], and Kilsoo Chun[1]

[1] Korea Information Security Agency(KISA), Korea
{hsju,kschun}@kisa.or.kr
[2] Center for Information and Security Technologies
{david_kdy,donghlee,jilim}@korea.ac.kr

Abstract. This paper presents the first mediated certificateless public key encryption and signature schemes. We also extend our schemes into hierarchical schemes. Our schemes does not suffer from the key escrow property that seems to be inherent in the mediated identity-based schemes. Key escrow is not always a good property for all applications because the exposure of a master key enable all the users' private keys to be leaked. Our mediated certificateless public key encryption and hierarchical schemes also support role based access control (RBAC) without the key escrow to manage the access to resources of a system. We finally describe security of our schemes and compare our schemes with the mediated identity based schemes from efficiency points of view.

1 Introduction

Revocation is one of the main difficulties faced in implementing Public Key Infrastructures (PKIs). Boneh et al. [1, 2] introduced an efficient method for obtaining instantaneous revocation of a user's public key and is called the mediated RSA(mRSA). Their method was to use an online semi-trusted entity called the SEcurity Mediator(SEM) which has a piece of each user's private key. In such a setting, a signer can't decrypt/sign a message without a token information generated by the SEM. Instantaneous revocation is obtained by instructing the mediator to stop helping the user decrypt/sign messages. This approach has several advantages over previous certification revocation techniques such as Certificate Revocation Lists(CRLs) and the Online Certificate Status Protocol(OCSP): fast revocation and fine-grained control over users' security capabilities.

Recently Libert and Quisquater showed that the SEM architecture in a mRSA can be applied to the Boneh-Franklin identity-based encryption and GDH signature schemes [3]. Nali, Miri and Adams have shown that Libert and Quisquater's mediated identity-based encryption scheme is suitable for the cryptographic support of role-based access control(RBAC) in [7]. They also presented the first mediated hierarchical identity based encryption and signature schemes by extending the mediated identity-based encryption scheme of Libert and Quisquater in [8].

Unfortunately, all identity based cryptographic schemes have inherent weakness, a key escrow property. Our main contribution is to remove key escrow property in the above mentioned mediated identity-based schemes and the hierarchical schemes by designing the first mediated certificateless public key encryption, signature and hierarchical schemes. A cryptosystem with the key escrow property has some serious disadvantages. For example once the master key is exposed, all the users' private keys are leaked and all the prior communication information is under the threat of exposure. In mediated identity based encryption and signature schemes, the trusted authority $\mathcal{TA}$ (in fact PKG) issues a partial private key $d_{ID,user}$ for user and another partial private key $d_{ID,sem}$ for SEM from a private key d_{ID} using its master secret key. As the result, the $\mathcal{TA}$ is still able to decrypt or sign any messages. Moreover, the hierarchical schemes of [8] still have an undesirable escrow property. Here, we present the hierarchical encryption scheme which will protect all users against any of their ancestors and against the root $\mathcal{TA}$.

Certificateless public key cryptography (CL-PKC) [5] do not explicitly provide revocation of users' security capabilities. This is natural since it aims to avoid the use of certificates as in ID-PKC. On the other hand, revocation is often necessary and even imperative. The way to obtain revocation in CL-PKC systems can be handled in the same way as in Boneh-Franklin IBE [6]. Their method is to require time dependent public keys, e.g., public keys derived from identities combined with time or date stamps. This has an unfortunate consequence of having to periodically reissue all private keys in the system. Moreover, these keys must be periodically and securely distributed to individual users. In contrast, our mediated CL-PKC schemes inherits its fine-grained revocation as in mRSA. The remainder of this paper is organized as follows: In section 2, we present our scheme, mCL-PKE to show how to remove the key escrow property. A security analysis of our scheme is given in section 3. In section 4 and 5 we present a signature scheme and show how to extend our schemes into hierarchical schemes. Section 6 discusses application aspects of our schemes. Finally, we conclude in section 7.

2   Our Mediated Certificateless Public Key Encryption

Our mediated certificateless public key encryption (mCL-PKE) scheme uses two private keys $< D_{ID}, x >$. The first private key issued by a trusted authority $\mathcal{TA}$ is used to inherit one to one mapping between public ID and private key D_{ID}. We use the second private key $x \in_R Z_q^*$ in order to remove a key escrow property and generate the corresponding pubic key P_{pub}. We name the private keys $< D_{ID}, x >$ as $< PrkeyI, PrkeyII >$.

In our scheme the first private key D_{ID} are split into shares via a one out of two secret sharing scheme, with one share held by the user and the other by the SEM. The encryption is then performed as a function of ID, P_{pub} as in the CL-PKE scheme [5]. The decryption requires the cooperation of a user and a SEM created by a $\mathcal{TA}$. Our scheme using this mediator SEM supports instantaneous revocation by instructing the SEM to stop interacting the user. Moreover, in

our scheme the key escrow property is removed because only the user knows the second private key x corresponding to P_{pub}.

1. **Setup**. Given a security parameter k, a $\mathcal{TA}$:
   (a) Run $\mathcal{IG}$ with input k in order to generate output $< G_1, G_2, \hat{e} >$ where G_1 and G_2 are groups of some prime order q and a bilinear map $\hat{e} : G_1 \times G_1 \rightarrow G_2$.
   (b) Choose an arbitrary generator $P \in G_1$.
   (c) Picks s uniformly at random from Z_q^* and computes $P_0 = sP$.
   (d) Choose cryptographic hash functions
   $H_1 : \{0,1\}^* \rightarrow G_1^*$, $H_2 : G_2 \rightarrow \{0,1\}^n$
   $H_3 : \{0,1\}^n \times \{0,1\}^n \rightarrow Z_q^*$, $H_4 : \{0,1\}^n \rightarrow \{0,1\}^n$
   where n denotes the size of plaintexts. The system's public parameters are $params =< G_1, G_2, \hat{e}, n, P, P_0, H_1, H_2, H_3, H_4 >$ while the master key $s \in Z_q^*$ is kept secret by the $\mathcal{TA}$.

2. **Key Generation**. Given a user of identity ID_A, the $\mathcal{TA}$ computes $Q_A = H_1(ID_A) \in G_1^*$ and $D_{ID} = sQ_A$ as a $PrkeyI$. Then it chooses random numbers $D_{ID,user} \in_R G_1$ and computes $D_{ID,sem} = D_{ID} - D_{ID,user}$. The $\mathcal{TA}$ gives the partial private key $D_{ID,user}$ to the user A and $D_{ID,sem}$ to the SEM over a confidential and authentic channel. The user A selects $x_A \in Z_q^*$ as his $PrkeyII$ and construct his public key as $P_{pub} =< X_A, Y_A >=< x_A P, x_A P_0 >$.

3. **Encrypt**. To encrypt $M \in \mathcal{M}$ for a user A with identifier $ID_A \in \{0,1\}^*$ and public key $P_{pub} =< X_A, Y_A >$, perform the following steps:
   (a) Check that $X_A, Y_A \in G_1^*$ and that the equality $\hat{e}(X_A, P_0) = \hat{e}(Y_A, P)$ holds. If not, return "Error" and abort encryption.
   (b) Compute $Q_A = H_1(ID_A) \in G_1^*$.
   (c) Choose a binary string $\sigma \in \{0,1\}^n$ and compute $r = H_3(\sigma, M)$.
   (d) Compute $U = rP \in G_1$, $g = \hat{e}(Y_A, Q_A)^r \in G_2$.
   (e) The ciphertext is $C =< U, V, W >=< rP, \sigma \oplus H_2(g), M \oplus H_4(\sigma) >$.

4. **Decrypt**. When receiving $C =< U, V >\in \mathcal{C}$, the user A forwards it to the SEM. They perform the following steps in parallel.
   – SEM: 1. Check if the user's identity ID_A or public key is revoked. If it is, return "Error".
   2. Compute $g_{sem} = \hat{e}(U, D_{ID,sem})$ and send it to the user A.
   – USER: 1. A computes $g_{user} = \hat{e}(U, D_{ID,user})$.
   2. When receiving g_{sem} from the SEM, A computes $g' = g_{sem}^{x_A} g_{user}^{x_A}$.
   3. A computes $\sigma' = V \oplus H_2(g')$ and then $M' = W \oplus H_4(\sigma')$.
   4. A checks the ciphertext's validity: $U = r'P$ with $r' = H_3(\sigma', M')$.

Remark 1. In our scheme, the $\mathcal{TA}$ is trusted to not replace the public keys of users and to issue only one private key $PrkeyI$ to the correct user. This means that a cheating $\mathcal{TA}$ impersonates an honest user by replacing a user's public key by one for which it knows the private key $PrkeyII$. We can use a binding technique of [5] for these problems. This binding restricts A to using a single public key. Moreover, if $\mathcal{TA}$ impersonates a user A there will be two

private keys for ID_A with different public keys and the $\mathcal{TA}$ can be identified as having misbehaved.

3   The mCL-PKE's Security

This section briefly discusses the security of our mCL-PKE scheme. We will show that the mCL-PKE is weakly semantically secure against inside attackers with an argument similar to the one provided for the proof of Theorem 4.1 of [3]. Insider attackers can access only the user part of the private key $PrkeyI$ corresponding to any identity but the one on which they are challenged. We also consider type I adversary $\mathcal{A}_\mathcal{I}$ and Type II adversary $\mathcal{A}_\mathcal{II}$ with and without the master key respectively as in a security model for CL-PKE [9].

Theorem 1. *If there exists a Type I or Type II IND-mID-CCA adversary with non-negligible advantage against mCL-PKE, then there exists an adversary $\mathcal{B}$ with non-negligible advantage against the CL-PKE scheme.*

Proof. (Sketch) Only outline is presented here for space limitation. Please refer to the full version of this paper for detailed proof. To prove Theorem 1, we first modify the notion of weakly semantically security against insider attacks (denoted by IND-mID-wCCA) given in [3] to consider type I and type II adversaries. In the random oracle model, we will use the attacker $\mathcal{A}$ against mCL-PKE scheme to build an adversary $\mathcal{B}$ that is able to distinguish ciphertexts produced by the CL-PKE. Overall the proofs are similar to the one of Theorem 4.1 in [3]. □

Remark 2. Our scheme assumed that users' private keys must be protected to ensure chosen ciphertext security as in Libert and Quisquater's scheme [3]. That is, our scheme is secure against chosen ciphertext attack in a weaker sense. Note that this weak assumption can be strengthened in a strong sense by using the ciphertext format of Baek and Zheng's mediated scheme [4].

4   A Mediated Certificateless Signature Scheme

We will describe a mediated certificateless public-key signature (mCL-PKS) scheme that is based on a provably secure ID-PKC signature scheme [10].

1. **Setup**. This is identical to Setup of our scheme mCL-PKE, except that there is only one hash function $H : \{0,1\}^* \times G_2 \to Z_q^*$.
2. **Key Generation**. Identical to mCL-PKE.
3. **Sign**. To sign $M \in \mathcal{M}$, the user A with identifier $ID_A \in \{0,1\}^*$ and private key D_{ID}, perform the following steps:
   (a) Chooses a random number $a \in Z_q^*$.
   (b) Computes $r = \hat{e}(aP, P) \in G_2$ and $v = H(M, r) \in Z_q^*$.
   (c) Sends v to SEM and perform the following steps in parallel.
       – SEM: 1. Check if the user's identity ID_A or public key is revoked. If it is, return "Error".
       2. Compute $U_{sem} = vD_{ID,sem}$ and send it to the user A.

- USER: 1. A computes $U_{user} = vD_{ID,user}$.

  2. When receiving U_{sem} from the SEM, A computes $U = x_A(U_{sem} + U_{user}) + aP \in G_1$.

  3. Returns $< U, v >$ as a signature of M.

4. **Verify**. When receiving $< U, v >$ on a message $M \in \mathcal{M}$ for identity ID_A and public key $< X_A, Y_A >$, the verifier :

   (a) Check that the equality $\hat{e}(X_A, P_0) = \hat{e}(Y_A, P)$ holds. If not, return "Error" and abort verification.

   (b) Compute $r' = e(U, P) \cdot e(Q_A, -Y_A)^v$.

   (c) Accepts the signature if and only if $v = H(M, r')$ holds.

5   Mediated Hierarchical CL-PKE

This section describes a mediated hierarchical certificateless public key encryption scheme denoted by mHCL-PKE. Our mHCL-PKE scheme eliminate all kinds of key escrow to any ancestor of an user in the mHIDE scheme [8]. We assumes that there exist two disjoint tree-shaped hierarchies of users and SEMs, respectively. Moreover, the root node of two hierarchies is a root $\mathcal{TA}$ and a set of users is associated to each SEM. We denote by $Level_t$ the set of nodes located at t^{th} level of both hierarchies. Except the root node, every node is identified by an ID-tuple $\overline{ID_t} = (ID_1, ID_2, \ldots, ID_t)$. The major steps of our scheme are identical to the ones in [8].

1. **Root Setup**. This algorithm is identical to Setup for mCL-PKE, except that now the ciphertext space for a level t. For ease of presentation, we denote the master key by x_0 instead of s. So we have $P_0 = x_0P = (x_{0,user} + x_{0,sem})P$, $P_{0,user} = x_{0,user}P$ and $P_{0,sem} = x_{0,sem}P$.

2. **Key Generation**. Given each of its child-user $\overline{ID_t} = (ID_1, ID_2, \ldots, ID_t)$, the $User_{t-1}$ selects $x_{t-1,user} \in_R Z_q^*$ and computes $Q_t = H_1(ID_1|| \ldots ||ID_t) \in G_1^*$, a partial private key $D_{t,user} = D_{t-1,user} + x_{t-1,user}Q_t$ of $PrkeyI$. The $User_{t-1}$ gives $D_{t,user}$ to its child $\overline{ID_t}$ over a confidential and authentic channel. Then he computes $R_{j,user} = x_{j,user}P$ for $1 \le j \le t$, publicly gives $R_{j,user}$ to its child $User_t$. The user $\overline{ID_t}$ selects $z_t \in Z_q^*$ as his $PrkeyII$ and constructs his public key as $P_t = < X_t, Y_t > = < z_tP, z_tP_0 >$. Given each of its child-SEM_t associated with $User_t$ the SEM_{t-1} selects $x_{t-1,sem} \in_R Z_q^*$ and computes $Q_t = H_1(ID_1|| \ldots ||ID_t) \in G_1^*$, a partial private key $D_{t,sem} = D_{t-1,sem} + x_{t-1,sem}Q_t$ of $PrkeyI$. The SEM_{t-1} gives $D_{t,sem}$ to its child SEM_t over a confidential and authentic channel. Then the SEM_{t-1} computes $R_{j,sem} = x_{j,sem}P$ for $1 \le j \le t$, publicly gives $R_{j,sem}$ to its child SEM_t.

3. **Encrypt**. To encrypt $M \in \mathcal{M}$ for a user $\overline{ID_t}$ with public key $P_t = < X_t, Y_t >$, perform the following steps :

   (a) Check that $X_t, Y_t \in G_1^*$ and that the equality $\hat{e}(X_t, P_0) = \hat{e}(Y_t, P)$ holds. If not, return "Error" and abort encryption.

   (b) Compute $Q_i = H_1(ID_1||ID_2|| \ldots ||ID_i) \in G_1^*$ for each $2 \le i \le t$.

   (c) Choose a random value $r \in Z_q^*$.

(d) Compute $g = \hat{e}(Q_1, Y_t)$ and $V = M \oplus H_2(g^r)$.

(e) Compute $U_0 = rP$ and $U_i = rQ_i$ for $2 \leq i \leq t$.

(f) Set the ciphertext $C = <U_0, U_2, U_3, \ldots, U_t, V> \in \mathcal{C}$.

4. **Decrypt**. When receiving $C = <U_0, U_2, \ldots, U_t, V> \in \mathcal{C}$, the user with $\overline{ID_t}$ proceeds as follows:

(a) Check $(U_0, U_2, U_3, \ldots, U_t) \in G_1^t$. Otherwise, reject C.

(b) Forwards C to the SEM_t , so that the following steps be performed in parallel:
   - SEM_t :
     1. Checks if the user's identity $\overline{ID_t}$ or public key is revoked. If it is, returns "Error".
     2. Computes $g_{sem_t} = \hat{e}(U_0, D_{t,sem})(\prod_{i=2}^{t} \hat{e}(R_{i-1,sem}, U_i))^{-1}$ and send it to the user $\overline{ID_t}$.
   - USER $\overline{ID_t}$:
     1. Computes $g_{user_t} = \hat{e}(U_0, D_{t,user})(\prod_{i=2}^{t} \hat{e}(R_{i-1,user}, U_i))^{-1}$.
     2. When receiving g_{sem_t} from SEM_t, $\overline{ID_t}$ computes $g^r = g_{user_t}^{z_t} \cdot g_{sem_t}^{z_t}$.
     3. Computes $M = V \oplus H_2(g^r)$.

6   Application

Using our mCL-PKE scheme of section 2, we modify Nali, Adams and Miri's RBAC scheme [7] to remove the inherent key escrow property as follows: The role manager (RM) has the role of the $\mathcal{TA}$ in our mCL-PKE scheme. For each role ID_i, the RM generates a $D_{ID_i} = sH_1(ID_i)$ and a pair $(D_{ID_i,user}, D_{ID_i,sem})$. The RM gives $D_{ID_i,user}$ to the user and $D_{ID_i,sem}$, its sub role decryption key shares to the SEMs associated with ID_i. Each user chooses a private key x and publish $<xP, xP_0>$ by itself or via a directory service.

Database Manager (DBM) obtains from the RM the keys of all roles required to access m and encrypts m using these role keys with virtual identities $ID_{i_1}, ..., ID_{i_k}$. The DBM computes $Q_{ID_{i_j}} = H_1(ID_{i_j})$ for $j = 1, .., k$ and the virtual identity $Q_{VID} = \sum_{j=1}^{k} Q_{ID_{i_j}}$. Then the DBM obtains and stores in the cipher table (CT) the ciphertext $C = (U, V, W) = (rP, \sigma \oplus H_2(g), M \oplus H_4(\sigma))$ where $g = \hat{e}(Q_{VID}, xP_0)^r$.

When the user A wants to access a protected M, A obtains C and the list of roles from the DBM. A sends them to a minimum number of SEMs whose roles are all ancestors of the roles. The SEMs check whether the user's identity, public key or any of the role identities is revoked or an RBAC separation of duty is broken. If these conditions holds, the SEMs compute their partial decryption of C by using their shares of the role keys and send it to the user A. A complete the decryption of C by using his shares of the role keys.

Note that our scheme is the same as the RBAC scheme using mIBE except using the pubic key which matches the private key. The application of a mediated hierarchical identity based encryption (mHIDE) which extends mIBE is to support information access control in hierarchical structured communities of

users whose access privileges change very dynamically [8]. Using our mHCL-PKE scheme of section 5, we can modify the application to remove key escrow property similarly.

7 Conclusion

In this paper, we have shown that the method of mRSA to allow fast revocation of RSA keys can be used by certificateless public key cryptography (CL-PKC). The mediated CL-PKC (mCL-PKC) that combines CL-PKC and mediated cryptography tackles the issues asociated with certificate management in PKIs and supports fine grained revocation. Moreover, we have shown that our mCL-PKC is more suitable for the cryptographic support of role-based access control than mediated ID-PKC because our schemes do not suffer from the key escrow property of mediated ID-PKC.

One possible goal for future research is to design and analyze a mediated selective ID-secure schemes in the standard model without random oracles. Another aim is to design and investigate the applications of a mediated certificateless public key cryptography.

References

1. D.Boneh, X.Ding, G.Tsudik, and C.Wong, *A method for fast revocaiton of public key certificates and security capabilities*,In Proceedings of the 10th USENIX Security Symposium, USENIX, 2001.
2. D.Boneh, X.Ding, and G.Tsudik, *Fine-grained control of security capabilities*,ACM Transactions on Internet Technology (TOIT) Volume 4, Issue 1, February 2004.
3. B. Libert, J.-J. Quisquater, *Efficient revocation and threshold pairing based cryptosystems*, Symposium on Principles of Distributed Computing-PODC'2003, 2003.
4. Joonsang Baek and Yuliang Zheng, *Identity-Based Threshold Decryption*, Proceedings of the 7th International Workshop on Theory and Practice in Public Key Cryptogrpahy (PKC'04), LNCS, vol. 2947, Springer-Verlag, 2004, pp. 262-276.
5. S.S. Al-Riyami and K.G. Paterson, *Certificateless public key cryptography*, In Advances in Cryptology-ASIACRYPT 2003, volume 2894 of LNCS, pages 452-473. Springer-Verlag, 2003.
6. D.Boneh and M.Franklin, *Identity-based encryption from the Weil pairing*, In Advances in Cryptology-CRYPTO 2001, volume 2139 of LNCS, pages 213-229. Springer-Verlag, 2001.
7. D. Nali, C. Adams, and A. Miri, *Using Mediated Idenitity-Based Cryptography to Support Role-Based Access Control*, (ISC 2004), 2004.
8. D. Nali, A. Miri, and C. Adams, *Efficient Revocation of Dynamic Security Privileges in Hierarchically Structured Communities*, Proceedings of the 2nd Annual Conference on Privacy, Security and Trust (PST 2004), Fredericton, New Brunswick, Canada, October 13-15, 2004, pp. 219-223.
9. C.Gentry and A.Silverberg, *Hierarchical ID Based Cryptography*, In Advances in Crypotology-ASIACRYPT 2002, LNCS, pages 548-566, Springer-Verlag, 2002.
10. F. Hess, *Efficient identity based signature schemes based on pairings*, In K. Nyberg and H. Heys, editors, Selectd Areas in Cryptography 9th Annual International Workshop, SAC 2002, volume 2595 of LNCS, pages 310-324. Springer-Verlag, 2003.

A Study on Privacy-Related Considerations in Ubiquitous Computing Environment
(A Focus on Context Aware Environment Applied in Intelligent Technology)

Jang Mook Kang[1], Jo Nam Jung[2], Jae-Sang Cha[1], and Chun su Lee[3]

[1] Dept. Of Computer Engineering, Seokyeong University
16-1 Jung-Ryung dong, Seoul, South Korea
`mooknc@paran.com, chajs@skuniv.ac.kr`
[2] Dept. Of Computer Science & Information Engineering, Inha University
253 Yong-Hyun dong, Incheon, South Korea
`jjn10@korea.com,`
[3] Dept. Of Institute for Business Research & Education, Korea University
1, 5-ga, Anam-dong, Seoul, South Korea
`cybertrade@korea.ac.kr`

Abstract. It is said that the UC(Ubiquitous Computing) Environment will be come by the technological revolution in the future. This paper firstly points out the problem that engineer thinks that Ubiquitous computing connected to network have become more complexes and distributed, but most system-related research is still focused on the 'performance' or 'efficiency'. And that most studied papers in privacy have been focused on encryption technology and protocols for securing the personal information data transaction. This paper secondly points out the problem that law-related researcher thinks that law can controller technology beyond human control. Today, due to the complexity of the problem, new privacy-related research based on total solution is integrative needed to consider laws, technology, economy, norms and so on. This paper will focus on analysis about Privacy in Ubiquitous computing Environment and total protection considerations on privacy. We present new privacy-related considerations in UC (Ubiquitous computing) Environment, point out the mechanism which new re-design privacy concept (protection scope) is affected from the UC (Ubiquitous computing) Environment and provide the total considerations for Privacy protection especially personal information protection by improve the intelligent technology applied in context aware technology.

1 Introduction

It is said that the ubiquitous network environment will be come by the new technological revolution in the near future. For many people, the idea that technology is moving faster than society can adapt seems like common sense. The Ubiquitous computing embedded in everyday objects has been dubbed 'the invisible computing' or 'disappearing computing'. This paper will focus on analysis about Privacy in Ubiquitous computing Environment and total protection considerations on privacy. We present new privacy-related considerations in Ubiquitous computing Environment, point out the mechanism which new re-design privacy concept is affected from the Ubiquitous computing Environment and provide the total considerations for Privacy protection especially personal information protection applied in intelligent technology.

2   UC(Ubiquitous Computing) Environment

2.1   The Characteristics About Ubiquitous Computing Environment

The ultimate goal of ubiquitous computing is to place computers everywhere in the real world environment, proving ways for them to interconnect talk and work together [1],[2]. The concept is now shifting computing paradigm from machines in a room to the augmented contexts in the real world [3]. Now, new computing paradigm or ubiquitous computing environment will make privacy risk everywhere possible.

2.2   A Move to Ubiquitous Society Applied in Intelligent Technology

Currently, Korea is going through a rapid transition from an IT-based society to a ubiquitous one with its government at the forefront of the effort. The ubiquitous society creates a ubiquitous space by organically connecting cyber space to physical space and limited time to real time. Supposing that IT-based space is a mere extension of a cyber space, the ubiquitous space is an entirely new space created from the actual networking between cyber space and physical space (as shown in Figure 1). Cyber space, physical space, and ubiquitous space are created beyond the realm of conventional time and space. Consequently, ubiquitous computing environment entails a world driven by technologically based real time mobility and pervasiveness of computing functions. So much of the technology needed to protect ubiquitous network systems already exits. Cryptographic techniques can be applied in support of authentication, authorization, integrity, confidentiality and so on [4]. To be useful, however, the infrastructure supporting these technologies must be put in space (cyber-space, real-space) and the intelligent technology must be integrated with social law, norm, and commerce to give protection for individual privacy.

3   Privacy Problems

3.1   Privacy Crisis in Ubiquitous Computing Environment

Cooley's definition of privacy-the "right to enjoy life and be let alone" [5] - has our preference. Current state of the art of pervasive computing does not properly address security and privacy [6].

3.2   Definition of Privacy

The debate on privacy was initiated by the publication of Samuel D. Warren and Louis D. Brandeis. They claimed that the state of mind, emotion, and intelligence can be protected by suing against the libel disgrace of honor and ownership of a copyright law, but the privacy law should be upheld to protect the individual light to be let alone. In Korea, private life is protected under the privacy law of the right not to be invaded regarding residence (Constitution Article 16) and the right not to be invaded in terms of privacy and freedom. Article 18 of the Constitution protects the privacy in communication by stipulating, shall not be violated in terms of privacy in communication. Ubiquitous era extended the concept of privacy through the active control over personal information applied in context aware environment by used intelligent technology.
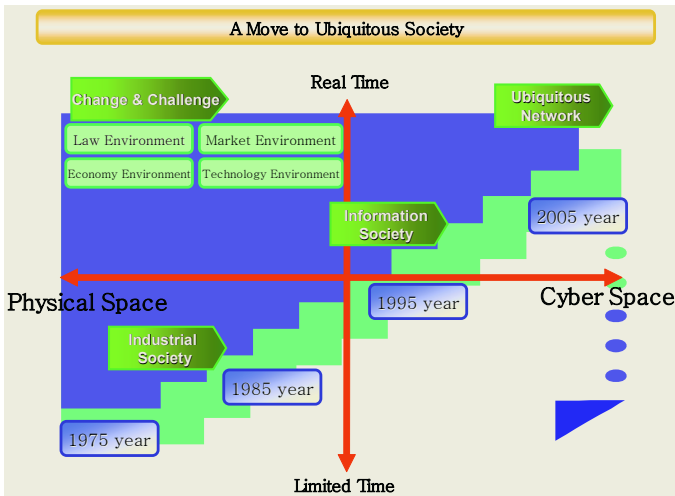
Fig. 1. A move to ubiquitous society

4   Privacy-Related Considerations

4.1   Personal Information Control

What is meant by 'personal information control'? A simple but useful definition of personal information control is 'the ability of an individual to control the terms under which their personal information is acquired and used.'[7]. Personal Information Control can be defined as a right to gain control of his or her own personal information. It is classified as a right to know and have access to the personal information. It is a way of active protection against privacy disclosure, which makes availability of competing values in privacy.

4.2   Changing Scope of Privacy

In our previous work, we can extension for scope of privacy as shown fig. 2. Ubiquitous Technology will make new space which is augmented real space by improved intelligent technology. So scope of privacy necessary change and adapt new technology-context aware environment. New protection scopes of privacy are physical space, cyber space and augmented space.

5   Conclusions

The concerns are not merely about security and privacy of ubiquitous computing, but about something much more significant; trust in the ubiquitous society. Therefore there is no simple solution that meets every security and privacy concern. The first step for addressing best strategy of Ubiquitous-based security and privacy is to build security into the re-definition of privacy, rather than attempt to introduce emerging technology or law. So, we present new privacy-related considerations in UC (Ubiquitous computing) Environment, point out the mechanism which new re-design privacy
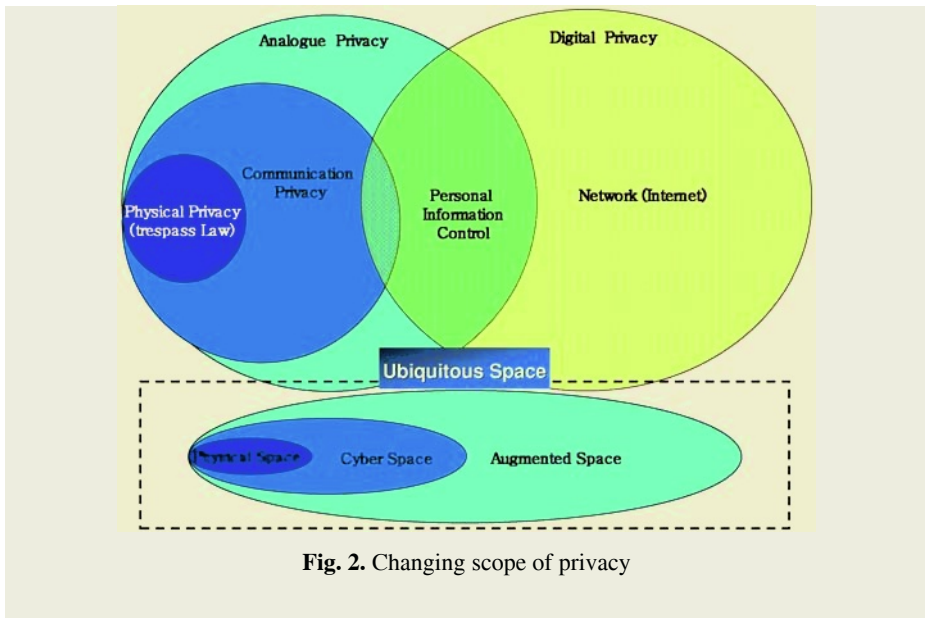
Fig. 2. Changing scope of privacy

concept (protection scope) is affected from the UC (Ubiquitous computing) applied in intelligent technology.

Of course, non-technical aspects are also important although scientists are outside the scope of their research. To create trustworthy ubiquitous computing environment applied in intelligent technology, a systematic approach that combines context aware technology, law, social norm and commerce imperatives are necessary.

Acknowledgement

References

1. M. Weiser. The Computer for the 21st Century. Scientific American, pages 66-75, September 1991.
2. M. Weiser. Some Computer Science Issues in Ubiquitous Computing. Communication of the ACM, Vol. 36 (7), pages 75-84, 1993.
3. Kimio Kuramitsu, Ken Sakamura. Towards ubiquitous database in mobile commerce. Proceedings of the 2nd ACM international workshop on Data engineering for wireless and mobile access, Pages 84-85, 2001.
4. Clifford Neuman.: Security, Payment, and Privacy for Network Commerce. IEEE Journal on selected areas in communications. Vol. 13(8), pages1530, 1995.
5. T. M. Cooley. A Treaties on the Law of Torts. Callaghan, Chicago, 1888.
6. M. Langheinrich Privacy by Design – Principles of privacy – Aware Ubiquitous systems. In Proceedings of Ubicomp 2001: Ubiquitous Computing: Third International Conference, LNCS 2001, Pages 273-291, Springer Verlag, Heidelberg, 2001.
7. Culnan MJ.: Protecting privacy online:is self-regulation working? J Public Policy Market. Vol 19(1), pages 20-26, 2000.

Optimal Operation for Cogenerating System of Micro-grid Network

Phil-Hun Cho[1], Hak-Man Kim[2,*], Myong-Chul Shin[1], and Jae-Sang Cha[3]

[1] Dept. of Electronic and Electrical Eng., SungKyunKwan Univ., Suwon, 440-746, Korea
`{Cho,Shin,youinaru}@kepco.co.kr`
[2] KERI, 28-1 Sengju-dong, Changwon, Kyung-Nam, 641-120, Korea
`kerihifi@naver.com`
[3] Dept. of Inform. And Commun. Eng. Seokyeong Univ., Seoul, 136-704, Korea

Abstract. This paper presents a mathematical model for optimal operating cogeneration of Micro-Grid Network. The electrical and thermal energy production relationship of cogeneration is based on its feasible operation region. In the proposed model, auxiliary boilers as well as cogererators are considered as thermal energy sources. Cost functions of energy sources and the loss function of electrical power are formulated by quadratic functions. The solution of Optimal operation for cogenerating system of micro-grid network is solved by quadratic program. Some examples are tested by proposed model and method.

1 Introduction

The portion of cogeneration in Micro-Grid Network is increasing as an electrical energy source. Expanded mathematical models of economic operation and planning for power system with cogenerators have been studied[1-9].

This paper describes a mathematical model for economic dispatch of the power system with cogenerators. To establish the mathematical model, the thermal and electrical energy production relationship of cogeneration and the thermal energy balance relationship between supply and demand are added to the conventional economic dispatch model for Micro-Grid Network. The formulation of the thermal and electrical energy production relationship of cogeneration is based on the cogeneration feasible operation region. To formulate the thermal energy balance relationship between supply and demand, it is assumed that the transmission of thermal network is lossless. Since the thermal energy is generally delivered through a local network, the assumption is reasonable. Auxiliary boilers as a thermal energy source are included for the thermal energy balance. In this paper, the energy cost functions of generators, cogenerators and auxiliary boilers are assumed as a quadratic form. The loss function of electrical network is also assumed as a quadratic form. As a result, the mathematical model of economic dispatch for cogenerating systems is represented as a quadratic programming problem.

Two cases are studied with the proposed mathematical model. In the first case, the effect of optimization of the electrical and thermal energy production ratio of cogenerators is considered. In the second case, the proposed mathematical model is applied to a cogenerating system with multiple thermal local areas.

---

* Correspondence author

2 Mathematical Model of Economic Dispatch for Cogenerating Systems

You are The objective function which is given by eq.(1) simply expresses electrical energy generation costs of generators, electrical and thermal energy generation costs of cogenerators and thermal energy generation costs of auxiliary boilers.

$$Min \sum_i c_{G,i}(p_i) + \sum_j c_{C,j}(p_j, q_j) + \sum_k c_{B,k}(q_k) \tag{1}$$

where,

$c_{G,i}(p_i)$: the energy generation cost function of the i-th generator

$c_{C,j}(p_j, q_j)$: the energy generation cost function of the j-th cogenerator

$c_{B,k}(q_k)$: the energy generation cost function of the k-th auxiliary boilers

i, j, k: indices identifies generators, cogenerators, auxiliary boilers
p: electrical energy
q: thermal energy

Constraints of economic dispatch for cogenerating systems can be formulated as follows.

Electrical energy produced by generators is limited by lower and upper bounds.

$$p_i^{min} \leq p_i \leq p_i^{max} \tag{2}$$

where,

p_i^{min} : lower bound of electrical energy generation of the i-th generator

p_i^{max} : upper bound of electrical energy generation of the i-th generator

Electrical energy produced by cogenerators is limited by lower and upper bounds.

$$p_j^{min} \leq p_j \leq p_j^{max} \tag{3}$$

where,

p_j^{min} : lower bound of electrical energy generation of the j-th cogenerator

p_j^{max} : upper bound of electrical energy generation of the j-th cogenerator

The thermal energy output of cogenerators must be within lower and upper limits resulting from selected Electrical-to-Steam ratio(E/S) limits [4].

$$q_j \geq \frac{1}{E_{Hj}} p_j \tag{4}$$

$$q_j \leq \frac{1}{E_{Lj}} p_j \tag{5}$$

where,

E_{Lj} : lower limit of E/S of the j-th cogenerator

E_{Hj} : upper limit of E/S of the j-th cogenerator
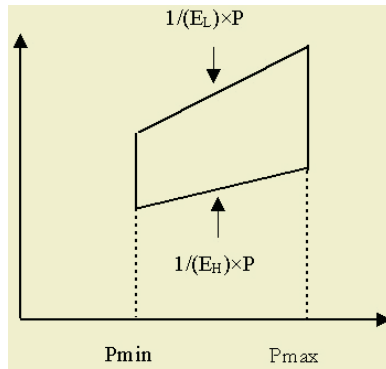
Fig. 1. A feasible operation region of cogenerators

Fig.1 shows a feasible operational region for cogenerators which is obtained from the above mentioned conditions [4].

The thermal energy produced by auxiliary boilers is limited by lower and upper bounds.

$$q_k^{\min} \leq q_k \leq q_k^{\max} \tag{6}$$

where,

$q_k^{\min}$: lower bound of thermal energy generation of the k-th auxiliary boiler

$q_k^{\max}$: upper bound of thermal energy generation of the k-th auxiliary boiler

Since thermal energy is generally delivered through a local area network, the thermal energy through the network is assumed to be lossless. The thermal demand of a local area is supplied from thermal energy sources of the area. With this assumption, the thermal energy balance of supply and demand is formulated as follows.

$$\sum_{j \in A} q_j + \sum_{k \in A} q_k = q_D^A \tag{7}$$

where,

q_D^A : thermal load of the A-th area

$\displaystyle\sum_{j \in A} q_j$: thermal energy supplied by cogenerators of the A-th area

$\displaystyle\sum_{k \in A} q_k$: thermal energy supplied by auxiliary boilers of the A-th area

Finally, the electrical energy balance of supply and demand is formulated as follows.

$$\sum_i p_i + \sum_j p_j = p_L + p_D \tag{8}$$

where,

p_D : electrical load

p_L : the loss of electrical network

3   Simulation and Results

3.1   Case Study #1

In this case study, the effect of energy production cost reduction according to optimal electrical and thermal energy production ratio, E/S, of cogeneration was studied.

Table 1 shows data and configuration of a given test system. In Table 1, G1 and G2 represent generators and CG1 represents cogenerator.

Table 1. Data and configuration of a test system

Facilities	Cost Function	Limits	Area
G1	$561+7.92p+0.001562p^2$	$150 \le p \le 600$	
G2	$310+7.85p+0.00194p^2$	$100 \le p \le 400$	
CG1	$750+5.62p+0.001255p^2+5.53q+0.001145q^2$	$5 \le p \le 50$ $E_L:0.8\ E_H:1.5$	A

In the test system, the electrical demand is 600[KW] and thermal demand is 40[KW]. The transmission loss of electrical energy is assumed as follows:

$$p_L = 0.00003 p_{G,1}^2 + 0.00006 p_{G,2}^2 + 0.00009 p_{C,1}^2 \tag{9}$$

where,

$p_{G,i}$: electrical energy generated by the i-th generator

$p_{C,j}$: electrical energy generated by the j-th cogenerator

Fixed E/S values of 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5 were selected between the lower limit(0.8) and the upper limit(1.5) of the E/S. Table 2 shows simulation results. The optimal E/S was selected as 1.25. The total energy generation cost with the optimal E/S, 1.25 was the lowest. Especially, there was no solution in some fixed E/S values(1.3 ~ 1.5).

From simulation results, it was verified that E/S gave effect on the total energy generation cost. Therefore, the optimization of E/S is a key factor to economic dispatch for cogenerating systems.

Table 2. Simulation results

E/S	0.8	0.9	1.0	1.1
Cost	6847.97	6834.61	6821.36	6808.20
E/S	1.2	1.3 ~ 1.5	1.25 (optimal)	
Cost	6795.15	infeasible	6788.66	

3.2   Case Study #2

For economic dispatch of cogenerating systems for multiple local thermal areas, test system data were summarized in Table 3 in which AB represents a auxiliary boiler.

In this test system, the electrical demand is 950[KW]. The thermal demand of the area A is 50[KW] and the thermal demand of the area B is 80[KW]. The transmission loss in the electrical network is represented as follows:

Table 3. Data of model system

Facilities	Cost Function	Limits	Area
G1	$561+7.92p+0.001562p^2$	$150 \leq p \leq 600$	
G2	$310+7.85p+0.00194p^2$	$100 \leq p \leq 400$	
CG1	$750+5.62p+0.001255p^2+5.53q+0.001145q^2$	$5 \leq p \leq 50$ E_L:0.8 E_H:1.5	A
CG2	$641+7.97p+0.001765p^2+6.43q+0.001241q^2$	$3 \leq p \leq 30$ E_L:0.6 E_H:1.2	A
CG3	$730+5.92p+0.00123p^2+5.58q+0.001243q^2$	$5 \leq p \leq 50$ E_L:0.4 E_H:0.9	B
AB	$540+5.63q+0.00155q^2$	$1 \leq q \leq 15$	B

$$p_L = 0.00003\,p_{G,1}^2 + 0.00006\,p_{G,2}^2 + 0.00009\,p_{C,1}^2 + 0.00012\,p_{C,2}^2 + 0.00008\,p_{C,3}^2 \qquad (10)$$

Table 4 shows the optimal solution of economic dispatch for cogenerating systems. For optimal plan, as in Table 4, cogenerator #1 produced 33.33[KW] and cogenerator #2 produced 16.67[KW] for thermal demand of area A. For thermal demand of area B, cogenerator #3 produced 65[KW] and the auxiliary boiler produced 15[KW]. For electrical demand, generator #1 produced 479.34[KW], generator #2 366.07[KW], cogenerator #1 produced 50[KW], cogenerator #2 produced 20[KW] and cogenerator #3 produced 50[KW].

Table 4. Simulation results of model system with two thermal areas

Facilities	Generation [KW]
G1	479.34
G2	366.07
CG1(Elec.)	50.0
CG1(Ther.)	33.33
CG2(Elec.)	20.0
CG2(Ther.)	16.67
CG3(Elec.)	50.0
CG3(Ther.)	65.0
AB	15.0
E/S(CG1)	1.5
E/S(CG2)	1.2
E/S(CG3)	0.77

4   Conclusions

In this paper, a economic dispatch model for cogenerating systems was proposed. The solution of optimal operation for cogenerating system of micro-grid network is solved by quadratic program. Some examples are tested by proposed model and method. In the first case, the effect of optimization of the electrical and thermal energy production ratio of cogenerators is considered. In the second case, the proposed mathematical model is applied to a cogenerating system with multiple thermal local areas. The proposed algorithm can be usefully used (a) considerations of multiple thermal local areas and (b) optimization of E/S based on the cogeneration feasible operation region.

References

1. S.Ashok and R.Banejee, "Optimal Operation of Industrial Congeneration for Load Management",IEEE Trans.on.Power System, Vol.18, no.2, May 2003
2. Shun-Hsien Huang, Wei-Jen Lee,"Optimal Operation Stratagy for Cogeneration Power Plants", Industry Applications Conference, 2004. 39th IAS Annual Meeting. Conference Record of the 2004 IEEE Vol. 3, Oct. 2004
3. N.N.Bengiamin,"Operation of Cogeneration Plants with Power Purchase Facilities", IEEE Trans. PAS, Vol. PAS-102, No.10, 1983.
4. 4. P.R.MacGregor and H.B.Püttgen, "Optimum Scheduling Procedure for Cogenerating Small Power Producing Facilities", IEEE Trans. Power System, Vol.4, No.3, pp.957-964, 1989.
5. P.R.MacGregor,"The net utility revenue impact of Small Power Poducing Facilities Operating Under Spot Pricing Policies", Ph.D dissertation, Georgia Institute of Technology, 1989.
6. Moslehi, M.Khadem, R.Bernal and G.Hernadez, "Optimization of Multiplant Cogeneration System Operation Including Electric and Steam Networks", IEEE Trans. Power System, Vol.6, No.2, pp.484-490, 1991.
7. Frans J. Rooijers and Robert A. M. van Amerongen, "Static Economic Dispatch for Cogeneration Systems", IEEE Trans. Power System, Vol. 9, No.3, pp.1392-1398, 1994.
8. Toshiyuki Tamura, Hiroyuki Kita, Ken-ichi Nishiya, Jun Hasegawa,"A Method for Determing the Optimal Operation of Dispersed Cogenerating Sources in Coordination with Boilres", T.IEE Japan, Vol. 114-B, No.12, 1994.
9. H.M. Kim, M.C. Shin, C.H. Kim, C.Y. Won, "Optimum Operation of Small Power Producing Facilities Interconnected with Power System", Tans. KIEE, Vol. 44, No. 4, pp. 409-417, 1995.
10. Allen J. Wood, Bruce F. Wollenberg, Power Generation Operation & Control, New York: John Wiley and Sons. Inc. 1996.
11. Frederick S. Hillier, Gerald J. Lieberman, Introduction to Operations Research, McGraw-Hill, 1990.
12. G.V. Reklaitis, A. Ravindran, K.M. Ragsdell, Engineering Optimization-Methods and Applications, John Wiley & Sons, Inc., 1983.

An Efficient and Secured Media Access Mechanism Using the Intelligent Coordinator in Low-Rate WPAN Environment⋆

Joon Heo[1] and Choong Seon Hong[2]

School of Electronics and Information, Kyung Hee University
1 Seocheon Giheung Yongin Gyeonggi 449-701 Korea
{heojoon,cshong}@khu.ac.kr

Abstract. The convenience of IEEE 802.15.4 (Low-rate Wireless Personal Area Network) based wireless access network will lead to widespread deployment in the evolutionary computing and adaptive systems. However, this use is predicated on an implicit assumption of confidentiality and availability. In this environment, malicious device may obtain an unfair share of the channel bandwidth. For example, IEEE 802.15.4 requires nodes competing for getting access to the channel to wait for a 'backoff' interval, randomly selected from a specified range, before initiating a transmission. Malicious devices may wait for smaller 'backoff' intervals than well-behaved devices, thereby obtaining an unfair advantage. Such misbehavior can seriously degrade the throughput of well-behaved devices. In this paper, we proposed an intelligent coordinator, a modification to IEEE 802.15.4 specification to simplify the prevention of such malicious devices. Proposed coordinator is able to ascertain whether to apply PTP (Priority Time Period). Simulation results indicate that our prevention mechanism is successful in handling MAC layer misbehavior. We plan to foster our solution in the evolutionary computing systems through further development.

1 Introduction

The growing importance of small and cheap wireless devices demands a common platform so that the device can communicate with each other. The IEEE 802.15.4 (Low-Rate WPAN) describes wireless and media access protocols for personal area networking devices. This standard specifies the physical (PHY) layer and Medium access control (MAC) sub-layer of a low cost, low power consumption and ad hoc wireless network. The MAC layer provides services that higher layers can use to access the physical radio. MAC layer protocols provide a means for reliable, single-hop communication links between devices [1][5][6]. This specification is meant to support a variety of applications, many of which are security sensitive. Wireless Medium Access Control (MAC) protocols such

---

as IEEE 802.15.4 and IEEE 802.11 use distributed contention resolution mechanisms for sharing the wireless channel. The contention resolution is typically based on cooperative mechanisms (e.g., random backoff before transmission) that ensure a reasonably fair share of the channel for all the participating nodes. In this environment, some malicious devices in the network may misbehave by failing to adhere to the network protocols, with the intent of obtaining an unfair share of the channel. The presence of malicious devices that deviate from the contention resolution protocol can reduce the throughput share received by conforming devices. A misbehaving node may obtain more than its fair share of the bandwidth by

- Selecting backoff values from a different distribution with smaller average backoff value (e.g., by selecting backoff values from range $[0, \frac{CW}{4}]$ instead of $[0, CW]$, where $CW(ContentionWindow)$ is variable maintained by each node.
- Using a different retransmission strategy that does not double the CW value after collision.

Such malicious misbehavior can seriously degrade the throughput of well-behaved devices [2][3]. In this paper, we propose modification to IEEE 802.15.4, for simplifying the prevention of such misbehaving devices. This paper is organized as follows. Section 2 includes characters of IEEE 802.15.4 MAC specification for channel access. Section 3 describes the proposed mechanism for prevention against attacks. Also, we describe how to implement the proposed mechanism with the existing IEEE 802.15.4 protocol. We can use reserved bits in the frame without violating the IEEE 802.15.4 MAC frame format. This means the proposed mechanism does not modify the frame structure and is compatible with legacy devices. Performance results of proposed mechanism are presented in section 4. Finally, we give some concluding remarks.

2 Media Access of IEEE 802.15.4

In this section we explain some defined functions of IEEE 802.15.4 MAC specification for channel access. In this specification, the device uses CSMA-CA for resolving contention among multiple nodes accessing the channel. A device (sender) with data to transmit on the channel selects a random backoff value from range $[0, 2^{BE}\text{-}1]$. BE is the backoff exponent, which is related to how many backoff periods a device shall wait before attempting to assess a channel. Each time device wishes to transmit data frames during the CAP (Contention Access Period), it shall locate the boundary of the next backoff slot and then wait for a random number of backoff slots. If the channel is busy, following this random backoff, the device shall wait for another random number of backoff slots before trying to access the channel again. If the channel is idle, the device can transmit on the next available backoff slot boundary. Acknowledgment and beacon frames shall be sent without using a CSMA-CA mechanism [1]. Also, the MAC sub-layer needs a finite amount of time to process data received by the PHY. To allow
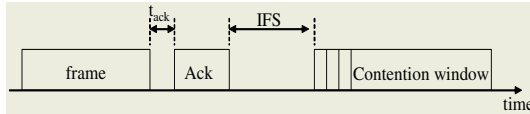
Fig. 1. Interframe Space

this, transmitted frames shall be followed by an IFS period; if the transmission requires an acknowledgment, the IFS shall follow the acknowledgment frame [1][8]. These concepts are illustrated in Figure 1. The IEEE 802.11 specification uses four different interframe spaces. Varying interframe spacings create different priority levels for different type of traffic. The logic behind this is simple: high-priority traffic doesn't have to wait as long after the medium has become idle. Therefore, if there is any high-priority traffic waiting, it grabs the network before low-priority frames have a chance to try. To assist with interoperability between different data rates, the interframe space is a fixed amount of time, independent of the transmission speed [8][9].

3   Media Access Mechanism Using Priority Time Period

The proposed mechanism is designed to require minimal modifications to IEEE 802.15.4 MAC specification. As mentioned above, a malicious device selects backoff values from a different distribution with smaller average backoff value, than the distribution specified by standard (e.g., by selecting backoff values from $[0, \frac{2^{BE}-1}{3}]$, instead of $[0, 2^{BE}-1]$). Such misbehavior of malicious device can seriously degrade the performance of well-behaved devices. Also well-behaved devices should consume the energy to acquire channel. The proposed mechanism is designed to prevent misbehavior of malicious device using the PTP (Priority Time Period) which is decided by coordinator. If well-behaved device repeatedly can't access the channel, device uses the PTP to transmit the frame; this frame should be authenticated by coordinator. The start of PTP (Priority Time Period) can locate in IFS (Interframe Space) as shown in Figure 2. The $time_{start}$ is the time to start PTP after acknowledgment frame and the $length$ is the maintenance duration of PTP; both $time_{start}$ and $length$ are decided by coordinator. The coordinator transmits these values to well-behaved devices using the encryption algorithm to protect these values. Key management between coordinator and devices may be provided by the higher layer, but are out of scope of this paper. Generally well-behaved device uses the CSMA-CA algorithm to acquire the channel. If the value of NB is more than 3, the device can use the PTP. In IEEE 802.15.4 specification, NB is the number of times the CSMA-CA algorithm was required to backoff while attempting the current transmission; this value shall be initialized to 0 before each new transmission attempt and the maximum value is 6. At first, the coordinator decides whether apply the PTP as shown in Figure 3. Priority Time Period is 1 bit in length and shall be set to 1 if coordinator applys the PTP. Otherwise, Priority Time Period subfield shall be set to 0.
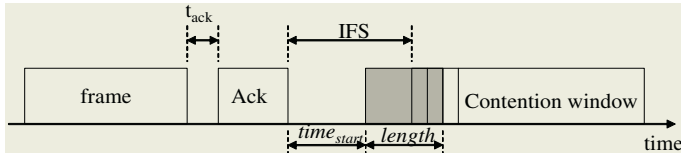
Fig. 2. Proposed Priority Time Period

Fig. 3. Beacon frame format for PTP

Fig. 4. Encrypted information to use PTP

As shown in Figure 4, The accurate information to use PTP (the $time_{start}$ and the $length$) are encrypted by shared key between coordinator and well-behaved devices. These values are included in beacon payload field.

- encrypted beacon payload: $E_{key}\{time_{start}, length\}$

A malicious device can't know these values, because it has not shared key. Beacon enabled network use a slotted CSMA-CA channel access mechanism, where the backoff slots are aligned with the start of the beacon transmission. Each time a device wishes to transmit data frames during the CAP, it shall locate the boundary of the next backoff slot and then wait for a random number of backoff slots. If the channel is busy, following this random backoff, the device shall wait for another number of backoff slots before trying to access the channel again. If the value of NB is more than to 3, the device can use the PTP. As mentioned above, NB is the number of times the CSMA-CA algorithm was required to backoff while attempting the current transmission. Also, the device includes hash value for authentication as shown in Figure 5. The coordinator can authenticate using this hash value whether the device is well-behaved or not. If malicious

Format of the frame control field

Bits: 0-2	3	4	5	6	7-9	10-11	12-13	14-15
Frame type	Security enable	Frame pending	Ack. request	Intra-PAN	Reserved	Dest. Addressing mode	Reserved	Source Addressing mode

$Hash_{key}\{time_{start}\}$

Fig. 5. Hash value for authentication
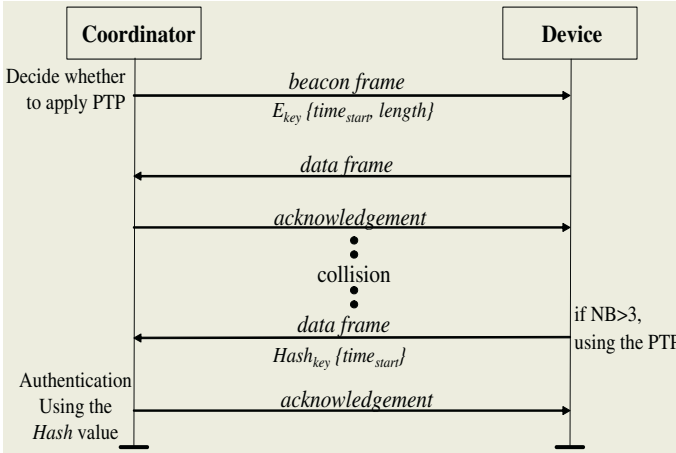
Fig. 6. Process for channel acquirement

device is detected, the coordinator can disconnect with that device. Figure 6 shows an entire process between coordinator and device.

4   Performance Analysis

In this section, simulation results are presented to demonstrate the effectiveness of the proposed mechanism. The performance metrics we use include 'Delivery Ratio of data frames of well-behaved devices', which is the ratio of well-behaved device's frame throughput over their sending rate. All the simulation results are based on a modified version of ns-2 network simulator. The traffic sources are chosen to be constant bit rate (CBR) source using packet-size of 20 bytes. Figure 7(a) shows the delivery ratio of well-behaved devices depending on the backoff range ratio of malicious device (for example, if malicious device select backoff value from $[0, \frac{2^{BE}-1}{4}]$, backoff range ratio is 0.25).

Also, Figure 7(b) shows the delivery ratio of well-behaved devices depending on the number of devices. The simulation results demonstrate that severe contention among frames of devices will cause significant performance degradation without PTP (Priority Time Period).

(a) Delivery ratio vs. backoff period ratio of malicious device   (b) Delivery ratio vs. number of devices
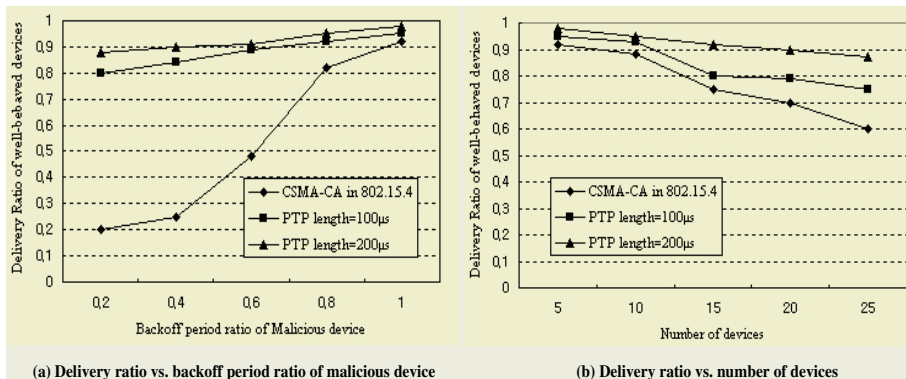
Fig. 7. Delivery Ratio Performance Comparison

5   Conclusion and Future Works

Preventing MAC layer misbehavior is an important requirement in ensuring a reasonable throughput share for well-behaved nodes in the presence of misbehaving devices. In this paper, we have presented modification of IEEE 802.15.4 MAC protocol to use PTP. Also, we have described how to implement the proposed mechanism with the existing IEEE 802.15.4 protocol. We can use reserved bits in the frame without violating the IEEE 802.15.4 MAC frame format. Simulation results indicate that our prevention mechanism is successful in handling MAC layer misbehavior. We plan to foster our solution in the evolutionary computing systems through further development.

References

1. IEEE Standard for Wireless Medium Access Control and Physical Layer Specification for Low-Rate Wireless Personal Area Networks, P802.15.4, 2003.
2. Pradeep Kyasanur, Nitin H. Vaidya, "Detection and handling of MAC layer misbehavior in wireless networks", Dependable Systems and Networks 2003, pp.173 - 182, June 2003.
3. V. Gupta, S. Krishnamurthy, M. Faloutsos, "Denial of service attacks at the MAC layer in wireless ad hoc networks", MILCOM 2002 Proceedings, pp.1118 - 1123, vol.2, Oct 2002.
4. G. Noubir, G. Lin, "Low-Power DoS Attacks in Data Wireless LANs and Countermeasures", ACM MobiHoc, Poster Session, 2003.
5. TS Messerges, J Cukier, et al, "A security design for a general purpose, self-organizing, multihop ad hoc wireless network", Proceedings of the 1st ACM workshop Security of Ad Hoc and Sensor Network, 2003.
6. N. Sastry, D. Wagner, "Security Consideration for IEEE 802.15.4 Networks", WiSe'04 Proceeding, pp.32-42, 2004.
7. Hao Yang, Haiyun Luo, Fan Ye, Songwu Lu, Lixia Zhang, "Security in Mobile Ad Hoc Networks: Challenges and solutions", Wireless Communications, Volume 11, Issue: 1, pp 38-47, Feb 2004.

8. Imad Aad, Claude Castelluccia, "Differentiation mechanisms for IEEE 802.11", IEEE INFOCOM, April 2001.
9. Michael Barry, Andrew T. Campbell, Andras Veres, "Distributed Control Algorithms for Service Differentiation in Wireless Packet Networks", IEEE INFOCOM, April 2001.
10. V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, E. Knightly, "Distributed Multi-Hop Scheduling and Medium Access with Delay and Throughput Constraints", MOBICOM, August 2001.

Pilot-Symbol Aided SFBC-OFDM Channel Estimation for Intelligent Multimedia Service

Sang Soon Park[1], Juphil Cho[2], and Heung Ki Baik[3]

[1] Department of Electronic Engineering, Chonbuk National University, Jeonju, Korea
spwman@chonbuk.ac.kr
[2] School of Electronics and Information Engineering Kunsan National University, Kunsan, Korea
[3] Division of Electronics and Information, Chonbuk National University, Jeonju, Korea

Abstract. Next generation intelligent multimedia communications require high quality voice, graphic, video and data service over any service environment. The MIMO-OFDM systems are well known systems which are attractive for intelligent multimedia service. Among the various MIMO-OFDM systems, the SFBC-OFDM coded in space and frequency domain is a more effective system for mobile environments. In fast time-varying channels, exact channel estimation, tracking and compensation of channel distortion is essential to STBC-OFDM system The conventional channel estimation method with least-squares (LS) algorithm estimates the channel frequency response. In this paper, we propose the pilot-symbol aided channel impulse response estimator for SFBC-OFDM. Our channel estimator is effective in mobile multimedia systems.

1 Introduction

The STBC was applied to the OFDM (Orthogonal Frequency Division Multiplexing) as an attractive solution for a high bit rate data transmission in a multipath fading environment. This system was referred to as the space-time block coded OFDM (STBC-OFDM)[1][2]. And the space-frequency block coded OFDM (SFBC-OFDM) has been proposed where the block codes are formed over the space and frequency domain[3]. The orthogonal condition for the SFBC-OFDM is the constant complex channel gains over successive subcarriers, and that for the STBC-OFDM is the constant complex channel gains over successive OFDM symbols. Therefore, SFBC-OFDM is the more efficient method in fast fading mobile channels.

For more efficient SFBC-OFDM systems, accurate channel state information(CSI) is required at the receiver. The conventional least-squares(LS) method estimates the channel frequency response[4]. Also channel impulse response estimator for MIMO-OFDM was proposed[5]. In OFDM system over fast-varying fading channels, channel estimation and tracking is generally carried out by known pilot symbols in given position of the frequency-time grid. In this paper, we propose the pilot-symbol aided channel impulse response estimator for SFBC-OFDM.

2 Pilot-Symbol Aided SFBC-OFDM Channel Estimation

Fig. 1 shows the SFBC-OFDM transmission scheme with 2 transmit antennas and 1 received antenna. We can expressed the received signal $y_k(n)$ as follows.

$$y_k(n) = \sum_{i=1}^{2} h_k^{(i)}(n) x_k^{(i)}(n) + z_k(n) \tag{1}$$

where the k is subcarrier index and n is OFDM symbol index. Also i and j denote the transmit and receive antenna index respectively. h is channel frequency response.

MIMO-OFDM channel impulse response estimator with the optimal training sequence has been present in [5].
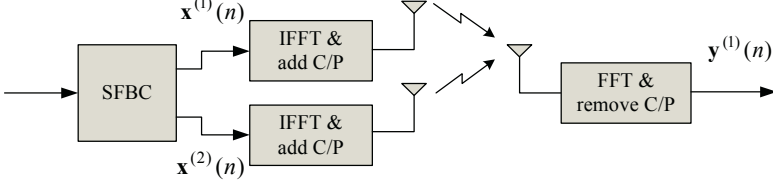
Fig. 1. SFBC-OFDM transmission scheme

2.1 Pilot-Symbol Aided Channel Impulse Response Estimation

Define i'th channel impulse response as follows.

$$\mathbf{g}^{(i)}(n) = \left[g_0^{(i)}(n) \quad g_1^{(i)}(n) \quad \cdots \quad g_{L-1}^{(i)}(n) \right]^T \tag{2}$$

And the pilot-symbol aided channel impulse response estimator $\tilde{g}$ for MIMO-OFDM is acquired by

$$\sum_{i=1}^{2} \sum_{l=0}^{L-1} \tilde{g}_l^{(i)}(n) q_{l_0-l}^{(ij)}(n) = p_{l_0}^{(j)}(n) \tag{3}$$

Where $j = 1, 2$ and $l_0 = 0, 1, \cdots, L-1$. Also define p and q as follows.

$$p_l^{(i)}(n) \equiv \sum_{k_1=0}^{\frac{N}{N_f}-1} \left(y_{N_f k_1+k_0}(n) x_{N_f k_1+k_0}^{(i)*}(n) W_N^{-(N_f k_1+k_0)l} \right)$$

$$q_l^{(ij)}(n) \equiv \sum_{k_1=0}^{\frac{N}{N_f}-1} \left(x_{N_f k_1+k_0}^{(i)}(n) x_{N_f k_1+k_0}^{(j)*}(n) W_N^{-(N_f k_1+k_0)l} \right) \tag{4}$$

Where k_0 is the starting position of pilot symbol and $N/N_f \geq L$. If we define $\mathbf{Q}^{(ij)}(n)$, $\mathbf{p}^{(i)}(n)$ and $\mathbf{g}^{(i)}(n)$ as

$$\mathbf{Q}^{(ij)}(n) \equiv \begin{bmatrix} q_o^{(ij)}(n) & q_{-1}^{(ij)}(n) & \cdots & q_{-L+1}^{(ij)}(n) \\ q_1^{(ij)}(n) & q_0^{(ij)}(n) & \cdots & q_{-L+2}^{(ij)}(n) \\ \vdots & \vdots & \ddots & \vdots \\ q_{L-1}^{(ij)}(n) & q_{L-2}^{(ij)}(n) & \cdots & q_0^{(ij)}(n) \end{bmatrix}$$

$$\mathbf{p}^{(i)}(n) \equiv \left[p_0^{(i)}(n) \quad p_1^{(i)}(n) \quad \cdots \quad p_{L-1}^{(i)}(n) \right]^T \tag{5}$$

$$\tilde{\mathbf{g}}^{(i)}(n) \equiv \left[\tilde{g}_0^{(i)}(n) \quad \tilde{g}_1^{(i)}(n) \quad \cdots \quad \tilde{g}_{L-1}^{(i)}(n) \right]^T$$

and equation (3) can be expressed as following matrix equation.

$$\mathbf{p}^{(1)}(n) = \mathbf{Q}^{(11)}(n)\widetilde{\mathbf{g}}^{(1)}(n) + \mathbf{Q}^{(21)}(n)\widetilde{\mathbf{g}}^{(2)}(n)$$
$$\mathbf{p}^{(2)}(n) = \mathbf{Q}^{(12)}(n)\widetilde{\mathbf{g}}^{(1)}(n) + \mathbf{Q}^{(22)}(n)\widetilde{\mathbf{g}}^{(2)}(n)$$

(6)

Therefore MIMO-OFDM channel impulse estimator is as follows.

$$\widetilde{\mathbf{g}}(n) = \mathbf{Q}^{-1}(n)\mathbf{p}(n)$$

(7)

Where

$$\widetilde{\mathbf{g}}(n) \equiv \begin{bmatrix} \widetilde{\mathbf{g}}^{(1)}(n) \\ \widetilde{\mathbf{g}}^{(2)}(n) \end{bmatrix}, \ \mathbf{p}(n) \equiv \begin{bmatrix} \mathbf{p}^{(1)}(n) \\ \mathbf{p}^{(2)}(n) \end{bmatrix}, \ \mathbf{Q}(n) \equiv \begin{bmatrix} \mathbf{Q}^{(11)}(n) & \mathbf{Q}^{(21)}(n) \\ \mathbf{Q}^{(12)}(n) & \mathbf{Q}^{(22)}(n) \end{bmatrix}$$

(8)
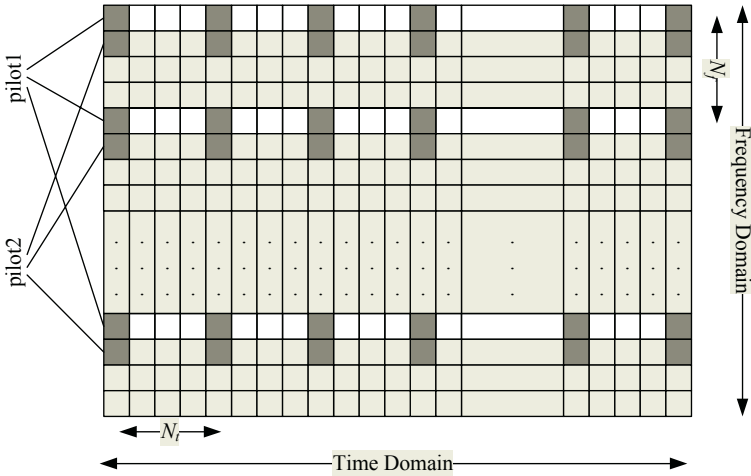
Fig. 2. Pilot symbol position in frequency-time grid

Fig. 2 shows the pilot symbol position for SFBC-OFDM. Therefore, we can estimate the channel impulse responses using the pilot1 and pilot2. Let $\widetilde{\mathbf{g}}_{\text{pilot1}}$ and $\widetilde{\mathbf{g}}_{\text{pilot2}}$ be the respective channel impulse response estimators using the pilot1 and pilot2. Where the starting position k_0 of pilot1 is 0 and the starting position k_0 of pilot2 is 1.

$$\mathbf{p}_{pilot1}(n) = \mathbf{Q}_{pillot1}(n)\widetilde{\mathbf{g}}_{pilot1}(n)$$
$$\mathbf{p}_{pilot2}(n) = \mathbf{Q}_{pillot2}(n)\widetilde{\mathbf{g}}_{pilot2}(n)$$

(9)

where

$$\widetilde{\mathbf{g}}(n) = \widetilde{\mathbf{g}}_{pilot1}(n) = \widetilde{\mathbf{g}}_{pilot2}(n)$$

(10)

Therefore $\widetilde{\mathbf{g}}(n)$ is

$$\widetilde{\mathbf{g}}(n) = (\mathbf{Q}_{pillot1}(n) + \mathbf{Q}_{pilot2}(n))^{-1}(\mathbf{p}_{pilot1}(n) + \mathbf{p}_{pilot2}(n))$$
$$\equiv \overline{\mathbf{Q}}^{-1}(n)\overline{\mathbf{p}}(n)$$

(11)

2.2   Pilot Symbol Design

For low complexity channel estimation, we should design the pilot symbol. At first, define pilot symbol vector as follows

$$
\begin{aligned}
\text{pilot1}^{(i)}(n) &= \begin{bmatrix} c1_0^{(i)} & c1_1^{(i)} & \cdots & c1_{N/N_f-1}^{(i)} \end{bmatrix}^T \\
\text{pilot2}^{(i)}(n) &= \begin{bmatrix} c2_0^{(i)} & c2_1^{(i)} & \cdots & c2_{N/N_f-1}^{(i)} \end{bmatrix}^T
\end{aligned}
\tag{12}
$$

where pilot1$^{(i)}$ and pilot2$^{(i)}$ are the transmit pilot symbol vectors from the antenna i. If we design the pilot vectors as follows.

$$
\begin{aligned}
&\left| c1_{k_1}^{(i)}(n) \right| = 1, \ \left| c2_{k_1}^{(i)}(n) \right| = 1 \\
&c1_{k1}^{(1)}(n) = -c2_{k1}^{(1)}(n) = c1_{k1}^{(2)}(n) = c2_{k1}^{(2)}(n)
\end{aligned}
\tag{13}
$$

And $\overline{\mathbf{Q}}(n)$ is as follows.

$$
\overline{\mathbf{Q}}(n) = \frac{2N}{N_f} \mathbf{I}_{2L}
\tag{14}
$$

where $\mathbf{I}_{2L}$ denote the $2L \times 2L$ identity matrix. Therefore equation (9) is expressed as

$$
\widetilde{\mathbf{g}}(n) = \frac{2N}{N_f} \mathbf{I}_{2L} \overline{\mathbf{p}}(n)
\tag{15}
$$

3   Simulation Results

In this section, we illustrate the results of computer simulation. The COST207 six-ray typical urban (TU) channel power delay profile was used throughout the simulation[6].

The OFDM systems employed 512 subcarriers at a symbol rate of 5×10^6 symbols per second on each subcarrier. Carrier frequency is 2 GHz and bandwidth is 5 MHz. The SFBC-OFDM transmission scheme in this section uses the 2 transmit antennas and 1 receive antenna. And each channel is uncorrelated. Also we use the 1 dimensional linear interpolation.

Fig. 3 shows the simulation results of LS frequency response estimator and impulse response estimator for pilot pattern with QPSK modulation at the velocity = 120km/h (a), 250km/h (b). Fig. 3 shows the BER curves of channel estimators for various pilot patterns. The pilot rate of this systems is 1/16. In this figures, we can see that performance results of proposed impulse response estimator are better than those of frequency response estimator. Also pilot pattern affects the performance results of channel estimators.

Fig. 4 shows the BER curves of LS frequency response estimator and impulse response estimator with 16-QAM modulation at the mobile speed v = 120km/h, 250km/h. Also pilot interval in frequency domain (N_f) = 8, and pilot interval in time domain (N_t) = 4.

In Fig. 4, we can see that performance results of proposed impulse response estimator are better than those of frequency response estimator.
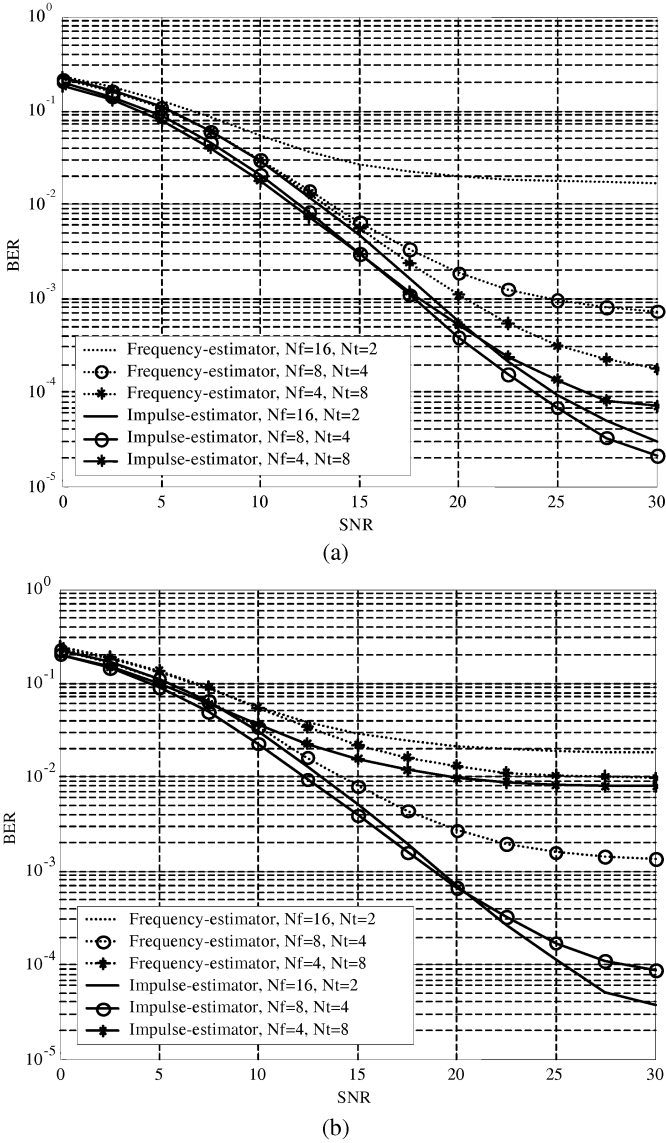
Fig. 3. BER curves of channel estimators for pilot patterns (a) v=120km/h (b) v=250km/h

4   Conclusions

Next generation intelligent multimedia service demands high speed, reliable and spectrally-efficient communications over the mobile wireless environment. The SFBC-OFDM is known as an effective coding scheme for mobile multimedia service.

In this paper, we have proposed pilot-symbol aided channel impulse response estimation for SFBC-OFDM system in high mobile environments. We have illustrated
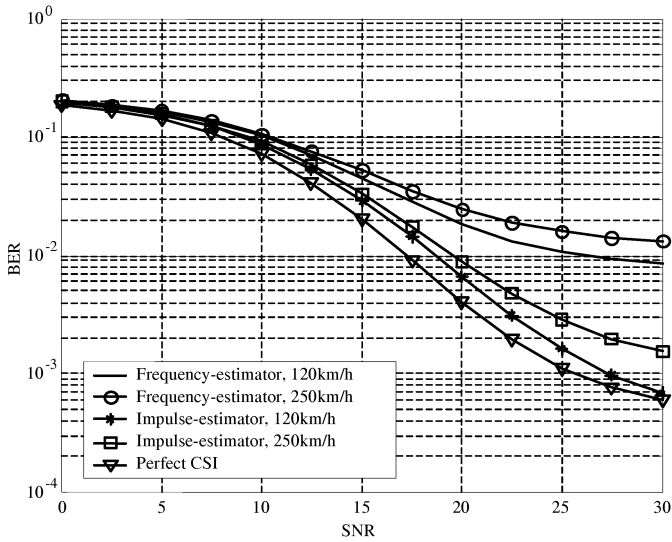
Fig. 4. BER curves of channel estimators for mobile speed

the performance of the LS frequency response estimator and impulse response estimator for mobile broadband wireless access. Through the computer simulations, we have shown that the performance of impulse response estimator is better than that of frequency response estimator in high mobile environments. Also we can lower the computational complexity of channel estimation through the pilot design. Therefore our proposed pilot-symbol aided channel impulse response estimator is an efficient method for high speed mobile multimedia service.

References

1. S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.
2. K. F. Lee and D. B. Williams, "A space-time coded transmitter diversity technique for frequency selective fading channels," *IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 149-152, Cambridge, MA, Mar. 2000.
3. K. F. Lee and D. B. Williams, "A space-frequency transmitter diversity technique for OFDM system," *IEEE GLOBECOM 2000*, vol. 3, pp. 1473-1477, San Francisco, USA, Nov. 2000.
4. W.G. Jeon, K.H. Paik, and Y.S. Cho, "An efficient channel estimation technique for OFDM system with transmitter diversity," *IEICE Transaction on Communications*, vol. E84-B, no. 4, pp. 967-974, April 2001.
5. Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Transactions on Wireless Communications*, vol. 1, no, pp. 67-75, Jan. 2002
6. COST207 TD(86)51-REV 3 (WG1), "Proposal on channel transfer functions to be used in GSM tests late 1986," Sept. 1986.

Cryptographic Protocol Design Concept with Genetic Algorithms

Kyeongmo Park[1] and Chuleui Hong[2]

[1] School of Computer Science and Information Engineering
The Catholic University of Korea
Yeouido P.O. Box 960, Yeongdeungpo-Gu, Seoul, 150-010, Korea
`kpark@catholic.ac.kr`
[2] Software School, Sangmyung University, Seoul, Korea
`hongch@smu.ac.kr`

Abstract. This paper describes design concepts of genetic algorithms in the cryptographic protocols for a fault-tolerant agent replication system. The principles of cryptographic protocols are outlined, followed by the specification of modal logic that is used to encode the belief and knowledge of communicating parties. Replication and voting do not suffice and cryptographic support is required. Secret sharing is a fundamental notion for secure cryptographic design. In a secret sharing protocol a dealer shares a secret among n servers so that subsets of $t + 1$ or more servers can later reconstruct the secret, while subsets of t or less servers have no information about it. This is what it is called a threshold secret sharing scheme. A verifiable secret sharing protocol adds fault tolerance to the above tasks. Our protocol can be used by asynchronous fault-tolerant service replicas.

1 Introduction

Computing over the Internet is not dependable and unreliable. Hosts connected via the Internet constantly fail and recover. The communication links go down at any time. Due to high communication load, link failures, or software bugs, transient communication and node failures are common in the Internet. Information transferred over the Internet is insecure and the security of an agent is not guaranteed. Fault tolerance guarantees the uninterrupted operation of a distributed software system, despite network node failure. Therefore, reliability is an important issue for Internet applications [2], [6], [8]. In this work, we address the security and fault tolerance issues for mobile agent systems running across the Internet. We present a cryptographic method in the replication and voting protocols for the agent replication extension system. The system makes mobile agents fault-tolerant and also detects attacks by malicious hosts.

In a mobile agent based distributed system, agents comprising an application must survive possibly malicious failures of the hosts they visit, and they must be resilient to the potentially hostile actions of other hosts. Correctness of a computation should depend only on hosts that would be visited in a failure-free

run. The replication and voting are necessary to survive malicious behavior by visited hosts. However, faulty hosts that are not visited by agents can confound a naive replica management scheme by spoofing. We have been investigating protocols for replication and voting in a family of applications. This problem can be solved by cryptographic protocols.

The design ideas of cryptographic techniques in our protocols are discussed. In a fault-tolerant mobile agent replication system, proactive secret sharing and threshold signing take on an important role in facilitating mobile processes by distributed authentication. The remainder of this paper is organized as follows. Section 2 describes the system model, a fault-tolerant mobile agent computation with replication and voting, and assumptions. Section 3 describes cryptographic protocols with secret sharing. Finally, in Section 4 we present our conclusions.

2   The System Model

Let us consider a distributed system which is a collection of processors interconnected by a communication network. We will use processors with nodes interchangeably. For each window of vulnerability, there are n processors that are responsible for the service during the period of time. While a main node is responsible for the service, we call it a server. Servers hold the shares of a secret. Each node is assumed to have an individual public/private key pair. Each server is assumed to know the public key of all other nodes in the system. Cryptographic techniques are employed to provide confidentiality and authenticity for passing messages. It is assumed that the adversary is computationally bound; the factoring problem is hard so that the adversary cannot subvert these cryptographic techniques.

Our protocol is intended for use in the Internet environment, where failures and attacks can invalidate assumptions about timing. Thus, we assume an asynchronous system where there is no bound on message delay or processor speed. We also assume that the network by which the nodes are connected is composed of fair links. A fair link is a communication channel between nodes that does not deliver necessarily all messages sent. If a node sends many messages to a single destination, then one of those messages is correctly delivered. Messages in transit may be read or altered by the adversary.

Replication of agents and data at multiple processors is a key to providing fault tolerance in distributed systems. Replication is a technique used widely for enhancing Internet services. The motivations for replication are to make the distributed system fault-tolerant, to increase its availability, and to improve a service's performance.

A replication algorithm proposed in [1] for Byzantine fault tolerance in asynchronous systems offers correctness guarantees provided that no more than 1/3 of servers fails. The amount of time an adversary has to compromise more than 1/3 of the servers is called the window of vulnerability. Less than 1/3 of the servers in a window of vulnerability can be faulty. A faulty server can stop executing the protocol, deviate from its specified protocol in an arbitrary manner,

and corrupt or disclose locally stored information. A correct server follows the protocol and does not corrupt or disclose locally stored information. A server is considered correct in a time interval t if and only if it is correct throughout interval t. Otherwise it is faulty in time t. With the assumptions described, to cause the damage to the replicated service, an adversary is allowed to do the following:

- compromise up to f faulty servers within any window of vulnerability,
- delay messages or correct servers by arbitrarily finite amounts,
- launch eavesdropping, message insertion, corruption, deletion, and replay attacks

To give a precise definition for the window of vulnerability, we assign version numbers $(v_0, v_1, v_2, ...)$ to shares, secret sharing, runs, and participating service replicas. Each run of the protocol generates a new secret sharing. A sharing is composed of the set of random shares which form a secret. Server initially store shares with version number v_0. If a run is executed with shares having version number $v_{old} = v_i$, then this run and the resulting new shares have version number $v_{new} = v_{i+1}$. A secret sharing is assigned the same version number as its shares.
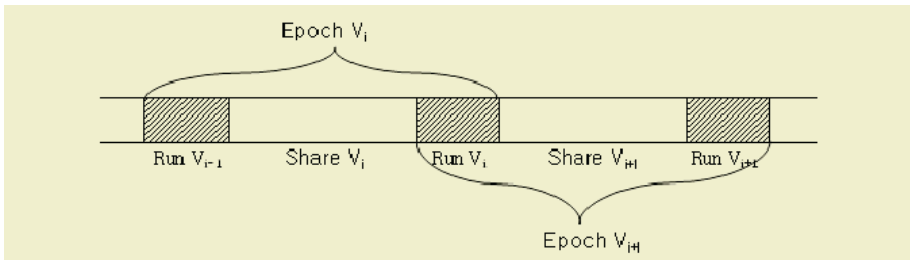
Fig. 1. The relations between old and new runs, shares, and epochs

Run v_{new} starts when some correct server initiates run v_{new} locally. Run v_{new} terminates locally on a correct server once the server has forgotten all information pertinent to v_{old} shares. Run v_{new} terminates globally at the earliest time t that the run has terminated locally on each correct server participated in run v_{new}. An epoch v_i is defined to be the interval from the start of the run v_i to the global termination of run v_{i+1}. The relationship between old and new runs, shares, and epochs, and their surrounding runs, shares, and epoches is illustrated in Fig. 1. In a successful attack, an adversary must collect $f + 1$ or more shares all with the same version number. Since $f + 1$ or more servers must be compromised during the same epoch to collect these shares, we define the window of vulnerability to be an epoch.

One way to share a secret and achieve a threshold access structure is to use the combinatorial secret sharing [4] technique. We split the secret up into m shares, where $m = combination(3f + 1, f)$. A share is given a label from 1 to m. We find all possible combinations of f nodes from our $3f + 1$ nodes, and we

label each group of f nodes with a number from 1 to m. Then, a node n_i gets a share s_i with label x if node n_i is not in a group with x.

For $f = 1, m = 4$ so we have four shares: S_1, S_2, S_3, S_4. We generate all possible groups of $f = 1$ servers and label them with numbers 1 to m, i.e., $g_1 = n_1, g_2 = n_2, g_3 = n_3, g_4 = n_4$. Then we give share s_1 to all nodes not in group g_1. Similarly, we give share s_2 to all nodes not in group g_2, s_3 to all nodes not in g_3, and s_4 to all nodes not in g_4. Now, by construction, we summarize: 1) Each share is held by $2f + 1$ nodes. 2) No set of less than $f + 1$ nodes have all of the shares. 3) Any set of $f + 1$ or more nodes will have all of the shares. For any given share, only f nodes do not have the shares, every other node has it. Then, for each share, given a group of any $f + 1$ of the $3f + 1$ nodes, at least one of the nodes has the share.

One of the goals in this work is to provide fault tolerance to mobile multi-agents through selective agent replication. Multi-agent applications reply on the collaboration among agents. If one of the involved agents fails, the whole computation can get damaged. The solution to this problem is replicating specific agents. One must keep the solution as independent and portable as possible from the underlying agent platform, so as to be still valid even in case of drastic changes of the platform. This offers interoperability between agent systems. The properties of agent systems are dynamic and flexible. This increases the agent's degree of proactivity and reactivity. Note that replication may often be expensive in both computation and communication. A software element of the application may loose at any point in progress. It is important to be able to go back to the previous choices and replicate other elements.

A simple agent computation might visit a succession of hosts, delivering its result messages to an actuator. The difficulties here arise in making such a computation fault-tolerant. The agent computation of interest can be viewed as a pipeline, depicted in the shaded box of Fig. 2. Nodes represent hosts and edges represent movement of an agent from one host to another. Each node corresponds to a *stage* of the pipeline. S is the *source* of the pipeline; A is the *actuator*. The computation is not fault-tolerant. The correctness of a stage depends on the correctness of its predecessor, so a single malicious failure can propagate to the actuator. Even if there are no faulty hosts, some other malicious host could disrupt the computation by sending an agent to the actuator first.
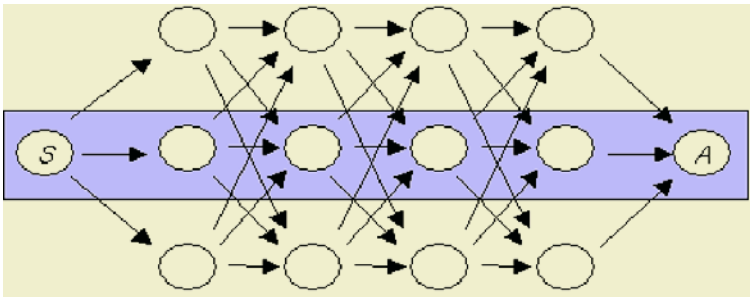
Fig. 2. Fault-tolerant agent computation using replication and voting

One step needed to make fault-tolerant is replication of each stage. We assume that execution of each stage is deterministic, but the components of each stage are not known *a priori* and they depend on results computed at previous stages. A node m in stage k takes as its input the majority of the inputs it receives from the nodes comprising stage $k-1$. And then, m sends its output to all of the nodes that it determines consisting of $k+1$. Fig. 2 illustrates such a fault-tolerant execution. The replicated agent computation with voting tolerates more failures than an architecture where the only voting occurred just before the actuator. The voting at each stage makes it possible for the computation to recover by limiting the impact of a faulty host in one stage on hosts in subsequent stages. It is assumed for every stage, i.e., an execution session on one host, a set of independent replicated hosts, i.e., hosts that offer the same set of resources, but do not share the same interest in attacking a host, because they are operated by different organizations. Every execution step is processed in parallel by all replicated hosts. After execution, the hosts vote about the result of the step. At all hosts of the next step, the votes, the resulting agent states, are collected. The execution with the most votes wins, and the next step is executed.

3   Cryptographic Techniques for Distributed Protocols

The agent computation using replication should tolerate one malicious host per stage of the pipeline. Any two faulty hosts could claim to be in the last stage and foist a majority of bogus agents on the actuator. These problems are avoided if the actuator can detect and ignore such bogus agents. This could be accomplished by having agents carry a privilege from the source to the actuator.

The privilege can be encoded as a secret initially known to the source and the actuator. It is necessary to defend against two attacks. First, a malicious host might misuse the secret and launch an arbitrary agent to the actuator. Second, a malicious host could destroy the secret, making it impossible for the remaining correct agents to deliver a result.

To simplify the discussion, we start with a scheme that prevents misuse of the secret by a malicious host. This scheme ensures that if a majority of replicas visit only correct hosts, no hosts except the source and the actuator will learn the secret. Agent replicas cannot carry copies of the secret, since the secret could then be stolen by any faulty host visited by a replica. It is tempting to circumvent this problem by the use of an *(n,k)* threshold secret sharing [7] scheme to share the secret embodying the privilege. In an *(n,k)* threshold scheme, a secret is divided into n fragments, where possession of any k fragments will reveal the secret, but possession of fewer fragments reveals nothing. In a system having *2k-1* replicas, the source would create fragments using a *(2k-1, k)* threshold scheme, and send a different fragment to each of the hosts in the first stage. Each of these hosts would then forward its fragment to a different host in the next stage.

However, this protocol fails, due to the voting at intermediate stage of the pipeline. The voting should ensure that the faulty hosts encountered before a vote cannot combine with faulty hosts encountered after the vote to corrupt the

computation. However, in the scheme, minorities before and after the vote can steal different subsets of the secret fragments. If they together hold a majority of the secret fragments, then the faulty hosts can collude to reconstruct the secret. One way to stop collision between hosts separated by a vote is to redivide the secret fragments at each vote. The outline of a protocol for a system with *(2k-1)* replicas is, as follows.

- The source divides the secret into *2k-1* fragments and sends each fragment to one of the hosts of the first stage.
- A host in stage i takes the fragments it receives, concatenates them, and divides that into *2k-1* fragments using a *(2k-1, k)* threshold scheme. Each fragment is then sent to a different host in stage *i+1*.
- The actuator uses the threshold scheme backwards and recovers the original secret.

This protocol is inefficient because secret sharing scheme require that each fragment be the same size as the original secret. Thus, messages get longer at every stage of the pipeline. In fact, the message size is multiplied by at every stage.

Two protocols have been developed. They do not suffer from the exponential blowup in message size. In one protocol, message size grows linearly with the number of pipeline stages and the number of replicas. In the other one, message size remains constant but an initialization phase is required. The first scheme uses chains of authentication, instead of a secret privilege, to prevent masquerading. The second scheme renews [3], [5] the secret after each round, making it impossible for fragments from before a vote to be used together with fragments constructed after that vote. To address the second attack - destruction of the secret by a faulty host - it is necessary to replace the secret sharing in the protocols described with verifiable secret sharing. This scheme allows for correct reconstruction of a secret even when hosts including the source are faulty.

4   Conclusion

In this paper, we describes design concepts of genetic algorithms in the cryptographic protocols for a fault-tolerant agent replication system. The principles of cryptographic protocols are outlined, followed by the specification of modal logic that is used to encode the belief and knowledge of communicating parties. Replication and voting do not suffice and cryptographic support is required. Secret sharing has been employed in the face of Byzantine faults. In a secret sharing protocol, a dealer shares a secret among n servers so that subsets of $t+1$ or more servers can later reconstruct the secret, while subsets of t or less servers have no information about it. This is what it is called a threshold secret sharing scheme. A verifiable secret sharing protocol adds fault tolerance to the above tasks. Our protocol can be used by asynchronous Byzantine fault tolerant service replicas to perform secret sharing.

References

1. M. Castro and B. Liskov: Practical Byzantine Fault Tolerance. In *OSDI: Symp. on Operating Systems Design and Implementation*, USENIX Association, 1999.
2. P. Dasgupta: A Fault Tolerance Mechanism for MAgNET: A Mobile Agent based E-Commerce System. In *Proc. Sixth Int. Conf. Internet Computing*, Las Vegas, NV, June 24-27, 2002, 733–739.
3. A. Herzberg, S. Jarecki, H. Krawczyk, and M. Yung: Proactive secret sharing or: How to cope with perpetual leakage. In *Proc. 15th Int. Conf. on Advances in Cryptology*, Springer, 1995, 339–352.
4. M. Ito, A. Saito, and T. Nishizeki: Secret sharing scheme realizing general access structure. In *IEEE Globecom*, 1987, 99–102.
5. S. Jarecki: Proactive Secret Sharing and Public Key Cryptosystems. Master's thesis, MIT, Cambridge, MA, Sep. 1995.
6. S. Mishra et al.: DaAgent: A Dependable Mobile Agent System. In *Proc. of the 29th IEEE Int. Symp. Fault-Tolerant Computing*, Madison, WI, June 1999.
7. A. Shamir: How to share a secret. *Communications of the ACM* 22, Nov. 1979, 612–613.
8. M. Strasser and K. Rothermel: Reliability Concepts for Mobile Agent. *Int. Journal Cooperative Information Systems*, 7(4): Dec. 1998, 355–382.
9. Z. Michalewicz: Genetic Algorithms + Data Structures = Evolution Programs. 3rd Ed., Springer-Verlag, 1996.

An Efficient Backoff Scheme
for IEEE 802.11e EDCF Differential Service

Ho-Jin Shin, Jang-Sub Kim, and Dong-Ryeol Shin

School of Information and Communication Engineering, Sungkyunkwan University
300 cheoncheon-dong, jangan-gu, suwon, Korea
{hjshin,jangsub,drshin}@ece.skku.ac.kr

Abstract. Wireless LANs (WLANs) are one of the fastest growing wireless access technologies. Unfortunately, since they were developed to closely match existing wired LANs, the popular IEEE 802.11 standards have several problems in providing Quality of Service (QoS) to stations. The IEEE 802.11e standard is being developed to overcome these drawbacks. One of most important function in 802.11e is the contention-based channel access mechanism called enhanced distributed coordination function (EDCF), which provides a priority scheme by differentiating the inter-frame space and the initial window size. In this paper, we first consider the enhanced DCF access method of IEEE 802.11e. We then introduced analytical model, which can be used to calculate the traffic priority and throughput corresponding to the configuration of multiple DCF contention parameters under saturation throughput. This enhanced DCF access method differs from the standard backoff schemes in that the backoff time is not uniformly chosen in the contention window interval that results in reduced collision probability under high loads. Numerical analysis shows that saturation throughput of the proposed scheme outperform the earlier approach.

1 Introduction

The IEEE 802.11 enables fast installation, with minimum management and maintenance costs, and is a very robust protocol for the best-effort service in the wireless medium. As usage and deployment of wireless LANs (WLANs) increase, it is reasonable to anticipate that the demands to be able to run time-bounded applications on them will be comparable as on wired networks. The IEEE 802.11 protocol was developed as a standard for WLANs [1]. Nowadays, any network should aim to provide QoS capabilities to applications. QoS-sensitive services, such as VoIP and video streaming, are expected to be delivered via WLAN. However, bandwidth is a scare resource in a wireless environment. Therefore, an efficient medium access control (MAC) protocol is indispensable for a WLAN to effectively manage the channel access and provide differentiated QoS to different applications. Thus, the idea of providing QoS guarantees over the WLANs has been the subject of intense study recently.

Unfortunately, the current 802.11 MAC does not posses any effective service differentiation capability, because it treats all the upper layer traffic in the same fashion. Hence, a special working group, IEEE 802.11e, was established to enhance the 802.11 MAC to meet QoS requirements for a wide variety of applications.

We propose a new scheme deciding the contention window size and procedure decrementing the backoff counter, which gives preferences to the active nodes over colliding nodes to reduce the collision probability and thereby improving the throughput. This is the main contribution of the paper, which is to be detailed later.

This paper investigates the performance of the new proposed scheme in terms of throughput. Then comparing with other models, we demonstrate the effect of the proposed mechanism on network performance.

The rest of the paper is organized as follows. Section 2 investigates tunable parameters in the IEEE 802.11e EDCF access method. Section 3 describes the proposed EDCF analytical model and compares it with earlier models. Section 4 describes the numerical results. Finally, we conclude the paper in section 5.

2   IEEE 802.11e EDCF

The IEEE 802.11e provides applications with differential services and QoS support by priority-based contention service, i.e. EDCF. In EDCF, each station can have multiple queues that buffer packets of different priorities. Each frame from the upper layers bears a priority value, which is passed down to the MAC layer. Up to eight priorities are supported in the IEEE 802.11e station and they are mapped into four different access categories (AC) at the MAC layer according to Table 1 [2].

Table 1. Mapping between priories and ACs

Priority	Access Category(AC)	Designation(Infomative)
1	0	Best Effort
2	0	Best Effort
0	0	Best Effort
3	1	Video Probe
4	2	Video
5	2	Video
6	3	Voice
7	3	Voice

A set of EDCF parameters, namely the arbitration interframe space (AIFS[AC]), minimum contention window size (CW$_{min}$[AC]), maximum contention window size(CW$_{max}$[AC]), and the backoff *persistence factor* (PF), is associated with each access category to differentiate the channel access. The lower AIFS[AC] / CW$_{min}$ / CW$_{max}$ results in the higher probability of winning the channel contention. In EDCF, the contention window is expanded by the PF after collision. In original DCF, the contention window is always double after collision (PF=2), while in EDCF PF may be a different value. AIFS[AC] is the number of time slots a packet of a given AC has to wait after the end of a time interval equal to a short interframe space (SIFS) duration before it can start the backoff process or transmit. After i ($i \geq 0$) collisions, the backoff counter in IEEE 802.11e is selected uniformly from [1, 2^i*CW$_{min}$[AC]], until it reaches the backoff stage *i* such as that 2^i*CW$_{min}$[AC] = CW$_{max}$[AC]. At that point, the packet will still be retransmitted, if a collision occurs, until the total number of retransmissions equals the maximum number of allowable retransmissions (*RetryLimit*[AC]) specified in [2], with backoff counters always chosen from chosen from the range [1, CW$_{max}$[AC]].

Interested readers can refer to [3] [4] [5] for a more thorough explanation on how the EDCF mode operate.

3   EDCF Analytical Model Using Proposed Scheme

In our analysis, we assume that the IEEE 802.11e network under investigation is heavily loaded. This implies that there is always at least one packet awaiting transmission at each AC queue within a station. The WLAN operates in an ideal physical environment, meaning that frame errors, the hidden terminal effect or the capture effect will not be modeled in our Markov model. In addition, the Markov model in [6] does not consider the frame retries limits, and thus it may overestimate the throughput of IEEE 802.11. With this observation, our analysis is based on a more exact Markov chain model [7] taking into the consideration of the maximum frame retransmission limit. We assume a fixed number of wireless stations, each always having a packet available for transmission. In other words, the transmission queue of each station is assumed to be always nonempty.

Assume there are total T priority access categories (AC), and for each AC, the number of traffic flows is $N_j, j = 1, 2 \cdots$ T. The total number of traffic flow is

$$N = \sum_{j=1}^{T} N_j \tag{1}$$

The analysis is divided into two parts. First, we examine the behavior of a single station with a Markov model, and obtain the traffic priority τ_n (i.e. the stationary probability that the station transmits a packet in a generic slot time). This probability does not depend on the access mechanism employed as shown in [6]. Then, we express the AC throughput equation by studying the events that can occur in a generic slot time.

3.1   AC Priority Analysis

We modify the Markov chain model described in [6] [7] [9] by introducing differential service parameters as shown in Fig. 1. The nth traffic flow is associated with $AIFS_n$, $CW_{n,min}$, $CW_{n,max}$, pf_n and all other parameters that have a subscript n. Moreover, for the simplicity, the range of n is [1, N] for all equations in Section 4 unless specific note. Define for convenience $f_n=pf_n$, $W_n=CW_{n,min}+1$, $W_{n,I} = (f_n)^i W_n$, and $W_{n,m'} = (f_n)^{m'} W_n=CW_{n,max}+1$.

We have

$$\begin{cases} W_{n,i} = (f_n)^i W & i \leq m' \\ W_{n,i} = (f_n)^{m'} W & m' < i \leq m \end{cases} \tag{2}$$

where m+1 is the retry limit of the packet. So far in IEEE 802.11 short retry limit is 7, while the long retry limit is 4.

Each traffic flow has a conditional collision probability pn (the probability of collision seen by a packet being transmitted on channel) and the backoff state transition rate λ_n. Given a specific backoff window size $W_{n,i}$, the average backoff timer $\overline{bk_{n,i}} = (W_{n,i}/2)$ in slot time. The average backoff window size [8] of a MAC frame is calculated as

$$\overline{W_n} = (1-p_n)\frac{W_{n,0}}{2} + p_n(1-p_n)\frac{W_{n,1}}{2} + \cdots + p_n^{m-1}(1-p_n)\frac{W_{n,m-1}}{2} + p_n^m(1-p_n)\frac{W_{n,m}}{2} \tag{3}$$

The overall average backoff timer of a packet in n^{th} traffic flow is

$$\overline{bk_n} = (\overline{W_n} / 2) \times slotTime \tag{4}$$

Without the loss of generality, let us assume that from the first to the last traffic flow, we have the highest to the lowest priority, i.e. $AIFS_1 = AIFS_{min}$. We define $\delta_{n,aifs} = AIFS_n - AIFS_{min}$. In saturation condition, the transmission queues of the highest priority traffic flows will be paused/resumed repeatedly when it decreases. Therefore, we calculated λ_n as follows

$$\lambda_n = \begin{cases} \dfrac{(\overline{bk_1} - \delta_{n,aifs})}{\overline{bk_n}} & \overline{bk_1} > \delta_{n,aifs} \\ 0 & otherwise \end{cases} \tag{5}$$

In above equation, if $\overline{bk_1} \leq \delta_{n,aifs}$, which means the traffic flow with the highest priority always starts transmitting before the deferent of $AIFS_n$ is completed. Therefore, the backoff timer of the nth traffic flow is always in the pause state, i.e. the nth traffic flow will not be able to win the channel contention.

T different types of traffic flows are considered with N_j traffic flows for traffic type j (j= 1,2, $\cdots T$). In saturation condition, each station has a packet immediately available for transmission after the completion of each successful transmission. Moreover, being all packets consecutive, each packet needs to wait for a random backoff time before transmission.

Let $b(t)$ be the stochastic process representing the backoff time counter for a given station at slot time t. A discrete and integer time scale is adopted: t and $t+1$ correspond to the beginning of two consecutive slot times and the backoff time counter of each station decrements at the beginning of each slot time [5]. Let $s(t)$ be the stochastic process representing the backoff stage (0, … , m) of the station at slot time t. As in [5][7], the key conjecture in this model is that the probability p that a transmitted packet collides is independent on the stage $s(t)$ of the station. Thus, the bi-dimensional process $\{s(t), b(t)\}$ is a discrete-time Markov chain, which is shown in Fig. 1.

We thus propose new backoff time counter selection scheme. Unlike the schemes in [1][5][7], we give preferences to the stations actively competing with the colliding stations within current contention window for it, the contention window is constructed non-uniformly as follows:

$$CW_{n,i} = [R_{n,i}, W_{n,i} - 1] \qquad 0 \leq i \leq m$$

Here, if a station transmits a packet successfully in backoff stage-0, $R_0 = 0$. Otherwise, there is a busy channel to incur collision and thus sets a contention window to $[W_{n,i}/2, W_{n,i} - 1]$ rather than $[0, W_{n,i} - 1]$. In other words,

$$R_{n,i} = W_{n,i}/2 \qquad 1 \leq i \leq m$$

Comparing with the example above, the proposed scheme adapts a different mechanism for selecting backoff time counter value. The station then moves to the next backoff stage, generating backoff time counter from CW_1 (32~63).

This approach will reduce the probability of each station choosing the same backoff time counter value. Fig. 2 shows exponential increment of CW for the proposed backoff time counter scheme (PF = 2).

In order to validate and analyze the proposed scheme, we employ the same Markov chain models and assumptions in [7].
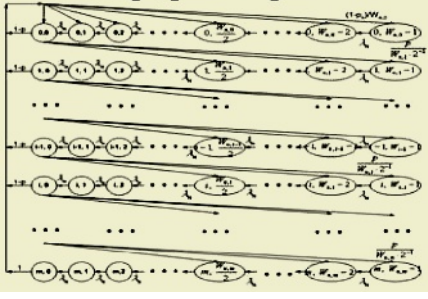
Fig. 1. The Markov Chain model
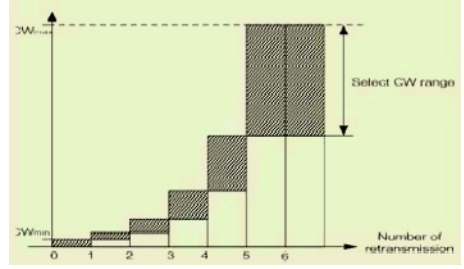
Fig. 2. The backoff counter range in proposed scheme

In the Markov chain of the nth traffic flow, the only non-null one-step transition probabilities are

$$
\begin{cases}
P\{i,k \mid i,k+1\} = \lambda_n & k \in [0,\ W_{n,i}-2]\ i \in [0,\ m] \\
P\{0,k \mid i,0\} = (1-p_n)/W_{n,0} & k \in [0,\ W_{n,0}-1]\ i \in [0,\ m] \\
P\{i,k \mid i-1,0\} = p_n /(W_{n,i}/2) & k \in [W_{n,i}/2,\ W_{n,i}-1]\ i \in [1,\ m] \\
P\{0,k \mid m,0\} = 1/W_{n,0} & k \in [0,\ W_{n,0}-1]
\end{cases}
\tag{6}
$$

The first equation in (6) accounts for the fact that, at the beginning of each slot time, the backoff time is decremented. The second equation accounts for the fact that a new packet following a successful packet transmission starts with backoff stage 0. In particular, as considered in the third equation of (6), when an unsuccessful transmission occurs at backoff stage (i-1), the backoff stage number increases, and the new initial backoff value is randomly chosen in the range $(W_{n,i}/2,\ W_{n,i}-1)$. This is different part from others in [5][6][7]. Finally, the fourth case models the fact that when the backoff stage reaches the maximum backoff stage, the contention window is reset if the transmission is unsuccessful or restart the backoff stage for new packet if the transmission is successful.

For the simplicity, we use $b_{i,k}$ instead of $b_{n,i,k}$ to be the stationary distribution of the Markov chain as $b_{i,k} = \lim_{t \to \infty} P\{s(t)=i, b(t)=k\}$. Using chain regularities in steady state, we can derive the following relations.

$$
b_{i-1,0} * p_n = b_{i,0} \quad 0 < i \le m
\tag{7}
$$

$$
b_{i,0} = p^i b_{0,0} \quad 0 \le i \le m
\tag{8}
$$

If a backoff time value k of a station at stage-i is selected among $CW_{n,i} = [W_{n,i}/2,\ W_{n,i}-1]$, the following relation can be derived.

$$
\begin{cases}
b_{i,k} = \dfrac{W_{n,i}-k}{W_{n,i}\lambda_n}(1-p_n)\sum_{j=0}^{m-1} b_{j,0} + b_{m,0} & i=0,\ k \in (0,\ W_{n,i}-1) \\[3mm]
b_{i,k} = \dfrac{p_n}{\lambda_n} b_{i-1,0} & 0 < i \le m,\ k \in (0,\ \dfrac{W_{n,i}}{2}-1) \\[3mm]
b_{i,k} = \dfrac{W_{n,i}-k}{(W_{n,i}/2)\lambda_n} p_n b_{i-1,0} & 0 < i \le m,\ k \in (\dfrac{W_{n,i}}{2},\ W_{n,i}-1)
\end{cases}
\tag{9}
$$

With (8) and transitions in the chain, equation (9) can be simplified as

$$b_{i,k} = \gamma \, b_{i,0} \qquad 0 \le i \le m \tag{10}$$

By using the normalization condition for stationary distribution, we have

$$
1 = \sum_{i=0}^{m} \sum_{k=0}^{W_{n,i}-1} b_{i,k} = \sum_{i=0}^{m} \sum_{k=0}^{W_{n,i}-1} \gamma b_{i,0} = \sum_{i=0}^{m} b_{i,0} \sum_{k=0}^{W_{n,i}-1} \gamma
$$
$$
= b_{0,0} \sum_{k=0}^{W_{n,i}-1} \gamma + \sum_{i=1}^{m} b_{i,0} \sum_{k=0}^{W_{n,i}-1} \gamma = b_{0,0} \cdot \frac{W_{n,0}+1}{2\lambda_n} + \sum_{i=1}^{m} b_{i,0} \sum_{k=0}^{W_{n,i}-1} \gamma \tag{11}
$$

With all the AC parameter known, a general function of $b_{n,0,0}$ with respect to p_n for a specific traffic flow can be found in (13). Let τ_n be the probability that the station transmits a nth traffic flow during a generic slot time. A transmission occurs when the backoff window is equal to zero. It is expressed as:

$$\tau_n = \sum_{i=0}^{m} b_{n,i,0} = \frac{1 - p_n^{m+1}}{1 - p_n} b_{n,0,0} \tag{12}$$

Then, using equation (2), (10) and (11), $b_{n,0,0}$ is computed as

$$
b_{n,0,0} = \begin{cases} \dfrac{4\lambda_n (1 - f_n p_n)(1 - p_n)}{\alpha} & m \le m' \\[2ex] \dfrac{4\lambda_n (1 - f_n p_n)(1 - p_n)}{\beta} & m > m' \end{cases}
$$

$$\alpha = 2(W_n + 1)(1 - f_n p_n)(1 - p_n) + 3W_n f_n p_n (1 - (f_n p_n)^m)(1 - p_n) \tag{13}$$
$$+ p_n (1 - p_n^{m})(1 - f_n p_n)$$

$$\beta = 2(W_n + 1)(1 - f_n p_n)(1 - p_n) + 3W_n f_n p_n (1 - (f_n p_n)^{m'})(1 - p_n)$$
$$+ p_n (1 - p_n^{m})(1 - f_n p_n) + 3W_n f_n^{m'} p_n^{m'+1} (1 - p_n^{m-m'})(1 - f_n p_n)$$

In the stationary state, a station transmits a packet of nth traffic flow with probability τ_n, so we have packet collision probability p_n as

$$p_n = 1 - (1 - \tau_n)^{j_i - 1} \prod_{o=1, o \ne i}^{T} (1 - \tau_o)^{j_o} \tag{14}$$

Therefore, equation (12), (13) and (14) represent a nonlinear system with two unknowns τ_n and p_n, which can be solved using numerical method. We must have $\tau_n \in (0,1)$ and $p_n \in (0,1)$.

Since the Markov chain of Fig. 1 is a little different from that of [5] [6] and [7], the result obtained for $b_{n,0,0}$ is different from that in [5] [6] and [7]. Similarly, different τ_n and p_n.

3.2 AC Throughput Analysis

Once τ_n is known, the probability $P_{n,tr}$ that there is at least one transmission in the considered slot time, and the probability $P_{n,s}$ that there is one and only one transmission occurring on the channel can be given as

$$P_{n,tr} = 1 - \prod_{o=1}^{T} (1 - \tau_o)^{j_o} \tag{15}$$

$$P_{s,t} = \frac{\sum_{o=1}^{T} \left\{ n_o \tau_o (1 - \tau_o)^{j_o - 1} \bullet \prod_{k=1, k \neq o}^{T} (1 - \tau_k)^{j_k} \right\}}{P_{n,tr}} \tag{16}$$

Moreover, we difined $P_{n,str}$ as the probability that there is one and only one transmission of a traffic flow with type n occurring on the channel, and we have

$$P_{n,str} = n_o \tau_o (1 - \tau_o)^{j_o - 1} \bullet \prod_{k=1, k \neq o}^{T} (1 - \tau_k)^{j_k} \tag{17}$$

We are finally in the position to determine the normalized system throughput, S, defined as the fraction of time the channel is used to successfully transmit payload bits, and express it as a fraction of a slot time

$$S = \frac{E[\text{payload information transmitted in a slot time}]}{E[\text{length of a slot time}]}$$

$$= \frac{P_{n,s} P_{n,tr} E[P_n]}{(1 - P_{n,tr})\sigma + P_{n,s} P_{n,tr} T_{n,s} + (1 - P_{n,s}) P_{n,tr} T_{n,c}} \tag{18}$$

Where $E[P_n]$ is the average packet length, $T_{n,s}$ is the average time the channel is sensed busy because of a successful transmission of a packet of a nth traffic flow, $T_{n,c}$ is the average time the channel is sensed busy by each station during a collision, and σ is the duration of an empty slot time. The times $E[P_n]$, $T_{n,s}$ and $T_{n,c}$ must be measured in slot times.

Let packet header be $H = PHY_{hdr} + MAC_{hdr}$ and propagation delay be δ. Then we must have the following expression for ACK timeout effect, which are same as in [5].

$$\begin{cases} T_{n,s}^{bas} = AIFS_n + H + E[P_n] + \delta + SIFS + ACK + \delta \\ T_{n,c}^{bas} = AIFS_n + H + E[P^*] + SIFS + ACK \end{cases} \tag{19}$$

Where *bas* means basic access method and $E[P^*]$ is the average length of the longest packet payload involved in a collision.

For the RTS/CTS access method,

$$\begin{cases} T_{n,s}^{rts} = AIFS_n + RTS + SIFS + \delta + CTS + SIFS + \delta \\ \qquad + H + E[P_n] + \delta + SIFS + ACK + \delta \\ T_{n,c}^{rts} = AIFS_n + RTS + SIFS + CTS \end{cases} \tag{20}$$

We suppose collision occurs only between RTS frames and consider only the CTS timeout effect.

4   Numerical Results

In this section, we present numerical results according to our analysis model and compare earlier model [5]. In our examples, we validate the analytical model of EDCF with by a heterogeneous traffic scenario with three traffic categories, i.e. voice, video and data, in one BSA. We assume that the number of data, voice and video traffics are integrated with a fixed ratio of 2:1:1. Different model are used to generate three types of traffic same as [5].

- Non-time-critical data traffic
- Voice traffic
- Video traffic

In order to better approximate the saturation condition, the inter-arrival time of above traffic model has been decreased when the number of traffic flows is low.

The parameters of EDCF access method used in analytical model is shown in Table 2.

We choose a PHY with 2Mbps capacity DSSS DQPSK. The PHY overhead is always transmitted at 1Mbps DBPSK [5][10]. The fragmentation threshold is chosen as the default value of 2312bytes. By setting the RTS threshold at 0 byte, the RTS/CTS access method is always used. We retain the persistence factor of all three ACs as the default value of 2 because of simple random number generation. The backoff window sizes are various for three ACs; while m'=4, m=6 for all three ACs.

Table 2. IEEE 802.11 parameters of analytical model

Slot Time	$20\mu s$
SIFS	$10\mu s$
AIFS(Voice/Video/Data)	$50/100/150\mu s$
Max Propagation Delay	$2\mu s$
RTS Retry Limit (long/short)	7/4
CW_min(Voice/Video/Data)	15/31/63 slot
CW_max(Voice/Video/Data)	255/511/1023 slot
Persistence Factor(Voice/Video/Data)	2/2/2
RTS_threshold	0 bytes
Fragmentation threshold	2312 bytes
DSSS PreambleLength	144 bits
DSSS PLCPHeaderLength	48 bits
PLCP Transmission Rate	1 Mbps(DBPSK)
PPDU Transmission Rate	2 Mbps(DQPSK)

Fig. 3(a) shows the numerical result of τ_n, the probability of transmitting a packet in any slot time of a specific ACs, i.e. the priority of AC. Given the particular configuration of AC parameters, the τ_n of voice AC is approximately 10 times higher than that of video AC. Similarly, the τ_n of video traffic is is approximately 10 times higher than that of data traffic. One can get different set of τ_n's of ACs by tuning the AC parameters such as AIFS[AC] and CW_{min}, etc.

Fig. 3(b) shows the value of conditional collision probability pn, i.e. the probability of collision seen by a packet being transmitted on channel. The highest priority AC (voice traffic) has less conditional probability of collision than the lower priority ACs (video and data traffic) in the saturation condition. From comparisons, the propose scheme is verified to reduce the chance of collision in heavy load.

Table 3. shows the numerical values of $T_{n,s}$ and $T_{n,c}$ for RTS/CTS access method computing according to (19) and traffic model described at the beginning of this section. Since we are studying the saturation condition, using RTS/CTS method will effectively decrease the time wasted in collisions.

(a) Probability of transmission   (b) Conditional probability of collision
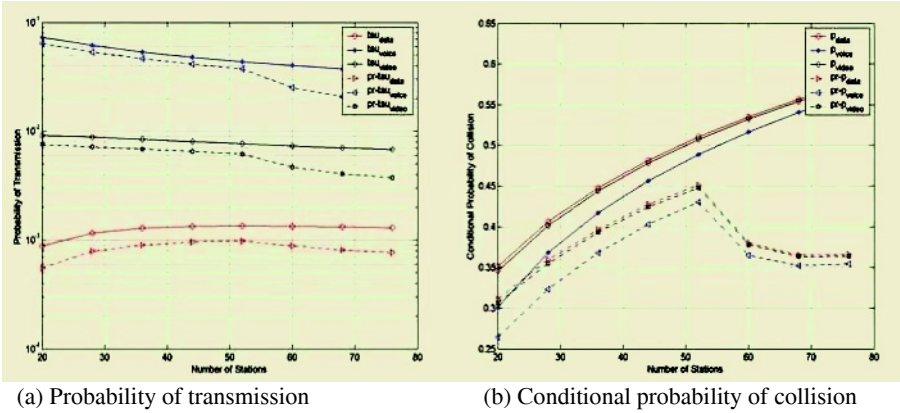
Fig. 3. EDCF AC Priority parameters

In Fig. 4, we compare result of normalized throughput in saturation condition with earlier model [5] and proposed scheme. The proposed shame allows higher throughput than [5]. The improvement in the throughput is due to reduced collision probability between active stations and collided stations by using the proposed backoff time counter scheme.

Table 3. $T_{n,s}$ and $T_{n,c}$ measured in μs

	Voice	Video	Data
Avg. Packet Payload [bits]	1312	13179	8192
$T_{n,s}$ (rts) [μs]	1836	7819	5376
$T_{n,c}$ (rts) [μs]	580	630	680
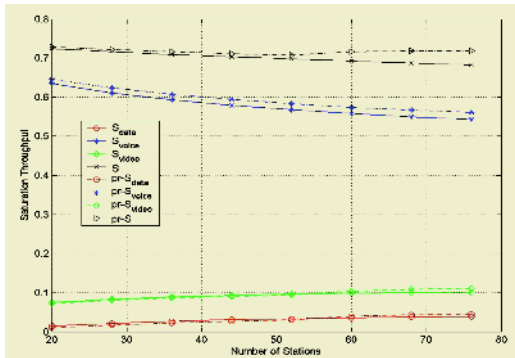
Fig. 4. EDCF nomalized saturation throughput (RTS/CTS access method)

5 Conclusions

This paper has proposed a new backoff scheme for IEEE802.11e EDCF differential services, analyzed the saturation throughput, and based on a bi-dimensional discrete

Markov chain. The analytical results showed that the proposed scheme allows higher saturation throughput and lower collision probability compared to the legacy mechanism [5] at high traffic conditions. Comparing with the model in [5], even though the proposed scheme shows better performance, fairness remains to be solved. The reason for unfairness in the new scheme is due to the fact that the selection is not made within the full range of the congestion window. Future works will be focused on the enhancement of the proposed scheme and the corresponding analytical model by guaranteeing the fairness.

References

1. IEEE std. 802.11-1999 Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, ISO/IEC 8802-11: 1999 Edition, Aug. 1999
2. IEEE 802.11e/D4.0, Draft Supplement to Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), November 2002
3. S. Choi, J. Prado, S. Shankar, and S. Mangold, " IEEE 802.11e Contention-Based Channel Access (EDCF) Performance Evaluation", Proc. IEEE ICC'03, Anchorage, USA, May, 2003
4. P. Garg, P. Doshi, R. Greene, M. Baker and X. Cheng, "Using IEEE 802.11e MAC for QoS over Wireless", Proc. IEEE International Performance Computing and Communications Conference 2003, Phoenix, USA, April 2003
5. H. Zhu, I. Chlamtac, "An Analytical Model for IEEE 802.11e EDCF Differential Services", Proc. the 12th International Conference on Computer Communications and Networks (ICCCN'03), Dallas, USA. Oct. 2003
6. G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function", IEEE Journal on Selected Area in Communication. pp. 514-519, Vol. 18, No. 3, Mar. 2000
7. H. Wu, Y. Peng, K. Long and J. Ma, "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement", Proc. IEEE Inforcom'02. 2002
8. Y. C. Tay, K. C. Chua, "A Capacity Analysis for the IEEE 802.11 MAC Protocol", Wireless Networks, Vol. 7 No. 2, pp. 159-171, Mar./April 2001
9. Ho-Jin Shin, Dong-Ryeol Shin and Hee-Young Youn, "An Efficient Backoff Scheme for IEEE 802.11 DCF", LNCS 3260, pp.180-193, 2004
10. M. Gast, 802.11 *Wireless Networks: the Definitive Guide*. O'Reilly, 2002

Some Properties of Grounding Modal Conjunctions in Artificial Cognitive Agents

Radosław Piotr Katarzyniak

Institute of Information Science and Engineering, Wrocław University of Technology
Wybrzeże Wyspiańskiego 27, 50-350 Wrocław
`radoslaw.katarzyniak@pwr.wroc.pl`

Abstract. In this paper properties of an originally introduced idea of epistemic satisfaction relation are discussed. It is assumed that this relation defines states of artificial cognition in which particular cases of modal conjunctions are well grounded. Presented theorems show that the epistemic satisfaction preserves the desirable similarity between artificially realized grounding and its natural counterpart known from the context of natural language discourse. The most important conclusion is that as regards to the set of modal conjunctions it is possible to develop an artificial cognitive agent that is able to carry out semantic communication reflecting the natural language communication. In order to achieve this property the artificial cognitive agent needs to be equipped with a certain system modality thresholds.

1 Introduction

The language grounding problem is a well known research issue considered by cognitive science and similar fields of artificial intelligence such as robotics e.g. [2,12,13]. A semantic language (of communication) is treated as well grounded if all its formulas are assigned certain meaning. This meaning is extracted from the content of agent-world interactions and stored (embodied) internally in private knowledge bases developed by the agent in the process of semiosis e.g. [12,13]. In case of natural cognitive agents the grounding of languages is a natural phenomenon related to the phenomenon of language meaning creation and is realized by specialized biological mechanisms e.g. by neural networks. This situation is different to the case of artificial cognitive agents for which the problem of grounding a certain artificial language is always an open question and requires individual solution.

In this paper the language grounding problem is considered for a certain class of artificial cognitive agents and a semantic language of communication consisting of conjunctions extended with modal operators of knowledge, belief and possibility. Following the semiotic approach to communication [1] the language grounding is treated here as strongly related to the so called semiotic triangle and is originally modelled by epistemic satisfaction relations. Figure 1 shows general outline of the proposed approach. In particular, the semiotic triangle consists of three elements: a formula, an object (external to the cognitive agent) and meaning (developed and stored internally by the agent). The role of the meaning is fundamental for the realization of semantic communication between agents. Namely, it is the developed and embodied meaning that makes it possible for agents to use a language as an external representation of certain knowledge.

The epistemic satisfaction relation is a concept that has been originally introduced in this approach in order to model the relation between an interpreted formula and its meaning. In this sense the epistemic satisfaction relation describes the way in which a certain formula is grounded in the overall knowledge stored in the internal cognitive space of the agent. Particular cases of grounding have already been considered elsewhere. For instance the original definitions for epistemic satisfaction of simple modalities, modal conjunctions, modal exclusive disjunctions and modal disjunctions can be found in [9], for modal implications in [8] and for modal equivalences in [5]. Related implementation issues are discussed in [3,7] and examples of applications to formal analysis of query answering in [4] and modelling imprecise knowledge extraction from organizational information resources in [11].
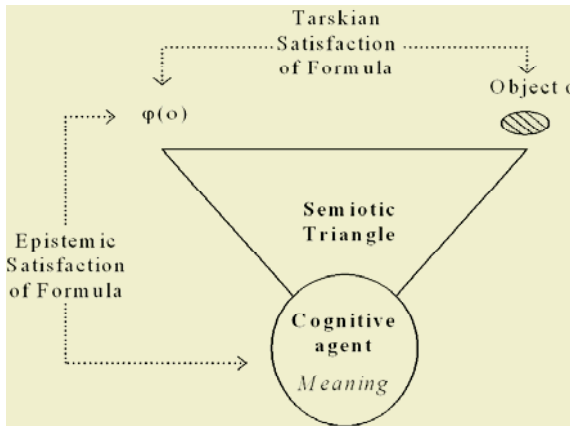
Fig. 1. Semiotic triangle and its relation to the idea of epistemic satisfaction relation (grounding relation) and Tarskian approach to defining the truth of languages

The results presented in this paper are related to the definition of epistemic satisfaction and grounding of modal conjunctions. In particular, the theorems given below say what is the degree of similarity between commonsense natural language rules of modal conjunctions usage and the semantics and pragmatics of artificial systems based on the proposed idea of epistemic satisfaction and grounding of modal conjunctions. Another issue that is discussed, too, is the possibility of creating an artificial cognitive agent that can utilize the idea of epistemic satisfaction of conjunctions in the way corresponding to a predefined set of commonsense language requirements.

2   Defining Epistemic Satisfaction and Grounding of Modal Conjunctions

The language of modal conjunctions is given in an usual way [9]. The alphabet is defined by the following sets of symbols:

- The set $O = \{o_1, \ldots, o_M\}$ consisting of all individual constants.
- The set $P = \{p_1, \ldots, p_K\}$ consisting of unary predicate symbols related to properties $P_1, \ldots, P_K$ of objects from O, respectively.

- The symbols Pos, Bel, Know called the modal operator of possibility, belief and knowledge, respectively.
- The symbols $\neg$ and $\wedge$ called the connective of logical negation and equivalence, respectively.

For each $p,q \in P$, $p \neq q$, and $o \in O$ the following formulas:

$p(o) \wedge q(o)$, $\neg p(o) \wedge q(o)$, $p(o) \wedge \neg q(o)$, $\neg p(o) \wedge \neg q(o)$,

$Pos(p(o) \wedge q(o))$, $Pos(\neg p(o) \wedge q(o))$, $Pos(p(o) \wedge \neg q(o))$, $Pos(\neg p(o) \wedge \neg q(o))$,

$Bel(p(o) \wedge q(o))$, $Bel(\neg p(o) \wedge q(o))$, $Bel(p(o) \wedge \neg q(o))$, $Bel(\neg p(o) \wedge \neg q(o))$,

$Know(p(o) \wedge q(o))$, $Know(\neg p(o) \wedge q(o))$, $Know(p(o) \wedge \neg q(o))$, $Know(\neg p(o) \wedge \neg q(o))$

are well built formulas.

The commonsense (intentional) meaning of these formulas is given in a usual way [9]:

The formula $p(o) \wedge q(o)$ represents the meaning communicated by the natural language statement "Object o exhibits the property P and object o exhibits the property Q.".

The formula $\neg p(o) \wedge q(o)$ represents the meaning communicated by the natural language statement "Object o exhibits the property P and object o does not exhibit the property Q.".

The formula $p(o) \wedge \neg q(o)$ represents the meaning communicated by the natural language statement "Object o does not exhibit the property P and object o exhibits the property Q.".

The formula $\neg p(o) \wedge \neg q(o)$ represents the meaning communicated by the natural language statement "Object o does not exhibit the property P and object o does not exhibit the property Q.".

The meaning of modal operators Know, Bel and Pos is given by "I know that…", "I believe that…" and "It is possible that …", respectively.

The above formulas are confronted with particular cognitive states of artificial agents. At each time point states of cognition are defined by structured set of so called base profiles. Base profiles are internal representations of individual observations made and stored by the cognitive agent (see [6]). Base profiles are distributed among non-conscious and conscious cognitive subspaces which role in artificial cognitive agents is assumed to be similar to the role of embodied non-conscious and conscious empirical material in natural cognitive agents. A broader discussion of this issue can be found in [5]. The internal representation of observation of the external world realized by the agent at the moment t is given by the relational system

$$BP(t) = <O, P_1^+(t), P_1^-(t), P_2^+(t), P_2^-(t),…, P_K^+(t), P_K^-(t) >\},$$

where the following interpretations and constraints are assumed:

- The set $O= \{o_1, …, o_M\}$ consists of all representations of atom objects $o \in O$, where the symbol o (used as a part of base profile) denotes a unique internal cognitive representation of a related atomic object located in the external world.
- For each $i=1,2,...,K$, both $P_i^+(t) \subseteq O$ and $P_i^-(t) \subseteq O$ hold.
- For each $i=1,2,...,K$ and $o \in O$ the relation $o \in P_i^+(t)$ holds if and only if the agent's point of view is that the object o exhibited the atomic property P_i and this fact was empirically verified at the time point t by the cognitive agent itself.

- For each i=1,2,...,K and o∈O the relation o∈ P_i^-(t) holds if and only if the agent's point of view is that the object o did not exhibit the atomic property P_i and this fact was empirically verified at the time point t by the agent itself.
- For each i=1,2,...,K, the condition P_i^+(t)∩ P_i^-(t)=∅ holds.

It is assumed that all base profiles are stored in the set KS(t) defining the overall content of strictly private empirical knowledge base. The complete description of cognitive states requires the introduction of the following four sets C^i(t), i=1,2,3,4 consisting of base profiles related to particular conjunctions:

C^1(t) = {BP(t_n): t_n≤^{TM}t and BP(t_n)∈ KS(t) and o∈P$^+$(t_n) and o∈Q$^+$(t_n) hold}
C^2(t) = {BP(t_n): t_n≤^{TM}t and BP(t_n)∈ KS(t) and o∈P$^+$(t_n) and o∈Q$^-$(t_n) hold}
C^3(t) = {BP(t_n): t_n≤^{TM}t and BP(t_n)∈ KS(t) and o∈P$^-$(t_n) and o∈Q$^+$(t_n) hold}
C^4(t) = {BP(t_n): t_n≤^{TM}t and BP(t_n)∈ KS(t_k) and o∈P$^-$(t_n) and o∈Q$^-$(t_n) hold}

Obviously, the sets C^1(t), C^2(t), C^3(t) and C^4(t) are related to conjunctions p(o)∧q(o), ¬p(o)∧q(o), p(o)∧¬q(o), ¬p(o)∧¬q(o), respectively. Moreover, each of them consists of all embodied empirical material in which related conjunctions are grounded and from which their meaning is extracted [9].

If the internal conscious and non-conscious cognitive spaces are denoted by WM(t) and PM(t), respectively, then at each time point t the distribution of the sets C^1(t), C^2(t), C^3(t) and C^4(t) over WM(t) and PM(t) can be described by pairs of sets WC^i = WM(t)∩C^i, PC^i = PM(t)∩C^i, where for each i=1,2,3,4 the equations WC^i∩PC^i=∅ and WC^i∪PC^i=C^i hold. In consequence the state of empirical material related to all considered conjunctions can be described by the set of base profiles' sets CS(t) = {WC^1(t), PC^1(t), WC^2(t), PC^2(t), WC^3(t), PC^3(t), WC^4(t), PC^4(t)}.

The epistemic satisfaction and grounding of modal conjunctions p(o)∧q(o), Pos(p(o)∧q(o)), Bel(p(o)∧q(o)), Know(p(o)∧q(o)) is given by the following definition [9]:

Definition. *(Epistemic satisfaction and grounding of p(o)∧q(o) and its modal extensions) Let the following be given: a time point t∈T, a state of cognition CS(t) = {WC^1(t), PC^1(t), WC^2(t), PC^2(t), WC^3(t), PC^3(t), WC^4(t), PC^4(t)} and a system of modality thresholds $0 < \lambda_{minPossibility} < \lambda_{maxPosssibility} < \lambda_{minBelief} < \lambda_{maxBelief} <1$. It is assumed for each pair of different properties P, Q∈{P_1,...,P_K} and each object o∈O that:*

The epistemic satisfaction relation CS(t)|=$_G$Pos(p(o)∧q(o)) holds if and only if o∈O/(P$^+$(t)∪P$^-$(t)), o∈O/(Q$^+$(t)∪Q$^-$(t)), WC^1≠∅ and λ_{minPos}≤λ(p(o)∧q(o))≤λ_{maxPos}.

The epistemic satisfaction relation CS(t)|=$_G$Bel(p(o)∧q(o)) holds if and only if o∈O/(P$^+$(t)∪P$^-$(t)), o∈O/(Q$^+$(t)∪Q$^-$(t)), WC^1≠∅ and λ_{minBel}≤λ(p(o)∧q(o))≤λ_{maxBel}.

The epistemic satisfaction relations CS(t)|=$_G$Know(p(o)∧q(o)) and CS(t)|=$_G$p(o)∧q(o) hold if and only if o∈P$^+$(t), o∈Q$^+$(t).

Similar definitions can be given in an obvious way for remaining conjunctions ¬p(o)∧q(o), p(o)∧¬q(o) and ¬p(o)∧¬q(o) (see also [10]).

Obviously in this definition the epistemic satisfaction and grounding are treated as equivalent. Namely, a modal conjunction is satisfied in the epistemic sense if and only if it is well grounded.

3   Properties of Epistemic Satisfaction and Grounding of Modal Conjunctions

The above definitions for epistemic satisfaction and grounding assure a few desirable properties of grounding language formulas, provided that all formulas are assigned assumed intentional interpretations. The first group of these properties relates to inherited interrelations between epistemic satisfaction of particular modalities:

Theorem 1. *Let $\varphi \in \{p(o) \wedge q(o), p(o) \wedge \neg q(o), \neg p(o) \wedge q(o), \neg p(o) \wedge \neg q(o)\}$.*

1.1. *If $CS(t)|=_G Pos(\varphi)$ holds, then $CS(t)|=_G Bel(\varphi)$ and $CS(t)|=_G Know(\varphi)$ do not hold.*
1.2. *If $CS(t)|=_G Bel(\varphi)$ holds, then $CS(t)|=_G Pos(\varphi)$ and $CS(t)|=_G Know(\varphi)$ do not hold.*
1.3. *If $CS(t)|=_G Know(\varphi)$ holds, then $CS(t)|=_G Pos(\varphi)$ and $CS(t)|=_G Bel(\varphi)$ do not hold.*

Properties mentioned in theorem fulfill the natural requirement that the concurrent use of two different modal extensions of the same conjunction $\varphi \in \{p(o) \wedge q(o), p(o) \wedge \neg q(o), \neg p(o) \wedge q(o), \neg p(o) \wedge \neg q(o)\}$ is irrational and therefore unacceptable. Proofs of these properties are direct consequences from the definitions for epistemic satisfaction of modal conjunctions and can be found in [10].

There are two important properties related to the grounding of knowledge extensions of conjunctions that correspond to the pragmatics of natural language discourse:

Theorem 2. *Let $\varphi \in \{p(o) \wedge q(o), p(o) \wedge \neg q(o), \neg p(o) \wedge q(o), \neg p(o) \wedge \neg q(o)\}$. The relation $CS(t)|=_G Know(\varphi)$ holds if and only if the relation $CS(t)|=_G \varphi$ holds.*

This desirable property results directly from the definition of epistemic satisfaction (see also [10]).

Theorem 3. *Let $\varphi \in \{p(o) \wedge q(o), p(o) \wedge \neg q(o), \neg p(o) \wedge q(o), \neg p(o) \wedge \neg q(o)\}$. For each state of cognition $CS(t)$ one and only one of the modal conjunctions $Know(p(o) \wedge q(o))$, $Know(p(o) \wedge \neg q(o))$, $Know(\neg p(o) \wedge q(o))$ or $Know(\neg p(o) \wedge \neg q(o))$ can be well grounded.*

This desirable property results directly from the definition of grounding. Again, the detailed proof can be found in [10]. The obvious conclusion from theorem 2 and theorem 3 is the following theorem:

Theorem 4. *Let $\varphi \in \{p(o) \wedge q(o), p(o) \wedge \neg q(o), \neg p(o) \wedge q(o), \neg p(o) \wedge \neg q(o)\}$. For each state of cognition $CS(t)$ one and only one of the conjunctions $p(o) \wedge q(o)$, $p(o) \wedge \neg q(o)$, $\neg p(o) \wedge q(o)$ or $\neg p(o) \wedge \neg q(o)$ can be well grounded.*

Other properties of epistemic satisfaction relation are related to the commonsense requirements underlying the logic of acceptable and unacceptable concurrent grounding of subsets of modal conjunctions.

Theorem 5. *Let four different conjunctions α, β, γ, $\chi \in \{p(o) \wedge q(o), p(o) \wedge \neg q(o), \neg p(o) \wedge q(o), \neg p(o) \wedge \neg q(o)\}$ be given.*

5.1. *The simultaneous grounding of more than one formulas $Bel(\alpha)$, $Bel(\beta)$, $Bel(\gamma)$ or $Bel(\chi)$ requires that $\frac{1}{2} < \lambda_{minBel}$ or $\lambda_{maxBel} < \frac{1}{2}$ holds.*
5.2. *The simultaneous grounding of all formulas $Pos(\alpha)$, $Pos(\beta)$, $Pos(\gamma)$ and $Pos(\chi)$ requires that $\lambda_{minPos} \leq \frac{1}{4} \leq \lambda_{maxPos}$ holds.*

Obviously, the theorem 5.1 can be used to define the system of modality thresholds which rejects the possibility of epistemic satisfaction of more than one modal belief conjunction and the theorem 5.2 says what is required to ensure the possibility of simultaneous grounding of all modal possibility conjunctions.

Theorem 6. *Let four different conjunctions α, β, γ, $\chi \in \{p(o) \wedge q(o),\ p(o) \wedge \neg q(o),\ \neg p(o) \wedge q(o),\ \neg p(o) \wedge \neg q(o)\}$ be given. It is possible to develop a system of modality thresholds that makes it possible to ground the following sets of modal conjunctions:*

$CS(t)|=_G Pos(\alpha)$	$CS(t)|=_G Pos(\alpha)$	$CS(t)|=_G Pos(\alpha)$	$CS(t)|=_G Pos(\alpha)$
$CS(t)|=_G Pos(\beta)$	$CS(t)|=_G Pos(\beta)$	$CS(t)|=_G Pos(\beta)$	
$CS(t)|=_G Pos(\delta)$	$CS(t)|=_G Pos(\delta)$		
$CS(t)|=_G Pos(\phi)$			

$CS(t)|=_G Bel(\alpha)$	$CS(t)|=_G Bel(\alpha)$	$CS(t)|=_G Bel(\alpha)$	
$CS(t)|=_G Pos(\beta)$	$CS(t)|=_G Pos(\beta)$	$CS(t)|=_G Pos(\beta)$	
$CS(t)|=_G Pos(\delta)$	$CS(t)|=_G Pos(\delta)$		
$CS(t)|=_G Pos(\phi)$			

and ensures the impossibility of simultaneous grounding of the following sets of modal conjunctions:

$CS(t)|=_G Bel(\alpha)$	$CS(t)|=_G Bel(\alpha)$	$CS(t)|=_G Bel(\alpha)$	$CS(t)|=_G Bel(\alpha)$
$CS(t)|=_G Bel(\beta)$	$CS(t)|=_G Bel(\beta)$	$CS(t)|=_G Bel(\beta)$	
$CS(t)|=_G Bel(\delta)$	$CS(t)|=_G Bel(\delta)$		
$CS(t)|=_G Bel(\phi)$			

Detailed proofs for situations mentioned in the theorem 6 can be found in [10].

4 Final Remarks

All theorems presented above show that the epistemic satisfaction relation preserves commonsense and desirable properties of rational language behaviour. Namely, the epistemic satisfaction is defined in a way that makes it impossible for an artificial cognitive agent to ground simultaneously sets of contradictory modal conjunctions provided that the contradiction of modal conjunctions is understood as at the intentional level of semantic communication. Two groups of properties can be determined as regards to the above problem. The first one is referred to in the theorems 1-4. Namely, in these theorems particular cases of contradictory sets of modal conjunctions are named. The general idea of this contradiction is that if a certain modal conjunction is perceived by an artificial cognitive agent as well grounded, then the remaining types of modal extensions of the same conjunction should not be allowed for not being grounded. At the same time all situations referred to in the theorems 5-6 are more subtle. The major conclusion that follows from theorems 5-6 is that as regards to the set of modal conjunctions it is always possible to develop an artificial cognitive agent that is able to simultaneously ground or simultaneously reject some sets of modal conjunctions. However, in order to achieve this possibility the artificial cognitive agent needs to be equipped with a very certain system of modality thresholds. If an internally embodied system of modality thresholds does not exhibit particular proper-

ties, then some rational language behaviour can be ever unavailable for the agent and some unacceptable language behaviour may be possible. A detailed discussion of these issues can be found in [10].

References

1. Eco, U.: La struttura assente, Tascabili Bompiani, Milano (1991)
2. Harnad, S.: The symbol grounding problem., Physica D 42, 335-346.
3. Katarzyniak, R., Pieczyńska-Kuchtiak, A.: Grounding languages in cognitive agents and robots., Systems Science, Vol. 29, no. 2 (2003), 113-127.
4. Katarzyniak, R., Pieczyńska-Kuchtiak, A.: Int. J. Applied Mathematics and Computer Science, Vol. 14, no. 2 (2004), 249-263.
5. Katarzyniak, R.: The language grounding problem and its relation to the internal structure of cognitive agents., Journal of Universal Computer Science, Vol. 11, no. 2 (2005), 357-374.
6. Katarzyniak, R., Nguyen, T.N.: Reconciling inconsistent profiles of agent's knowledge states in distributed multiagent systems using consensus methods., Systems Science, Vol. 26, No. 4 (2000), 93-119.
7. Katarzyniak, R., Pieczyńska-Kuchtiak, A.: A consensus based algorithm for grounding belief formulas in internally stored perceptions., Neural Network World, Vol. 5. (2002), 461-472.
8. Katarzyniak, R.: Extracting modal implications and equivalences from cognitive minds., DS2004, LNAI 3245 (2004), 420-428.
9. Katarzyniak, R.: Grounding modalities and logic connectives in communicative cognitive agents. [in] Nguyen N.T. (ed.) Intelligent technologies for inconsistent knowledge processing, Int. Series on Advanced Intelligence, Volume 10, Advanced Knowledge International, Adelaide (2004), 21-37.
10. Katarzyniak, R.: On some properties of grounding modal conjunctions in artificial cognitive agents., Technical Reports, Institute of Information Science and Engineering, Wrocław University of Technology, Wrocław, Poland (2005).
11. Katarzyniak, R.: Organizations as communicative cognitive agents., Proc. 7th Org. Semiotics OS 2004, 309-324.
12. Vogt, P.: Anchoring of semiotic symbols., Robotics and Autonomous Systems, Vol. 43, (2003), 109-120.
13. Vogt, P.: The Physical Grounding Symbol Problem., Cognitive Systems Research, Vol. 3, (2002), 429-457.

Multi-agent System for Web Advertising

Przemysław Kazienko

Wrocław University of Technology, Institute of Applied Informatics,
Wybrzeże S. Wyspiańskiego 27, 50-370 Wrocław, Poland
kazienko@pwr.wroc.pl

Abstract. The main aim of a personalized advertising system is to provide advertisements, which are most suitable for the given anonymous user navigating the web site. To achieve this goal, many sources of data are processed in one coherent vector space: the advertisers' and publisher's web site content, sessions of former users from the past, the history of clicks on banners and the current user behavior as well as the advertising policy promoting certain campaigns. The multi-agent system running mostly on the publisher site is introduced to organize personalized advertising. Each cooperating agent is responsible for a separate, specific task: web content and usage mining, click-through data exploration, user monitoring, advertisement recommendation and management.

1 Introduction

Web advertising is the major or often even sole source of income for recent online portals. There are two or three main participants of web advertising: advertisers, publishers and sometimes advertising brokers, which link both advertisers and publishers.

The advertisers want to attract people to visit their web sites, so they pay publishers for exposition their banners or other kinds of advertisements on publisher's pages. Users, that navigate such pages, sometimes click through on a banner, so they shift to advertiser's site. The main task of the advertising broker, such as Real Media [1], DoubleClick [17] or Google's AdSense [7], is to deliver banners appropriate for the given publisher's page.

There is no broker in two other models of advertising, in which the whole advertising management is performed by either advertisers or publishers themselves [4]. In this paper, we focus on this third model: the publisher possesses all advertising data, i.e. data about banners, and the publisher is responsible for presenting adequate advertisements on its pages. Such model enables the usage of specific information usually available only for the publisher: data about users and their behaviors.

There are two main research issues in web advertising: scheduling and personalization. The main aim of the former is to maximize the total click-through-rate for all advertisements by proper managing exposition time and advertising space on the web page [2, 5, 18]. The latter seems to be an important challenge for current advertisers. It appears to be more "individualized" than typical target advertising, which partitions customers in a market into specific segments [8]. It aims to assign a suitable advertisement to a single web user rather than to a group of individuals [3, 16].

This paper is part of research carried out on personalized recommendation systems [9, 10, 12, 13, 14, 15]. Advertisements are only one, although specific, kind of objects, which can be suggested to a user.

2   Personalized Web Advertising

The general concept of personalized online advertising was originally presented in [11]. There are five main factors taken into account at personalization: content of both publisher's pages and advertisers' portals, historical behaviors of users, the behavior of the current user, recently exposed advertisements, and advertising features (Fig. 1).
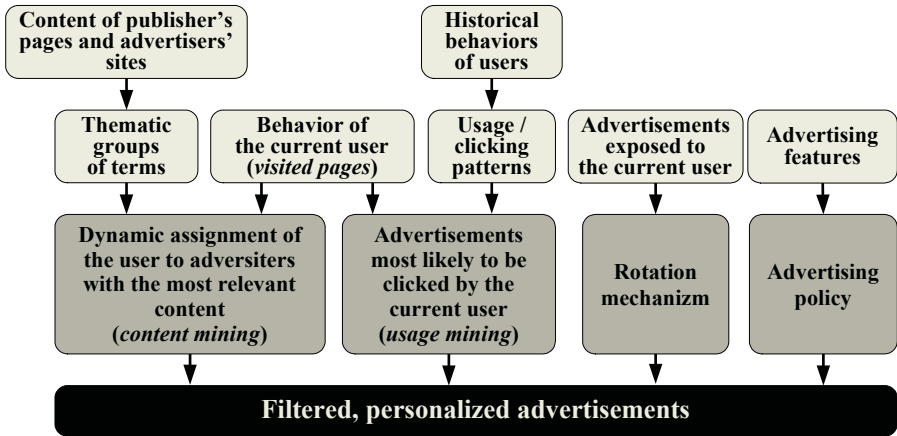
Fig. 1. Main factors of filtering and personalization of advertisement

Historical behaviors are derived from the past user sessions. A session is the set of pages watched by the user during one visit to the publisher's web site. Simultaneously, the data about banners clicked during the session is retained. Thus, we have exactly one ad usage session corresponding to each page user session. Usage patterns are obtained by clustering user sessions offline. Due to one-to-one relationship between user and ad session, we receive also clicking patterns related to typical usage patterns. This course of action is called *web* and *ad usage mining*.

The content is processed in a sense alike. Content features, in the form of terms, are extracted from HTML source of publisher's pages. Next, selected terms are clustered to achieve thematic groups, called *conceptual spaces* [13]. Since we cluster terms, individual publisher's pages may belong to many clusters. Besides, we take the content of the whole web sites linked by particular benners. Each advertiser's web site is treated like one page from publisher's portal. Advertiser's web sites that contain terms from a thematic group are allocated to this group. In this way we establish a one-to-one relationship between the content of publisher's site and advertisements.

Having content and usage clusters, we can dynamically assign each single user to the most relevant patterns based on their current behavior. This assignment is performed online and at each user HTTP request. In consequence, the user is suggested advertisements with the content most relevant to the content of pages viewed by them recently. Additionally, advertisements most likely to be clicked by the user have the greater chance to be exposed. It comes from the historical behavior of other users that are similar to the current one: they simply used to click certain banners. Note that the user is assigned to one usage pattern but this pattern corresponds to one clicking pattern.

Another personalization factor is related to rotation mechanism. All advertisements presented to the current user are retained to prevent from the emission of the same banner too frequently for one user. Furthermore, there are some advertising features, such as the maximum number of exposition per user, that force the system to monitor all advertisements suggested to each user. Additionally, advertising policy includes also data that enable to promote certain campaigns, e.g. those more profitable.

Some mentioned above processes such as content processing, clustering, i.e. pattern discovery, are carried out offline, while some are performed online: the assignment of the user to the most suitable patterns, monitoring of the current user behavior. Due to permanent changes in source data: new advertisements, updates in content, etc., the offline activities have to be periodically repeated.

All data in the system is processed in one coherent vector space – this facilitates the integration of heterogeneous data sources. There are two kinds of vectors: with dimension equal to either the number of pages in publisher's site or the number of advertisements.

3   The Multi-agent Architecture

The advertising system was designed using a multi-agent architecture, in which expert-agents cooperate with one another and may be distributed among many hosts [6]. Every agent is responsible for a single task and it encapsulates specific functions that are available for the rest of the system. Not only do they interchange agents' information, but they also possess their own knowledge. Some agents are involved in usage mining processes while the other ones are used to discover useful knowledge from the content (Fig. 2). The appropriate usage and content patterns are assigned to the current user so that the system is able to recommend the best, personalized advertisements and additionally the suitable hyperlinks i.e. to suggest "the next navigational steps". Let us focus on tasks accomplished by the individual agents.

Web Crawler retrieves the web site content, i.e. individual pages, using HTTP and following hyperlinks existing on pages. Next, it extracts terms from the HTML content and calculates the document-term frequency useful at clustering. *Web Crawler* is used to analyze both the publisher's and advertisers' web sites linked by the advertisements. It delivers data for *Publisher* and *Advertiser Content Manager*.

Search Engine is an auxiliary agent normally used for Information Retrieval within the content of the publisher's site. Nevertheless, it also retains all terms used by users for searching. *Publisher Content Manager* uses these terms at selection of the proper descriptors.

Ad Manager includes user interface and provides some advertising campaign management tools. It keeps all advertisement features such as the total number of expositions and number of expositions per user, etc.

Publisher Content Manager selects terms from the publisher's content, which have been delivered by *Web Crawler*, to extract only "good descriptors" of the content. Besides, it creates the vector representations of these terms and determines their significance according to their frequency.

Advertiser Content Manager uses capabilities of *Publisher Content Manager* and performs analogical operations on the terms retrieved by *Web Crawler* from whole

advertisers' web sites. To achieve it, the agent communicates with *Ad Manager* to obtain URL addresses of individual advertisements.

Content Miner clusters term vectors from publisher's pages and it calculates the mean vector (centroid) of each cluster.

Advertiser Content Miner assigns advertisers' sites to clusters provided by Content Miner based on the terms, which are common for both advertisers' and publisher's sites. In this way, we obtain equivalent content clusters for advertisements.
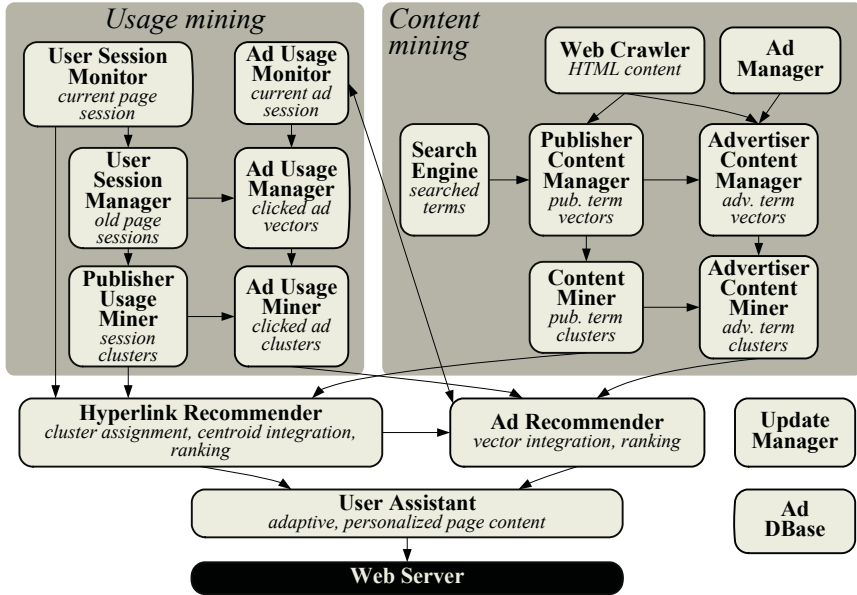
Fig. 2. Multi-agent architecture of AdROSA system

User Session Monitor keeps track of user activities i.e. visited pages. This information reflects the current user behavior and it is crucial at the assignment of the user to the appropriate content and usage patterns. Finally, the session data is stored by *User Session Manager* after the user leaves the publisher's web site.

User Session Manager creates and updates session vectors. Users often change their behavior, so we should not rely on older sessions with the same confidence as on newer ones. For that reason, *User Session Manager* periodically decreases coefficients of older sessions [10].

Publisher Usage Miner clusters historical user sessions and then calculates their centroids representing typical usage patterns. One of such patterns is assigned to a user at each HTTP request.

Ad Usage Monitor controls displayed and clicked banners: it checks exposition limits (total and per one user), and retains data useful for the rotation mechanism – the current ad session. It also relays data about exposed advertisements to *Ad Manager* and additionally, it passes on the clicked advertisements to *Ad Usage Manager* after the session is finished.

Ad Usage Manager creates and maintains the historical clicked ad vectors.

Ad Usage Miner clusters clicked ad vectors and it calculates clicking patterns - centroids. Clustering is based on the usage patterns previously delivered by *Publisher Usage Miner* rather than on source clicked ad vectors. Nevertheless, these vectors are used at calculating of centroids, which represent the advertisements most likely to be clicked by users navigating the web site according to the corresponding visiting (usage) pattern.

Hyperlink Recommender is responsible for assignment of the user to the appropriate usage (session cluster) and content pattern (term cluster). To achieve this goal, it makes use of knowledge maintained by *Content Miner* – content patterns, *Publisher Usage Miner* – usage patterns, and *User Session Monitor* – current user behavior. Having suitable patterns assigned, *Hyperlink Recommender* extracts from them the publisher's pages with the highest score. Such pages are relayed to the *User Assistant* as hyperlink recommendations. This process is carried out at each user request.

Ad Recommender performs the main online personalization process of advertising. It gains patterns (vectors) assigned to the current user by *Hyperlink Recommender*. Next, it gets the corresponding to them advertising patterns: the term and clicking ones, delivered by *Advertiser Content Miner* and *Ad Usage Miner*, respectively. *Ad Recommender* integrates obtained patterns and creates the advertising ranking. This ranking list is compared against the already exposed advertisements provided by *Ad Usage Monitor*. As a result, we accomplish the most advisable advertisements, which are returned to *User Assistant* and *Ad Usage Monitor*.

User Assistant dynamically incorporates hyperlinks and banners delivered by *Hyperlink* and *Ad Recommender*, respectively, into the HTML content of the page returned to the user.

Update Manager is responsible for monitoring the publisher's web site, the validity of campaigns and other factors of changes in source data. Its main task is to recognize whether changes are serious enough to start the offline update processes.

Ad DBase provides database connectivity. It performs operations requested by other agents on the database.

4   Knowledge Maintenance

Changes in data sources are the serious problem in almost all web environments. It makes the knowledge derived from various initial data to be periodically updated, so that the hyperlink and advertisement recommendations are reliable. There are many factors of changes with different intensity and importance onto the quality of recommendation, such as new advertisements, sets of new user sessions, new and updated publisher's pages, etc. (Fig. 3). There is a special agent – *Update Manager*, which task is: to monitor these factors; to decide, whether and what the update process should be carried out, and to synchronize the process. Note that the update is usually the very resource consumable and long-lasting process. The synchronization among all involved agents is necessarily because the system works continuously and the switch to the new knowledge should be done "at one moment". For example, a big number of new user sessions makes all usage and clicking patterns to be recalculated. This triggers four agents: *User Session Manager* and *Miner*, *Ad Usage Manager* and *Miner*. For that reason, *Update Manager* has to know all dependencies between the knowledge used by particular agents.

Update Manager can initiate the update of only (Fig. 3): clicking patterns (2), both usage and clicking patterns (1+2), advertisers' thematic groups (3), both publisher's and advertisers' thematic groups (4+3), or the whole system knowledge (1+2+3+4). Additionally, some changes are much or less easy to introduce, e.g. the insertion of new advertisements results only in processing of the new advertiser's web site (performed by *Web Crawler*) and the relatively simple extension of appropriate vectors (2+3) related to *Ad Usage Manager* and *Miner* as well as *Advertiser Content Manager* and *Miner*. In opposite, the large enough modification of content of publisher's site forces to rebuild all both publisher's and advertisers' thematic groups (3+4).
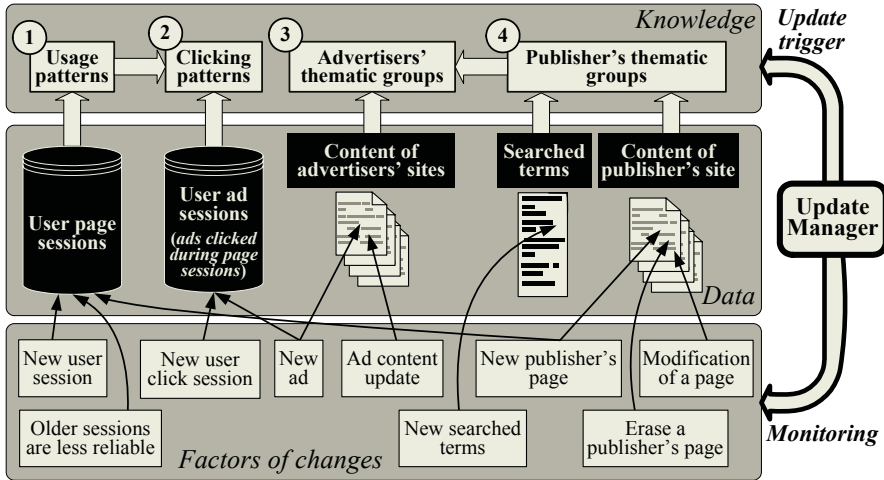
Fig. 3. Information flow and some factors of changes in personalized advertising

5   Conclusions and Future Work

The proposed system "personalizes" the online advertising using web usage and content mining. Additionally, it is able to recommend hyperlinks in the similar, adaptive way. Due to its multi-agent architecture, it is more flexible, scalable and open for new functions such as the sophisticated scheduling mechanisms and extension of personalization with knowledge derived from purchase history as well as product or page ratings gathered in e-commerce [15]. Furthermore, the multi-agent platform makes easier the introduction of diverse update processes, what is very important in the ever changing web environment. The system as the prototype was put into practice for the copy of www.poland.com web site.

References

1. 24/7 Real Media, Inc., http://www.realmedia.com/
2. Amiri A., Menon S.: Efficient Scheduling of Internet Banner Advertisements. ACM Transactions on Internet Technology **3** (4) (2003) 334–346.
3. Bae S.M., Park S.C., Ha S.H.: Fuzzy Web Ad Selector Based on Web Usage Mining. IEEE Intelligent Systems **18** (6) (2003) 62-69.

4. Bilchev G., Marston D.: Personalised Advertising – Exploiting the Distributed User Profile. BT Technology Journal **21** (1) (2003) 84-90.
5. Chickering D.M., Heckerman D.: Targeted Advertising with Inventory Management. 2nd ACM Conference on Electronic Commerce, EC00 (2000) 145-149; Interfaces **33** (5) (2003) 71-77.
6. Ferber J.: Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence. Addison Wesley Longman, Boston (1999).
7. Google Advertising Programs, http://www.google.com/ads/
8. Iyer G., Soberman D., Villas-Boas J.M.: The Targeting of Advertising. Marketing Science, (2005) to appear.
9. Kazienko P.: Multi-agent Web Recommendation Method Based on Indirect Association Rules. 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems, KES'2004, **LNAI 3214**, Springer Verlag (2004) 1157-1164.
10. Kazienko P.: Product Recommendation in E-Commerce Using Direct and Indirect Confidence for Historical User Sessions. 7th International Conference on Discovery Science, DS'04, **LNAI 3245**, Springer Verlag (2004) 255-269.
11. Kazienko P., Adamski M.: Personalized Web Advertising Method. Adaptive Hypermedia AH 2004, **LNCS 3137**, Springer Verlag (2004) 146-155.
12. Kazienko P., Kiewra M.: Integration of Relational Databases and Web Site Content for Product and Page Recommendation. 8th International Database Engineering & Applications Symposium, IDEAS '04, IEEE Computer Society (2004) 111-116.
13. Kazienko P., Kiewra M.: Personalized Recommendation of Web Pages. Chapter 10, in: Nguyen T. (ed.), Intelligent Technologies for Inconsistent Knowledge Processing, Advanced Knowledge International, Adelaide, South Australia (2004) 163-183.
14. Kazienko P., Kiewra M.: ROSA - Multi-agent System for Web Services Personalization. First Atlantic Web Intelligence Conference Proceedings, AWIC'03. **LNAI 2663**, Springer Verlag (2003) 297-306.
15. Kazienko P., Kołodziejski P.: WindOwls – Adaptive System for the Integration of Recommendation Methods in E-commerce. Third Atlantic Web Intelligence Conference Proceedings, AWIC'05, **LNAI 3528**, Springer Verlag (2005) 218-224.
16. Langheinrich M., Nakamura A., Abe N., Kamba T., Koseki Y.: Unintrusive Customization Techniques for Web Advertising. Computer Networks **31** (11-16) (1999) 1259-1272.
17. Online Advertising. DoubleClick Inc. (2004) http://www.doubleclick.com/us/products/online_advertising/
18. Tomlin J.: An Entropy Approach to Unintrusive Targeted Advertising on the Web. Computer Networks **33** (1-6) (2000) 767-774.

A Mobile Agent Approach to Intrusion Detection in Network Systems

Grzegorz Kolaczek, Agnieszka Pieczynska-Kuchtiak, Krzysztof Juszczyszyn,
Adam Grzech, Radosław Piotr Katarzyniak, and Ngoc Thanh Nguyen

Institute of Information Science & Engineering, Wrocław University of Technology, Poland
{grzegorz.kolaczek,agnieszka.pieczynska-kuchtiak,
krzysztof.juszczyszyn,adam.grzech,radoslaw.katarzyniak,thanh}
@pwr.wroc.pl

Abstract. A major problem with current approaches to the automatic intrusion detection is that it is difficult to make a difference between a normal user behaviour and potential attacker. In this case it is very difficult to conduct investigations of attacks without human intervention. In this paper a framework of an original proposal of the multi-agent system is presented. The aim of this system is security evaluation of the network system on the basis of intrusions pattern scenarios matching. Mechanism of intrusions' detection and defining new attacks' templates are introduced. In this proposal soft methods of computing, advanced methods of symbolic knowledge representation and the methods of semantic communication are used.

1 Introduction

There are two main reasons of violation of network security policy: the states of network system components from a technical point of view and on the other hand intentional human commitments. The second case is called an attack and a user that offends against system security policy is called intrusion.

A security management system must be able to automatically cope with intrusions. A major problem with current approaches to the automatically intrusions' detection is that it is difficult to make a difference between a normal user behaviour and potential attacker. The crucial point to maintain high level of network system security is the ability of system to identify any potential sources of attacks and determine if they offend against security system policy. Intrusion detection system must collect different information from variety system resources and be able to analyse them for detecting any symptoms of security breaches [1],[11].

In order to evaluate the current state of network system security and intrusion pattern detection it is necessary to define the parameters describing the states of the system that will be under control. In addition the time intervals in which the values of these parameters will be captured must be specified.

Due to the complexity of the problem of vulnerability assessment and intrusion detection in a network systems in our project we concentrate on two main problems: problem of detection of known patterns of intrusions attacks and problem of defining new patterns of security breaches. It is assumed that security system must be able to recognise abnormal states of the network system and point out the sources of violence. In other words it is necessary to make a system learn to detect known patterns of attacks and recognise new symptoms of security system breach.

2   Intrusions Detection in Network Systems

Intrusions detection systems (ang. Intrusions Detection Systems – IDS) have been proposed as an approach to cope with current security problems. The aim of the intrusions' detection is discovering of all abnormal states of the system in relation to the network traffic, users' activity and system configurations that may indicate violate of security policy [14].

The idea of security mechanisms is as follows: The intrusions' detection system collects the information about the states of a network system and analysesd them in order to detect symptoms of security breach [2]. But although the idea is very simple, practical solutions of such systems have to deal with a lot of theoretical difficulties. To solve these problems advanced information processing techniques need to be applied.

Difficulties with building intrusions detection systems result in a complexity of the structure of attacks' symptoms, distributed nature of the network systems and dynamics of the source of threats especially the problems of encoding new intrusions scenarios.

The security assessment of a network system requires:

− Applying complex and flexible mechanisms for monitoring values of system attributes that have an influence on the security level of all network system.
− Applying effective computational mechanisms for evaluating the states of system security on the basis of incomplete, uncertain and inconsistent resources.
− Applying the algorithms of machine learning to detect new intrusions pattern scenarios and recognise new symptoms of security system breach in order to update the security system knowledge base.

From the theoretical and practical point of view a lot of works related to the problem of automatic assessment of network security systems take into account distributed nature of such systems [3]-[7], [10], [12]-[15]. There is a quite new approach for building network security systems that is based on the idea of the multiagent systems that can be used in a security network systems area [8], [9], [16]. An example of such multiagent system that has been used to solve the problems with evaluation of the security states in a distributed environment is given in [10]. However the current solutions of the network security systems that apply the multiagent approach have not been put into practice. These solutions present from the conceptual point of view only general ideas of usage the multiagent architectures for automatic assessment of network security systems. And difficult theoretical and practical problems connected with the nature of the real time working distributed systems still remain opened. In such a situation problems of processing of incomplete, uncertain and inconsistent knowledge, the necessity of collective learning and distributed, dynamic knowledge base management need to be taken under deep consideration.

3   The Architecture of the Multiagent System

It is assumed that the network system is consisted of the set of nodes. There are also two types of agents in our multiagent system: monitoring agents (AM) and managing agents (AZ). Monitoring agents (AM) observe the nodes, process captured informa-

tion and draw conclusions that are necessary to evaluate the current state of system security. However managing agents (AZ) are responsible for managing the work of AM agents (Fig. 1). Each agent AM monitors its own area of responsibility consisting of the set of nodes. It is assumed that these areas might mutually overlap. As a consequence of this overlapping the credibility of AM agents' conclusions is rising, the source of attack might be quicker established and the risks of security breach of the main nodes is falling. Each agent AZ manages the set of agent AM. AZ agents' areas cannot mutually overlap. If the number of AM agents is quite small, there is the only one AZ agent in a system.

AM: monitoring agent **AZ:** managing agent
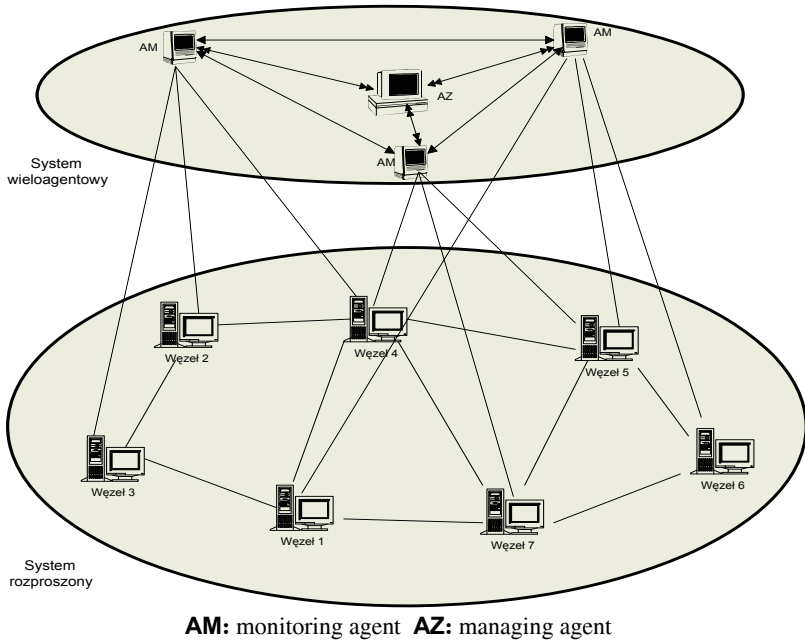
Fig. 1. The architecture of the multiagent system

The main activities of the monitoring agent AM are as follows: monitoring of its region, updating of database frequently or on the basis of a given criteria and updating of security ontology frequently or on the basis of a given criteria. General structure of AM agent is given in Fig.2. With reference to the monitoring task the agent AM establishes its monitoring schedule. In case of detecting any symptoms of security breach it draws a conclusion about a potential attack and sends warning messages to the AZ agent. It is also possible that the agent AM monitors selected nodes on a request of AZ agent.

AM agent's database is updated on the basis of its experiences or on the basis of other agents' knowledge. However security ontology is updated by the agent AM on the basis of its private experiences or as a execution of AZ agent's request that collects ontologies of all AM agents.

AZ agent is responsible for four main activities: creating attacks' profiles, updating its database, updating its ontology and AM agents' managing.
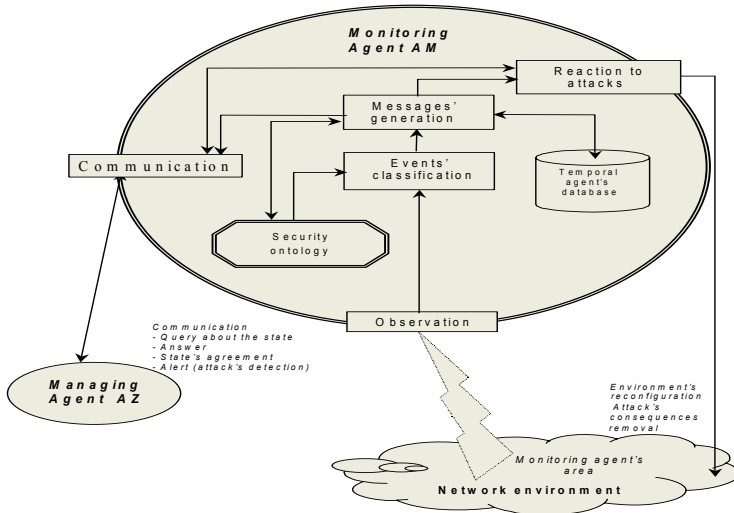
Fig. 2. Monitoring agent's architecture and its environment

In order to create the attack profile AZ agent collects systematically information from AM agents about the nodes in which security policy was broken. The information is gathered at given time intervals or when the attack occurs.

Data from AM agents are verified by AZ agent in order to detect conflict in a attack profile. If the there is a conflict of agents' observations AM agent determines the conflict parameters and degree of opinions' consistency. If the value of this degree is big enough AZ agent determines the consensus profile as a conflict resolution. In case of high degree of data inconsistency or data inadequacy AZ agent sends a request message to AM agents to complete data.

The idea of AZ agent's database and ontology updating is the same as in case of updating of AM agents' databases and ontologies. The only difference is that AZ agent exploits the true facts of security breach in a network system and also administrator's knowledge. AZ agent learns from the administrator of system patterns of intrusions scenarios and rules of how to verify policy. These rules are necessary to detect any violations of security policy in a network system.

The task of managing of AM agents includes verification of information provided by them, determining trust level for them, updating their knowledge bases and ontologies and solving knowledge conflicts between them.

4   Agent's Database and Communication

It is assumed that each AM agent uses relational data structure. It seems to be the best for describing behavior of network nodes and attacks' symptoms. The data collected by AM agent from nodes are intrinsically temporal and are stored in the form of incomplete tuples. Whereas AM agent knowledge consists of the set of rules which represents intrusions' scenarios. Each rule has a form of implication A$\rightarrow$ B, where A is a meta-tuple (complex tuple), and B is an attack's identifier. AM agent's knowledge base is created by: using patterns defined by AZ agent, rules' extracting by

means of data mining techniques and rules' generating on the basis of semantic communication between agents. The agent's ontology contains the set of concepts representing types of attacks, relations between them and the scheme of the knowledge base. Ontologies used by AM and AZ agents are similar, the only difference is that there are the concepts that describe AM agents in AZ agent's ontology.

The aim of agents' communication is information exchanging, warning the agents about the security breaches and attacks, commissioning AM agents to do some tasks, executing tasks, updating databases and ontologies and maintaining interaction mechanisms.

Communication in a multiagent system takes place in case of: 1) detecting by AM agent any symptoms of security breaching in its monitoring area, 2) necessity of AM agent's movement into another monitoring area, 3) necessity of gathering to AZ agent additional information, 4) necessity of updating AM agent's database, 5) necessity of updating AM agent's ontology.

The agents communicate using the modified version of ACL language (agent communication language) that is consisted of two main components: the language of message's content representation and KQML language of communication. The language of message's content representation is based on the first order logic and LISP might be an example of such language. The KQML performatives have been divided into four types: information performatives, queries, requests and warnings. Information performatives are used to exchange information between agents about the states of monitored nodes (*tell* performative, *deny* performative) and the ability to answer the query (*error* performative, *sorry* performative). For questions following queries performatives are used: *evaluate, ask-if, ask-in, ask-one, ask-all, stream-all.* Requests are sent by AZ agent to AM agent in case of:1) necessity of AM agent's ontology updating (*update-ontology* performative), 2) necessity of database updating (*update-database* performative), 3) necessity of attack's pattern updating (*update-attack-template* performative*)*, 4) necessity of AM agent's movement in order to monitor a node that is at risk of attack (*moveto* performative).

Warning performatives are used to warn agents about detected symptoms of security breach and any attacks against the nodes in a network system. AM agent sends a warning to AZ agent using *warn* performative in case of breaking security policy in a monitoring node i.e. stored agent's observations are not consistent with a security policy and it is not possible to classify them to the known patterns of intrusions' scenarios. If it is possible to classify the observations into at least one attack's template then AM agent sends a message to AZ agent using at*tack-detection* performative.

5   Conclusions

In this paper the idea of the security system based on the multiagent approach was presented. This solution solves the problems common for security systems. Thanks to applying the multiagent approach it is possible to conduct investigations of attacks without human intervention, learn a system too recognise new patterns of attacks, integrate different, distributed source of information and in this way to make a system be flexible and more effective. The conception of the mualtiagent security system, presented in this paper, is the first step of building a prototype of a real working system. For a future work it is necessary to work out: 1) patterns of known attacks, 2)

classification of attacks, 3) the conception of the automatic attacks' detection, 4) the language of agents' communication, 5) a structure of agents' database, 6) models of agents' behaviours, 7) domain ontologies and algorithms for its integration and evolution, 8) protocols of agents' co-operation and behaviours' co-ordination and finally 9) methods of conflicts' resolution.

References

1. Bace R., An Introduction to Intrusion Detection and Assessment: for System and Network Security Management, ICSA White Paper, (1998)
2. Błażków R., Gajewski D., Grzech A., Backup and archive services in corporate network., 3rd International Conference on Business Information Systems. BIS '99., Poznań, Poland, (1999) 33-38
3. Bonifácio Jr., Cansian, A., Moreira E., and de Carvalho A., An Adaptive Intrusion Detection System Using Neural Networks. Proceedings of the IFIP World ComputerCongress–Security in Information Systems, IFIP-SEC'98, Chapman & Hall, Vienna, Austria, (1998)
4. Crosbie, M., Spafford G.: Applying Genetic Programming to Intrusion Detection. In:Proceedings of the AAAI Fall Symposium on Genetic Programming. Cambridge, Menlo Park, CA, AAAI Press (1995)
5. Crosbie M, Dole B, Ellis T, Krsul I, Spa.ord E., IDIOT users guide. Technical Report TR-96-050, The COAST Project, Dept of Computer Science, Purdue University, West Lafayette, (1996)
6. Dasgupta, D.: Immunity-Based Intrusion Detection System: A General Framework. In: Proceedings of the 22nd National Information Systems Security Conference, USA (1999)
7. Denning DE, Edwards DL, Jagannathan R, Lunt TF, Neumann PG., A prototype IDES: A real-time intrusiondetection expert system. Technical report, Computer Science Laboratory, SRI International, Menlo Park, (1987)
8. Dobrowolski G., Technologie agentowe w zdecentralizowanych systemach informacyjno-decyzyjnych, Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków. (2002)
9. Ferber Jacques, Multi-Agent Systems, Addison-Wesley Longman (1999)
10. Gorodetski V., Karsaev O., Khabalov A., Kotenko I., Popyack L., Skormin V. Agent-based model of Computer Network Security System: A Case Study. Proceedings of the International Workshop "Mathematical Methods, Models and Architectures for Computer Network Security. Lecture Notes in Computer Science, vol. 2052, Springer Verlag, (2001) 39-50
11. Grzech A., A model for setting computer network security standards, Intelligent systems & applications. ISA '2000, Wollongong, Australia, (2000) 885-890.
12. Jagannathan R, Lunt TF, Anderson D, Dodd C, Gilham F, Jalali C, Javitz HS, Neumann PG, Tamaru A, Valdes A, System Design Document: Next-generation intrusion-detection expert system (NIDES). Technical report, Computer Science Laboratory, SRI International, Menlo Park (1993)
13. Klein, M. Combining and Relating Ontologies: an Analysis of Problems and Solutions. IJCAI'01 Workshop on Ontologies and Information Sharing, Amsterdam, (2001) 53–62.
14. Ko C. Execution Monitoring of Security-critical Programs in a Distributed System: A Specification-based Approach. PhD thesis, Dept of Computer Science, University of California at Davis, (1996)
15. Lee, W., Stolfo, S.J., Mok, K., A Data mining Framework for Building Intrusion Detection Model. In: Proceedings of the IEEE Symposium on Security and Privacy. (1999)
16. Manvi S.S., Venkataram P., Applications of agents technology in communication: a review, Computer Communications 27, (2004) 1493-1508

Non-textual Document Ranking Using Crawler Information and Web Usage Mining

Maciej Kiewra[1] and Ngoc Thanh Nguyen[2]

[1] Fujitsu Services, Spain, Camino Cerro de los Gamos,1 Spain
`mkiewra@mail.fujitsu.es`
[2] Institute of Information Science and Engineering, Wroclaw University of Technology, Poland
`thanh@pwr.wroc.pl`

Abstract. Document ranking creation has a big impact on the whole IR system. The characteristics of the local rankings differ from the global rankings, because the local system can access information unavailable for the global search engines. The purpose of this paper is to discuss non-textual factors that may influence on a document's ranking position and to present a new method of document ranking calculation (*RosaRanking*) that can be useful for local search engines, recommendation systems etc.

1   Introduction

Even a medium web site is composed of many documents related to different topics. The whole web site can be viewed as a mix of documents that may resemble a virtual jungle. For this reason, web sites are often enriched with extra IR (Information Retrieval) tools (search engine, hyperlink recommending system etc.) that facilitate access to the relevant documents, up-to-date information and are able to adapt themselves to users' needs and content changes. Normally, these tools determine a document ranking that reflects the relevance of the document in a given context. Document ranking creation has such a big impact on the whole IR system, that the document position in the ranking should be determined not only by the direct, contextual relevance, but also by document quality and its usage. The main stress in the IR research is laid on the methods applied to the global search engines (e. g. Google [1]). The authors of this paper believe that local search engines and recommendation systems should enhance those algorithms to adapt them for local purposes. Additionally, local systems ought to benefit from the data that are only available locally (for example site usage statistics).

The purpose of this paper is to discuss non-textual factors that influence on a document's ranking positions and to present a new method of document ranking calculation that can be useful for local search engines, recommendation systems etc. For this aim a system (called ROSA [6]) has been performed.

2   Related Works

Document ranking is widely used in IR. The most popular ranking algorithms are: HITS [7], PageRank [2], DirectHit [3]. The first two analyse hyperlinks between documents assuming that the more hyperlinks from "good" documents, the better the

target document is. This assumption is generally true in case of global search engines. In case of local search engines and recommendation systems, this type of ranking reflects rather site navigation model than documents' relevance. If a document A contains a hyperlink to a document B from the same site, it does not have to mean that the author of A thinks that the document B is worth citing. For example, many web masters include links to documents that contains site's privacy policy or a site map in every web page. As a consequence those documents possess high position in the ranking despite their possibly small relevance.

There are some works that try to enhance PageRank for the local search engines (e.g. [4] in which the PageRank algorithm is enriched with information about users' navigation). Similarly, HITS algorithm is also adapted to the local necessities [8].

On the other hand, DirectHit ranking is based on two usage factors: *click popularity* and *stickiness.* The former measures how many times a particular document was clicked from the search engine result page. The latter determines the amount of time that users spend on reading a particular document (when the user click another document from the result page, it is assumed that the previous page has been abandoned). *Click popularity* and *stickiness* may be laden with errors, because not all users enter the page by means of search engines. Moreover the user can open many browser windows and then return to the first open window in order to read its content.

3   Non-textual Factors of Document Ranking

Let the tuple $d=(url, t_0, t_u, m_d, H)$ be a document (web page), where url corresponds to the unique document's URL address that identifies the tuple d unambiguously, t_0 is the timestamp (date and hour) in which the given document d has appeared in the web site; t_u is the total amount of time (measured for example in seconds or hours) in which the document d has been unavailable; m_d is the number of modification that has suffered the document since it has appeared. Finally, $H=\{url_1, url_2, url_3,.., url_m\}$ is the set of all resources to which the document d points. Notice that a $url \in H$ does not have to come from the same web site nor be a document (cascade style sheet or multimedia files are also regarded as resources). Let $D=\{d_1, d_{2,..}, d_n\}$ be the set of all documents (web pages) available in the web site.

Let the tuple $v=(url_v, t_v)$ represent a single *document visit* (denominated also a document hit) where and t_v is a visit's timestamp and:

$$\underset{d \in D}{\exists} (d = (url_d, t_{0d}, t_{ud}, H_d) \wedge url_d = url_v) .$$

Let $s=(v_1, v_{2,...,} v_p)$ represent the *user session* - a sequence of a single document visits performed by a user during one site visit. Let the usage of document $d=(url_d, t_{0d,} t_{ud}, m_d, H_d)$, $d \in D$ be the set of *user sessions* in which the document d has been visited, that is:

$$U_d = \{s : s = (v_1, v_2,..., v_p) \wedge \underset{0 < i \leq p}{\exists} (v_i = (url_{vi}, t_{vi}) \wedge (url_{vi} = url_d)) \}.$$

Let t_c be the current timestamp. It is assumed that before any arithmetical operation is performed on timestamps, they have been converted into numbers. The following functions converting a timestamp to a number are used in this paper: $day(t)$ returns a number of days that has passed since the first January of 1970 to the timestamp t; $hour(t)$ is a number of hours that has passed since the midnight of the first January of

1970 to the timestamp t; $sec(t)$ is a number of seconds that has passed since the midnight of the first January of 1970 to the timestamp t.

Unlike global search engines (that are only able to gather click popularity), a local search engine can benefit from traffic popularity that can be regarded as the number of visits related to each document. This data can be captured on-line. The second option is to extract it from log files. The detailed description of extracting information from web logs can be found in [9]. The main problem with document visits is a cyclic reinforcing of frequently visited documents (since they appear as the first in the ranking, the users visit them, and as a consequence their position is getting higher, so the users click them etc.). One of the methods that weaken a bit this negative trend is a usage of temporal *traffic popularity* that reflects the mean document visits per day. The temporal *traffic popularity* of the document $d=(url_d, t_{0d}, t_{ud}, m_d, H_d)$ for $d \in D$ can be obtained from the following formula: $tp(d) = \dfrac{card(U_d)}{day(t_c) - day(t_{0d})}$ in order to normalize *traffic popularity* it can be divided by the maximal value that permits the *relative traffic popularity* to be obtained: $rtp(d) = \dfrac{tp(d)}{tp_{\max}(d')}$.

New documents do not possess high position in search engines document rankings due to two main reasons. Firstly, there are few pages that point to them (PageRank, HITS). Secondly, they have been visited less times that old ones (DirectHit). Underestimation of new documents' relevance is not a grave problem in the case of global search engines because many documents appear and disappear every day. In contrast, promotion of new documents is a very important task in the local systems.

Detection of new documents is a relatively simple, since almost every search engine contains (or cooperates with) a crawler – a program that periodically visits web pages in order to analyse their content and follow included hyperlinks. For the purpose of this paper a document is considered as new if it has appeared nt days or later (nt is a freshness threshold). *Freshness rate* of document $d=(url_d, t_{0d}, t_{ud}, m_d, H_d)$, $d \in D$ is calculated from the formula: $fr(d) = q^{day(t_c) - day(t_0)}$ where $q \in (0,1)$. For the practical reasons if $fr(d) < \varepsilon$ it is considered to equal 0 ($\varepsilon \approx 0.01$). In case of big portals or vortals in which documents appear and change very often, it is better to use *hour* projection instead of *day* projection. *Stickiness* is the amount of time that users spend on reading/viewing a particular document. Having access to detailed users activity, it is possible to calculate it better than in case of global search engines. The *stickiness* of the document $d=(url_d, t_{0d}, t_{ud}, m_d, H_d)$ in the session $s=(v_1, v_{2,...}, v_p)$, $s \in U_d$ is equal: $stick_s(d) = min(sec(t_{vi+1}) - sec(t_{vi}), timeout)$, where *min* is a function that returns the smaller value from its two parameters, $v_i = (url_{vi}, t_{vi})$, $url_{vi} = url_d$ and $i < p$. *Timeout* was introduced because of experiments which have revealed that there are the sessions in which the *stickiness* is greater than 10 minutes (the mean *stickiness* of the document is normally less than 1 minute). Single sessions with abnormal high *stickiness* influence negatively on the final results.

Due to the fact that *stickiness* of the document d is not defined for those sessions in which the document d is the last one, it has been defined the subset of the set U_d:

$$U_d' = \{s : s = (v_1, v_2, \ldots, v_p) \land \underset{0 < i < p}{\exists} (v_i = (url_{vi}, t_{vi}) \land (url_{vi} = url_d)) \}.$$

The mean stickiness of the document $d=(url_d, t_{0d}, t_{ud}, m_d, H_d)$, $d \in D$ is equal:

$$stick(d) = \frac{\sum\limits_{s \in U'_d} stick_s(d)}{card(U'_d)} \; .$$

This method of *stickiness* calculation (although more precise than in DirectHit) is also laden with errors. First of all, it is not possible to get the *stickiness* of the last document in the session. Moreover, we cannot be sure that a user was really reading the document (and not taking a coffee-break). For all these reasons the *stickiness* should not influence so firmly on a ranking position like, e. g., traffic popularity.

Statistics that visualize from which document the users begin their navigation within the site is widely used in applications that analyses web traffic (for example ROSA, WebTrend). Those documents can be regarded as gateways that link the site with the external world. Thanks to the analysis of three principal ways in which the user can enter the given site it is possible to understand why *session opening* ought to influence positively on the document's ranking position:

- The user has typed the URL address in the browser or has chosen it from the favourites list (the user must consider the document as important).
- The user has followed a link from an external web site (the author of the external page considers the document as important).
- The user has found the document in a global search engine (the user must consider the document as relevant to the query).

Opening rate of the document $d=(url_d, t_{0d}, t_{ud}, m_d, H_d)$, $d \in D$ can be defined as follows: $or(d) = \dfrac{card(U_{fd})}{card(U_d)}$, where $card(U_d) > 0$. For $card(U_d)=0$, $or(d)=0$. $U_{fd} \subseteq U_d$ has the following definition:

$$U_{fd} = \{s : s = (v_1, v_2, \ldots, v_p) \land v_1 = (url_{v1}, t_{v1}) \land (url_{v1} = url_d)\}.$$

Even the most relevant document is useless if it is not available. There are two methods to check document *availability*. The first of them uses a crawler (similarly to new document detection). If the crawler detects that some document is unavailable, it marks this document as deleted instead of removing it from the index. The timestamp of detection is saved to the future calculation of the total time of unavailability. Unfortunately, since crawler activity is resource consuming, web administrators allow the crawler to analyse the site only at nights or with x seconds pauses between 2 documents analysis. As a consequence some document unavailability can be missed.

The second method benefits from web server log files. Web servers usually log the HTTP response code that has been sent for the given request. Codes higher than 400 reflect unavailability of the given resource. It is important to emphasize that if nobody requests unavailable resources, this fact will not be detected. What is worse, the end of the period in which the document has been unavailable is also laden with error (for example the document has reappeared on Monday, but the first request related to it has taken place on Friday).

System ROSA uses its crawler to check documents (first option). *Availability score* is a percent of the document life when the document has been available. *Availability* of the document $d=(url_d, t_{0d,} t_{ud}, m_d, H_d)$, $d \in D$ is equal:

$$av(d) = \frac{t_{ud}}{hour(t_c) - hour(t_{0d})}.$$

Documents that are often updated are considered to have higher quality, but in fact, it is a controversial question that depends a bit on a character of a document. Obviously, main pages of site sections should be often modified to inform the users about new content of the section. On the other hand, frequent modification of "content pages" does not have mean that the document has poor quality. For this reason, the usage of this feature should not have negative impact. It seems that the best practice is to treat each modification as a form of new documents' promotion. Modification rate of the document $d=(url_d, t_{0d,} t_{ud}, m_d, H_d)$, $d \in D$ is defined as follows:

$$mr(d) = \frac{m_d}{hour(t_c) - hour(t_{0d})}.$$

One of most important tasks related to web site administration is the validation of hyperlinks correctness. Hyperlinks to the pages that do not exist or are temporally unavailable (denominated also *dead links*) may cause users' confusion and question site's credibility. Hyperlinks that pointed to all types of resources should be taken into account (not only HTML documents, but also images, cascade style sheet files etc.). Dead link rate of the document $d=(url_d, t_{0d,} t_{ud}, m_d, H_d)$, $d \in D$ can be calculated as:

$dl(d) = \dfrac{card(H_{ud})}{card(H_d)}$ where $card(H_d) > 0$. If $card(H_d) = 0$ then $dl(d) = 0$. $H_{ud} \subseteq H_d$ is a

set of those hyperlinks that have their source in the document d which are unavailable. From technical point of view, dead links detection does not vary from *availability* detection (in both cases the HTTP response code is analysed).

Analysing all factors influencing on a document position in the ranking it is indispensable to mention *click popularity* that measures how many times a particular document was clicked from the search engine result page. It is possible to extend this definition adding recommending system's list as an additional source of click popularity. Please note that click popularity is a part of *traffic popularity*. Moreover, introduction of click popularity restricts the usage of the presented ranking to the search engines and recommendation systems (it would not be possible to use it for deciding which web pages should be crawled more frequently). For all these reasons the ROSA system uses click popularity rather as a feedback that measures the ranking quality. Due to the space limits evaluation of ranking quality will be presented in the future works.

4   Ranking Calculation

All ranking components can be divided into two categories: quality (*availability*, *frequency of modification*, *freshness*, *dead links rate*) and web usage (*traffic popularity*, *stickiness*, *opening rate*) factors. The authors have decided to maintain natural dichotomy in the formula definition (except from freshness). This solution permits the influence of usage and quality factors to be measured.

The usage score of the document $d=(url_d, t_{0d}, t_{ud}, m_d, H_d)$, $d \in D$ is defined as usage$(d) = \max\{rtp(d), fresh(d)\} + \alpha \cdot stick(d) + \beta \cdot or(d)$, where rtp is a relative traffic popularity, fresh is a freshness, stick is stickiness and or is opening rate. Parameters: $\alpha \in (0,1)$, $\beta \in (0,1)$ regulate influence of the stickiness and opening score on the final result. The quality score of the document $d=(url_d, t_{0d}, t_{ud}, m_d, H_d)$, $d \in D$ is defined as follows: quality$(d) = av(d) + \chi \cdot (1-dl(d)) + \delta \cdot mr(d)$, where av is availability, dl is a dead link rate and mr is a modification rate. Parameters: $\chi \in (0,1)$, $\delta \in (0,1)$ regulate influence of the dead link rate and opening score on the final result. RosaRanking of document *d* is a linear combination of *usage score* and *quality score*: *rank(d)* = $\varepsilon \cdot$*quality*(*d*)+ $\phi \cdot$*usage*(*d*). Parameters: $\varepsilon \in (0,1)$, $\phi \in (0,1)$ regulate influence of the *stickiness* and *opening score* on the final ranking measure.

5   Experimental Results

To prove drawbacks of PageRank in a local site, PageRank was calculated for the site of Information Systems Department in Wroclaw University of Technology. The department web site uses the tomcat server that provides the on-line documentation. "The ranking winners" are, in fact, main pages of tomcat documentation, but they are not visited at all. Table 1 contains 11 documents with the highest PageRank and their corresponding visits ranking in November (the tomcat documentation was omitted).

Table 1. Documents with the highest PageRank (tomcat documatation was omitted)

Order num.	PageRank	Number visits	Visit ranking	URL
1	4.26	109	44	/rosasite/welcome.jsp
2	3.39	47	99	/zsi/eng/index.html
3	3.27	130	31	/zsi/index.html
4	3.09	799	3	/zsi/pracownicy/pracownicy.htm
5	3.06	148	23	/zsi/dzialalnosc/dzialalnosc.htm
6	3.02	21	171	/zsi/info.htm
7	3.02	152	21	/zsi/aktualnosci/aktualnosci.htm
8	2.89	293	8	/zsi/dydaktyka/dydaktyka.htm
9	2.86	102	46	/zsi/zaklad/zaklad.htm
10	2.61	27	148	/zsi/kontakt.htm
11	2.57	45	104	/zsi/mapa.htm

There are at least 3 documents that have obtained the high ranking position thanks to the characteristics of the navigation structure (site map - position 11, contact information - 10, general information 6). Additionally, the visits' number of all documents from the table is relatively small (the most visited document had 1403).

Finally, a document ranking using RosaRanking was calculated. All top-ten documents seem to possess high valuable information (Table 2).

6   Conclusions

Presented document ranking method can be used not only to sort search engine result lists, but also to choose the documents that are the most interesting to the user or to

Table 2. Documents with the highest RosaRanking

URL	RosaRanking
/	1
/pracownicy.htm	0.76
/zsi/dydaktyka/usm_program_szczegolowy.htm	0.59
/neuman/java/programowanie.htm	0.41
/neuman/znaczniki/spis.htm	0.40
/neuman/java/skladnia.htm	0.39
/neuman/kierunki/bezpieczenstwo.htm	0.32
/missi2000/referat26.htm	0.27
/zsi/dydaktyka/pmag.htm	0.25
/stopka/laboratorium/laboratorium.html	0.23

restrict the list of documents, in case of calculating the similarity matrix (there is no point in calculating similarity to low-quality documents). The future works will be concentrated in application of this ranking to a recommendation based on the web usage mining, presented in [5]. These works permits the ranking to be evaluated more precisely by means of *click popularity*. Another interesting research field is applying the method presented above to products available in a virtual product catalogue or an e-commerce site.

References

1. Brin S., Page L.: The Anatomy of a Large-Scale Hypertextual Web Search Engine. In: Proceedings of Seventh International World Wide Web Conference (2001) 107-117.
2. Brin S., Page L., Motwani R. and Winograd T.: The PageRank citation ranking: Bringing order to the Web. Technical report, Stanford University Database Group (1998)
3. DirectHit: http://www.directhit.com
4. Gui-Rong Xue et al.: Implicit link analysis for small web search. In: Proceedings of 26th ACM Conference on Research and Development in Information Retrieval (2003) 56-63
5. Kazienko P., Kiewra M.: Link Recommendation Method Based on Web Content and Usage Mining. In: Proceedings of IIPWM´03, Springer-Verlag (2003) 529-534.
6. Kazienko P., Kiewra M.: ROSA - Multi-agent System for Web Services Personalization. LNAI 2663, Springer Verlag (2003) 297-306.
7. Kleinberg J.: Authoritative sources in a hyperlinked environment. Technical Report RJ 10076, IBM, May 1997.
8. Miller J., Rae G., Schaefer F.: Modifications of Kleinberg's HITS algorithms Using Matrix Exponentiation and Web Log Records. In: Proc. of the 24th Annual International ACM SIGIR Conference on Research and Development in IR (2001) 444-445.
9. Mobasher B., Cooley R., Srivastava J.: Data Preparation for Mining World Wide Web Browsing Patterns. J. of Knowledge and Information Systems 1(1) (1999) 5-32

A Propagation Strategy Implemented in Communicative Environment

Dariusz Król

Institute of Applied Informatics, Wrocław University of Technology
Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland
`Dariusz.Krol@pwr.wroc.pl`

Abstract. The need for propagation mechanism in distributed environment is increasing in applications such as distributed databases, collaborative editing of Web documents, cooperative software developments etc. The propagation function required in these applications are often very different from the traditional approach. Hence, a flexible way of specifying and implementing propagation function on individual objects in distributed systems is required. We show how a propagation model with multi-class facilities can be constructed.
The motivation for the presented work is to make the propagation process as easy as possible. As a part of the data manipulation language, we have developed reusable operators for object propagation. The approach can be extended to object evolution in the context of mobile intelligent environment.

1 Introduction

In the last years, many research has been carried out in developing the Migration and Propagation Theory (M&PT). M&PT has emerged to be a powerful tool for solving various practical problems like schema evolution [13] and [17], constraint maintenance [15], supervised learning algorithms [16], activity coordination, collaboration and cooperation [12], error detection [11] and mobile data migration [10]. From the many object architectures proposed, the Multi-class Object Model is found to be effective for solving a number of real integration problems (see [9]). Several researchers have developed propagations algorithms for implementing in different types of distributed and mobile systems. A detailed survey on adequate examples is given in Section 4.

The most important feature for our propagation model is that it enables the use of existing manipulation language. Furthermore, these methods can be shared with other systems that require them, for example object migration and data replication. The intention of the presented work is to remove propagation traditional backward restrictions by focusing on the development of a efficient language extension.

To satisfy the growing needs of such advanced applications, the propagation model should be flexible, general, and extensible. Conventional model designs are neither general nor extensible. The object-oriented frameworks allow new

features of the distributed objects to be incorporated dynamically without much effort.

In order to fulfill the object evolution requirements [3], the value propagation approach is presented. Concerning physical implementation, propagation functions are specified to assure the correct object propagation and allow the user to handle all objects consistently in both backward and forward changing.

The main contributions of this paper are as follows:

1. The proposed evolution strategy is based on a uniform Multi-class Object Model with a solution for constraint definition.
2. In addition to data manipulation language, a definition of forward and backward propagate operator using the same basic algorithms is given.
3. The strategy supports propagation updates with a prototype implementation of SQL statements.

In this work, we present an adaptive propagation operator that can incrementally propagate data changes from the source to the target using multi-object framework as enabling technology. To our knowledge, this propagation strategy is the first that achieves unified value propagation: backward and forward.

The rest of this paper is structured as follows. Section 2 shows an overview of the value propagation example. The adequate operator is described in Section 3. Section 4 cites some related work. Section 5 summarizes the main ideas and give the main topics for the future.

2   Motivating Value Propagation Example

To illustrate the problem, consider the following example. Suppose we have a distributed system of a factory assembling vehicles. The vehicles are built up of components imported from various countries. The schema of all classes are shown in Figure 1, arrows indicate superclass-subclass relationships.
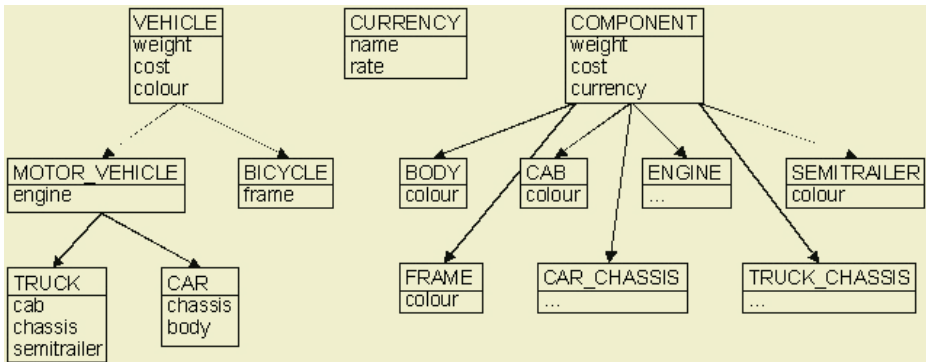
Fig. 1. The schema of objects

Class *Component* has two atomic attributes: *weight* and *cost*. Attribute *cost* stores a cost of production of a component expressed in a currency referred to by local currency attribute. Subclasses of *Component* are *Body*, *Cab*, *Frame*, *Semitrailer*, *Engine*, *Car_Chassis* and *Truck_Chassis*, of which first four contain atomic attribute *color*. Class *Currency* has attributes *name* and *rate*. Attribute *rate* is an exchange rate of a currency to national currency i.e. the currency of a country the factory is located in. Class *Vehicle* has three atomic attributes: *weight*, *cost* and *color*. Its subclasses *Bicycle*, *Motor_Vehicle*, *Truck* and *Car* have attributes: *frame*, *engine*, *cab*, *chassis*, *semitrailer* and *body*, which are references to appropriate objects of class *Component*'s subclasses.

In the presented situation we can see the following examples of value propagation:

1. The color of a vehicle is usually the color of one of its components.
   Eg. the color of a bicycle is the color of its frame, the color of a car is the color of its body and the color of a truck we consider either the color of its cab or its semitrailer (we can see that, it is a matter of a certain convention, which is, however, not important for our research).
   It is an example of a forward value propagation part-whole relationship, that is, in the direction from part to whole, but note, that a property of only one part object is propagated, properties of others are ignored.
   This connection between objects implies that if a property of part object is modified, the whole object should be updated, in order to keep consistency.
2. The weight of a vehicle is a sum of weights of all its components.
   In our simplified example the weight of a bicycle is a weight of its frame, the weight of a car is a sum of weights of its engine, chassis and body, and the weight of a truck is a sum of weights of its engine, chassis, cab and semitrailer.
   It is also an example of a forward value propagation part-whole relationship, but it takes into account properties of all part objects (in fact, all part objects belonging to a certain set).
   This connection between objects implies a need of whole object update in case of modification of any of its part objects (precisely, all part objects belonging to a certain set).
3. The cost of production of a vehicle is a sum of costs of production of all its components but also of transportation and assembling. Since those costs are expressed in different currencies, the overall cost have to be recalculated basing on exchange rates.
   The cost of a vehicle is a function of costs of its components (and other like assembling, etc.), so it is another case of forward value propagation part-whole relationship. In fact, it is the most general case of propagation, which covers two previous examples.
   Additionally, the function depends also on exchange rates i.e. objects of some other class, which is connected with a logical relationship rather than whole-part one (see Figure 2).
4. Either modification of a component object or a currency object implies a need of updating a vehicle object. But also implies a need of change the cost
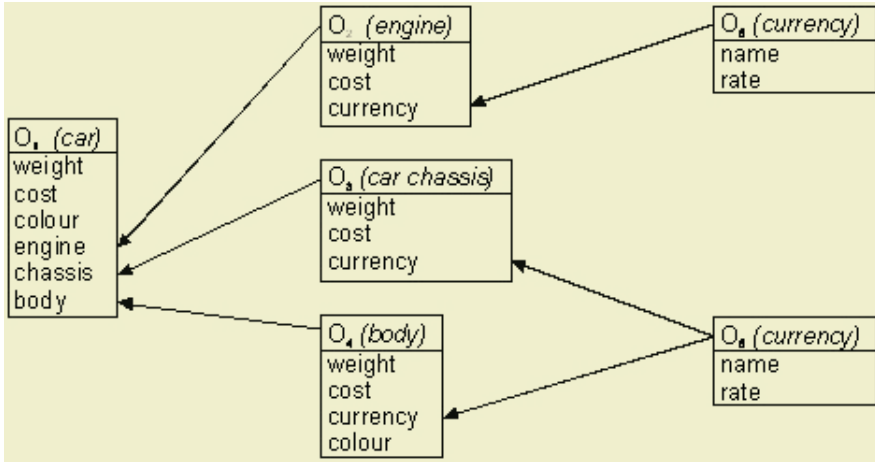
Fig. 2. Example of value propagation

of body, cab, engine, frame etc. This process is an example of a backward value propagation.

3   Value Propagation Operator

The goal of the propagation operator is to produces output a set of incremental changes that can be applied on the target objects. In this section we describe an adaptive propagation strategy for pushing a change from an input data to an output in both directions.

Propagation function must be specified in order to update the instance values. Two kinds of functions must be defined: forward and backward. The former describes the new value of the instances. The latter describes the opposite direction management. These propagate functions are reversible, which means the changes can be freely managed.

Constraints defined on a object composition hierarchy traditionally are maintained in the forward direction. No systems, known to date, provide a mechanism for enforcing integrity constraint in the backward direction along a object hierarchy, which we called the backward propagation.

The fundamental question is for data integrity after propagation process is completed. We use the following operators:

$$\frac{Constraint_1, Constraint_2, ..., Constraint_k}{Forward\ Propagation}\ O_1, O_2, .., O_m \tag{1}$$

$$O_1, O_2, .., O_n\ \frac{Constraint_1, Constraint_2, ..., Constraint_k}{Backward\ Propagation} \tag{2}$$

$$O_1, O_2, .., O_n\ \frac{Constraint_1, Constraint_2, ..., Constraint_k}{Propagation}\ O_1, O_2, .., O_m \tag{3}$$

Given our propagation model, we will now provide rules for forward and backward propagation and for integration with manipulation language.

It seems more natural to specify object dynamics along with the definition of the update to be customized. Thus, we propose to extend SQL so that the update statement includes forward and backward clauses. Towards this, we introduce the *PROPAGATION* condition that can specify which data should be changed by the *UPDATE* statement and how. Essentially, the *PROPAGATION* condition creates the customizable data modifications required by the changed in reality objects.

The extended SQL *UPDATE* statement offers additional but optional clauses such as: *FORWARD PROPAGATION*, *BACKWARD PROPAGATION* and *PROPAGATION*.

Example 1. Value update

$$UPDATE\ My\ car\ SET\ Weight:1400\ WITH\ FORWARD\ PROPAGATION$$

Example 2. Attribute deletion

$$UPDATE\ My\ car\ DEL\ Chassis\ WITH\ BACKWARD\ PROPAGATION$$

Example 3. Association update

$$UPDATE\ My\ car\ FROM\ Truck\ INTO\ Car\ WITH\ PROPAGATION$$

4   Related Work

The research concerned the integration of heterogeneous data has been the objective of many intensive studies. The problem of propagation is rapidly encountered when attempting to coordinate data in a distributed systems.

In addition to well known data integration during schema modifications, a system must reflect changes also in the instances. The system schema cannot be changed without considering what should happen to the stored data. If one of them is changed, both shall be ever consistent one with the other. For fulfilling this requirement, the following steps must be accomplished: identify the objects to be updated and propagate the changes to the instances wherever they are.

Three main approaches have been identified in this aspect: immediate and deferred update propagation and combination the above noted two methods. The former approach initiates an immediate conversion of all objets affected by the change. This causes delays during system modifications, but not during access to objects. The latter generates a conversion program to convert objects into the new schema. This conversion is not immediate, is delayed until an instance is directly accessed, then the conversion is invoked. Deferred conversion causes delays during object access. Most often this approach is used for data reorganization. The third, hybrid approach combines two of the above described methods. For example, changes could be propagated to instances through operation, which is selected by the user.

The dominant implementation of adaptive propagation follows an algorithmic approach in that they propose specific algorithms to compute changes. The authors of [1], and [4] analyze efficiency of the methods analytically and experimentally. At this opportunity several languages have been developed to integrate data sources (see [2]).

This work is based on previous works (see [7], [8], and [9]), which provide modeling of the migration process from one environment to another, but handles migration changes in addition to data updates.

5   Conclusion

The main advantage in the proposed concept is that a new language for manipulation of objects can be formulated on SQL base. From this, it is possible to manage the object dynamics for any given distributed configuration. This paper also present a simple mapping scheme for forward and backward propagation using multi-class object constraint schema.

Data propagation is not restricted only to the following statement: when an update is applied to an object, the update is not just applied to this object only, it may also propagate to its related objects in composition hierarchy according to defined rules, until all reached objects have been updated. Also objects not direct connected could be updated. This process could concern the state or behavior. In this case the multi-class operators could be absolute necessary.

The propagation management involves not only the schema change but also data manipulation operations, for example *UPDATE* statement. Thus, there is a direct relationship between propagation and transaction processing (see [5], and [14]). Since transactions change the state, we have to analyze their effects during propagation to assure correctness and consistency of data.

Currently, we are working on an extension of the present approach in two directions. First, we are specifying a mechanism to control concurrently propagations within the transaction processing. Second, a object definition language with integrated migration, propagation and replication is being specified. As a final result, we intend to propose a complete evolution model with dynamics. In addition, this propagation model will be implemented on top of a commercial system, whose main attempt is to simulate the feasibility of the proposed model.

References

1. Bae M., Hwang B.: An Update Propagation Method Based on the Tree of Replicas in Partially Replicated Databases. Lecture Notes in Computer Science **3320** (2004) 25–29
2. Bertino E., Guerrini G., Merlo I.: Extending the ODMG Object Model with Triggers. IEEE Transactions on Knowledge and Data Engineering **16** (2004) 170–188
3. Castellani X., Jiang H., Billionnet A.: Method for the analysis and design of class characteristic migrations during object system evolution. Information Systems **26** (2001) 237–257

4. Claypool K., Rundensteiner E.A.: AUP: Adaptive Change Propagation Across Data Model Boundaries. Lecture Notes in Computer Science **3112** (2004) 72–83
5. Corradi A.: Parallel Objects Migration: A Fine Grained Approach to Load Distribution. Journal of Parallel and Distributed Computing **60** (2000) 48–71
6. Jurisica I., Mylopoulos J., Yu E.: Ontologies for Knowledge Management: An Information System perspective. Knowledge and Information Systems **6** (2004) 380–401
7. Król D.: Migration mechanism for multi-class object in multi-agent systems. Lecture Notes in Artificial Intelligence **3214** (2004) 1165–1171
8. Król D., Możdżyński A. Indexing Object Propagation in Large Systems. Advanced in Soft Computing (2004) 434–438
9. Król, D., Nguyen N. T.: Migration Mechanism for Multi-class Objects in Distributed Database Systems. International Series on Advanced Intelligence **10** (2004) 239–253
10. Lauzac S.W., Chrysanthis P.K.: View Propagation and Incosistency Detection for Cooperative Mobile Agents. Lecture Notes in Computer Science **2519** (2002) 107–124
11. Lin J.W., Kuo S.Y.: Resolving error propagation in distributed systems. Information Processing Letters **74** (2000) 257–262
12. Mitschang B.: Data propagation as an enabling technology for collaboration and cooperative information systems. Computers in Industry **52** (2003) 59–69
13. Noy N.F., Klein M.: Ontology Evolution: Not the Same as Schema Evolution. Knowledge and Information Systems **6** (2004) 428–440
14. Sapir A., Gudes E. Dynamic Relationships and Their Propagation and Concurrency Semantics in object-Oriented Databases. Lecture Notes in Computer Science **1649** (1999) 94–111
15. Sato H, Makinouchi A.: Temporal Constraints for Object Migration and Behaviour Modeling Using Colored Petri Nets. Lecture Notes in Computer Science **1920** (2000) 80–95
16. Suresh S., Omkar S.N., Mani V.: Parallel Implementation of Back-Propagation Algoritm in Networks of Workstations. IEEE Trans. on Parallel and Distributed Systems **16** (2005) 24–34
17. Tamzalit D., Oussalah M.: A Conceptualization of OO Evolution. Lecture Notes in Computer Science **2817** (2003) 274–278

Using Recommendation to Improve Negotiations in Agent-Based Systems

Mateusz Lenar and Janusz Sobecki

Institute of Applied Informatics, Wroclaw University of Technology
Wyb.Wyspianskiego 27, 50-370 Wroclaw, Poland
{lenar,sobecki}@pwr.wroc.pl

Abstract. In this paper we deal with the problem of non-intuitive and low-efficient negotiations between agents in agent based system. We find recommendation techniques as a suitable method of making negotiation smarter and more efficient. We have introduced into the negotiation thread the following improvements: content-based recommendation and collaborative one. These two types of recommendation are discussed. On the base of a negotiation algorithm we have studied how introduced recommendation methods can improve the whole process of finding mutually acceptable agreements among agents.

1 Introduction to Negotiation in Agent-Based Systems

Agent based system can be defined as a common platform where autonomous agents are running concurrently and independently to achieve their individual goals. There are different types of agents' communication. One of them is negotiation, which not only plays a role in communication level but also in a semantic one. In agent based system the agents' goals are often contradictory and the designer's role is to build such a mechanism, which not only will resolve the conflicts but also do it efficiently and smart. That is why we introduced some recommendation techniques into negotiation mechanism. In this paper we will focus on this problem and propose a new approach.

Etymology of a term "negotiation" is strictly connected with trading. In Latin *negotiari* means to take care about a business. From ancient times people use negotiation techniques in different fields of human life: to bargain on the market, to resolve real-world conflicts, to establish rules of cooperation etc. According to the definition formulated by Nicolas Jennings, negotiation in agent based systems is a process, "by which group of agents communicate with one another to try and come to a mutually acceptable agreement on some matter" [3]. We employed negotiation as a form of extended communication in agent based system because these systems are working as human beings and like in real world, negotiation seems to be a natural and intuitive way of communication.

In the negotiation mechanism we encapsulate protocols, strategies and possible deals [6]. Principles and rules of an interaction between agents are described by a negotiation protocol. The protocol is designed and implemented using existing lower-level communication protocols. The negotiation protocol is focused on a definition of communicates semantics. In the protocol we specify number and types of participants, reachable states and rules saying how to change the state, a sequence of making offers and finally types of negotiation deals.

The negotiation strategy is a set of rules and functions that are used in the negotiation process. The strategy covers aims and manners to reach the mutually acceptable state of the system. We focus on a resource-dependent strategy, where an agent weakens its negotiation position not below specified level in a specified pace.

The formal model of negotiation contains: sets of agents and issues, and a negotiation thread [2]. There are the finite sets of agents and the negotiation issues. The issues are weighted in case of their importance and issues may become continuous or discreet values. A scoring function of an issue plays a main role in offers' creating and evaluating. The offer is a one-dimensional indexed vector, which contains the values of negotiation issues. The scoring function, which evaluates the offer, is defined as follows:

$$F^a = \sum_i w_i^a \cdot F_i^a(x_i)$$

(1)

In formula 1 x_i means the value of an issue i; w_i^a means the relative importance of an issue i to the agent a – all weights are normalized ($\sum_i w_i^a = 1$); F_i^a is a scoring function of the issue i evaluated by the agent a.

The negotiation thread between agents a and b is any finite sequence of the form $\{ x_{a \to b}^{t_1}, x_{b \to a}^{t_2}, ..., x_{a \to b}^{t_n} \}$, where $t_1 < t_2 < ... < t_n$. Interpretation Int of an offer vector x by an agent a at the time t'>t using scoring function V is:

-   accepted, when $F^a(x_{b \to a}^{t_k}) \geq F^a(x_{a \to b}^{t_{k-1}})$;
-   rejected, when there is reached compromise in the other concurrent thread, or one of the negotiation issues has reached its critical values (e. g. time deadline was reached);
-   active, when $F^a(x_{b \to a}^{t_k}) < F^a(x_{a \to b}^{t_{k-1}})$ and there is generated a counteroffer $x_{a \to b}^{t_{k+1}}$.

2   Recommendations in Information Systems

With growing popularity of the web-based systems that are devoted for very differentiated users the recommendation methods are getting also more popular. We can distinguish three basic recommendation approaches: demographic, content-based and collaborative.

In the demographic approach there is used a stereotype reasoning [5] in recommendations and is based on the information stored in the user profile that contains different demographic features [7]. The stereotype reasoning is a classification problem that is aimed at generating initial predictions about the user [5]. The demographic recommendations have however two basic disadvantages [7]: for many users they may be too general and they do not provide any adaptation to user interests changing over time.

Content-based recommendation takes descriptions of the content of the previously evaluated items to learn the relationship between a single user and the description of the new items [7], or in other words a user is supposed to like a new item if the item is similar to other items that are liked by the user [1]. This approach, however has some

disadvantages: it tends to overspecialize its recommendations and is based only on the particular user's relevance.

The collaborative recommendations are able to deliver recommendations based on the relevance feedback from other similar users. The main advantages over the content-based architecture are the following [7]: the community of users can deliver subjective data about items together with their ratings and it is able to offer completely new items to the particular user. Collaborative recommended agents also have some disadvantages: they offer poor prediction when the number of similar users is small and there is a lack of the transparency in the predictions.

The disadvantages of content-based and collaborative recommendation could be overcome by applying the hybrid solution [10]. In this work, however, we propose to apply collaborative approach on the holiday agent's side and contend-based approach on the customer side.

3   System Nntology

Ontology consists of terms, definitions and axioms. So to employ negotiation and recommendation into the process of holiday reservation we have to define domain ontology: The basics of the ontology contain definitions of the following elements: negotiated objects' models (i.e. holiday tours); negotiation participants (travel agent, customer); distances between negotiated objects and customers; rules of interaction.

3.1   Definition of Negotiated Object's Model and Negotiation Participants

Definition of negotiated object consists of the specification of issues and their values of the single distinguished object, i.e. a holiday tour. Each holiday tour is described by a tuple of values (as described below on the figure 1). The values are enumerated or the type of values is given.

Within the negotiation ontology we can distinguish two different participants: a travel agent and a customer one. The customer agent's description consists of three elements: a customer's demographic data (fig. 2), customer's preferences (fig. 1) and a customer's usage data in form of a list of former holidays together with their evaluation (from 0 to 5).

The travel agent's description contains preferences including minimal commission and an agent usage data in the form of a history of the former negotiations.

3.2   Distance Definition Between Holiday's Descriptions and Customer's Profiles

In the recommending systems the user profile could be represented by many different forms: binary vectors, feature vectors, trees, decision trees, semantic networks. In our paper we propose to describe profiles of different negotiation partners or objects as tuples defined in the following way: the finite set A of issues of these elements and the set V contains issues values, where: $V = \bigcup_{i \in A} V_i$ (V_i is the domain of issue i). The tuple p is a function $p : A \rightarrow V$ where $(\forall i \in A)(p(i) \in V_i)$.

Offer-id	Unique identifier
Destination	Set={Egypt-Hurgada, Egypt-Sharm el Sheik, Spain-Gran Canaria, Spain-Tererife, Cyprus-Larnaca, Tunisia-Tunis}
Departure-Airport	Set={Warsaw, Cracow, Wroclaw, Poznan, Katovice, Gdansk, Szczecin}
Hotel-ratings	Set={5 or higher, 4 or higher, 3 or higher, 2 or higher, 1 or higher}
Hotel-amenities	Vector=(fitness room, swimming pool, kid's activities, game room, golf course, dining room, tennis courts, pets allowed, internet access, diving, windsurfing, hotel beach)
Room-facilities	Vector=(bathroom, WC, mini-bar, TV, air-condition, balcony or terrace)
Room-type	Set={single, twin, double, triple, four person}
Number-of-participants	Non zero integer number
Number-of-children-under-15-years-old	Integer number
Start-of-the-holiday	date
Duration	Integer number
Price	Real number
Early-booking-discount	Percent
Commission	Real number

Holiday tour (left label bracket) **Customer preferences** (right label bracket)

Fig. 1. Holiday tour and customer preferences ontology

Customer Demographic data

Age	Integer number
Gender	Set={Female, Male}
Address-zip-code	Zip code

Fig. 2. Customer Demographic data

The distance between profiles of holidays' descriptions and customers' profiles is based on the distances between the values of corresponding issues. The distance function between values of each issue of the profiles is defined as a function $\delta : V_i \times V_i \to R_+ \cup \{0\}$ for all $i \in A$. This function should be given by the system designer and fulfils all the distance function conditions but not especially all the metrics conditions. The distance function for each issue may be different. The values of this function may be enumerated or given in any procedural form.

Despite utilization of distance function we can also apply similarity function $\sigma : V_i \times V_i \to R_+ \cup \{0\}$ that could be transform into any distance function and vice versa in the following way: $\sigma(c,d)=1/\delta(c,d)$ for $\delta(c,d) \neq 0$; $\delta(c,d)=1/\sigma(c,d)$ for $\sigma(c,d) \neq 0$.

The final form of distance or similarity depends on the microstructure of the issue values (i.e. atomic or subset value). Most of the most popular distance or similarity functions are defined for feature vector (binary) or weighted feature vector are Euclidean distance and Cosinus similarity. The distance between tuples l and k could be defined as a simple sum of distances between values of each issue or as a square root of sum of squares of distances (Euclidean distance). We can also indicate the importance of each issue a by using some weights defined as a function $c : A \to [0,1]$ so the distance is determined by the formula:

$$\delta(p_l, p_k) = \sum_{i \in A} [c(i) * \delta^{at}(p_l(i), p_k(i))]$$

(2)

3.3   Customer Clustering Based on the Customer Profile

Clustering problem could be defined as a partition of the given set of profiles into subsets such that a specific criterion is optimized. The criterion may be defined as the average squared Euclidean distance between a profile and the corresponding cluster center. To minimize this criterion we propose to use k-means clustering that partitions the set of the profiles into k non-overlapping clusters that are identified by their centers. However this problem is known to be NP-hard, there are known sub-optimal solutions that are attractive because of its simplicity and flexibility [4].

To solve the k-means clustering we propose to use Lloyd's algorithm [4] that is very popular in the field of data mining and this algorithm's steps are following. First, select randomly k elements as the starting centers of the clusters (centroides). Second, assign each element of the set to a cluster according to the smallest distance to its centroid. Third, recompute the centroid of each cluster, for example the average of the cluster's elements. Fourth, repeat steps 2 and 3 until some convergence conditions have not been met (for example centroides do not change). The attractiveness of this algorithm lies in its simplicity and its ability to terminate, however with many iterations with cost of the step 2 equal to $O(kdN)$, where d is the dimension of each element and N is the number of elements.

3.4   Application of Consensus Methods in the Holiday Travel Negotiation

The consensus methods [8] are used to find collaborative recommendation in the negotiation thread. In order to do this we must have the group of similar profiles of customers G, the customers are grouped using their profiles using k-means algorithm. According to this grouping we may find consensus of the holiday desription that represents the best all customers holiday preferences that belong to this group.

Let j be the index of the a group G member, p_j be the tuple representing the customer profile and r_j is the sub-tuple such that $r_j(i)=p_j(i)$ for all $i \in C$ where $C \subset A$ is a set of holiday description that where accepted by the customer j. Then to find the consensus (recommended holiday preferences for particular group of customers G) we must find the holiday description r that conform the formula 3 [9].

$$\min\left(\sum\nolimits_{j \in G} \delta(r, r_j)\right) \tag{3}$$

This problem could be computationally difficult, but according to [8] we can reduce the computation by finding the minimal value for all issues a of a tuple separately (see formula 4)

$$\min\left(\sum\nolimits_{j \in G} \delta(r(i), r_j(i))\right) \tag{4}$$

The consensus methods are also used to determine the values of the scoring function weights for such issues values that were not specified directly in the customers preferences. The consensus is determined on the basis of the holidays, which the customer spent in the past and evaluated them as being very pleasant.

4   Negotiation Algorithm with Recommendation

The negotiation algorithm with recommendation is based on the following assumptions: the travel agent's goal is to maximize the commission and the customer agent's goal is to maximize the discount and to find the travel offer closest to the customer profile. Customer profile includes weights of each issue. The unwanted values have attached the lowest value. Not evaluated issues have no weights; however they may be determined by means of the content-based recommendation. The customer may mark some issues that cannot be changed during negotiation. The customer sets also the weights of issues (destination – 4 etc) and the weights of issue elements (*Spain-Gran Canaria* – 3 etc). Based on weights and values of issues there is calculated the "best offer", which in the algorithm plays a role of a model offer. Moreover, the customer sets conditions for the "worst satisfying" offer, which is calculated in percent of the "best offer" value. This value gives the customer's agent to know about negotiation threshold. The simplified negotiation algorithm, which is applied both to the customer and the travel agent is presented below:

Step 1. The agent creates the "best offer" and sends it to the other agent. Agent waits for an answer or for the negotiation deadline. Go to step 2.

Step 2. If the negotiation deadline is not reached go to step 3 otherwise go to step 7

Step 3. Evaluate received offer using agent's own scoring function. Go to step 4

Step 4. If the offer is acceptable go to step 5 otherwise go to step 6

Step 5. The agent sends "accept" message. Go to step 8

Step 6. Create counter offer, sent it to the other agent and wait for an answer or for the negotiation deadline. Go to step 2

Step 7. Send "reject" message. End of negotiation with failure.

Step 8. End of negotiation with success.

We employ recommendation in steps 3 and 6. To evaluate the offer (step 3), customer agent uses its own scoring function (defined in paragraph 1) and content based recommendation. This will make this evaluation more realistic, because customer preferences are often incomplete. To create the counteroffer (step 6) the customer agent should consider lower demands, creating a closer offer to the counteroffer using a distance function value and an offer which is closer to the previous customer well evaluated tours and, of courser, a lower discount. To create the counteroffer (step 6) the travel agent should consider a closer offer to the counteroffer based on customer clustering, a reduction of a distance function value between created offer and the counteroffer, preference of an offer which is closer to the previous customer well evaluated tours and a lower commission. In both cases the negotiation strategy establishes the pace of reducing of the scoring function value.

5   Conclusions

In this paper we have presented the multi-agent system that supports bilateral multi-issue negotiation with recommendation applied to the holiday tours booking. Presented model allows us to employ different types of scoring functions to different issues and apply recommendation. The customer's agent is using content-based recommendation to enrich information for scoring function formulation and counter offer formulation. The travel agent is using collaborative recommendation to find such

counteroffer that still being acceptable for the agent is supposed to be acceptable for the customer. However the architecture has not been implemented yet, other implementations [10] have already shown the efficiency of recommended systems.

References

1. Dastani M, Jacobs N,. Jonker CM, Treur J (2001) Modelling User Preferences and Mediating Agents in Electronic Commerce. LNCS 1991:163-193
2. Faratin P, Jennings N, Sierra C (1997) Negotiation decision functions for autonomous agents", Elsevier Science.
3. Jennings N, Lomuscio A, Wooldridge M (2000) Agent-Mediated Electronic Commerce.
4. Kanungo T, Mount DM, Netanyahu, NS, Piatko C, Silverman R, Wu AY (2002) An Efficient k-means clustering algorithm: analysis and implementation. IEEE Tran. On Pattern Analysis And Machine Intelligence 24(7): 881-892.
5. Kobsa A, Koenemann J, Pohl W (2001) Personalized Hypermedia Presentation Techniques for Improving Online Customer Relationships. Knowledge Eng. Rev. 16(2): 111-155.
6. Lenar M, Zgrzywa A (2004) Modelling Multi-Aspect Negotiations in Multiagent Systems Using Petri Nets. Lecture Notes in Computer Science 3029:199-208.
7. Montaner M, Lopez B, de la Rosa JP (2003) A Taxonomy of Recommender Agents on the Internet. Artificial Intelligence Review 19: 285-330.
8. Nguyen NT (2001) Conflict Profiles' Susceptibility to Consensus in Consensus Systems. Bulletin of International Rough Sets Society 5(1/2): 217-224.
9. Sobecki J (2003) XML-based Interface Model for Socially Adaptive Web-Based Systems User Interfaces. Lecture Notes in Computer Science 2658: 1107-1116.
10. Sobecki J (2004) Hybrid Adaptation of Web-Based Systems User Interfaces. Lecture Notes in Computer Science 3038:505-512.

Fault Diagnosis of Discrete Event Systems
Using Place Invariants

Iwan Tabakow

Computer Science Department, Wroclaw University of Technology, Poland
`iwan.tabakow@pwr.wroc.pl`

Abstract. This paper describes a method of using Petri net P-invariants in system diagnosis. The considered discrete event system is modelled by a live, bounded, and reversible place-transition net. The notions of D-partition of the set of places P of a given place-transition net N and net k-distinguishability are first introduced. Two different diagnosis test strategies are discussed and several examples are given.

1 Introduction

The use of Petri net models in diagnosis and reliable design of event-driven systems is a subject of interest to researchers since more than twenty years. In general, the most of the studies in this area focus attention on dynamical analysis concerning specification and implementation of some fault detection, fault diagnosis and/or fault recovery procedures, e.g. using partially stochastic Petri nets [1], or also using trace analysis [6], etc. The study of the system fault indistinguishability properties seems to be important because of the following two reasons. First, we have an additional possibility of describing the critical components of the considered system. Second, there exists a possibility of using some simple and at the same time exact tools for improving the system (self-) diagnosis capabilities in the early stages of its design [2],[9],[10],[11],[12].

The main purpose of this paper is a brief introduction to some rather deterministic fault diagnosis approach concerning the inherent place-transition net fault indistinguishability. This paper describes a method of using Petri net P-invariants in system diagnosis. The considered discrete event system is modelled by a live, bounded, and reversible place-transition net. The notions of D- partition of the set of places P of a given place-transition net N and net k-distinguishability are first introduced. Two different diagnosis test strategies are discussed and several examples are given.

2 Basic Notions

In general any place-transition net $N =_{df} (T,P,A,M_0,K,W)$, where (T,P,A) is a finite net containing sets of *transitions*, *places*, and *arcs* called also *edges*, $K : P \rightarrow (IN_\omega - \{0\})$ and $W : A \rightarrow \mathbb{N}$ are the corresponding *place capacity* and *edge multiplicity* (called also *weight*) functions, respectively. The *initial marking vector* $M_0 : P \rightarrow IN_\omega$, where $\mathbb{N}$ denotes the set of all natural numbers, $IN =_{df} \mathbb{N} \cup \{0\}$, $IN_\omega =_{df} IN \cup \{\omega\}$, and ω is an infinite number such that: $\omega + k = \omega$

and　$k < \omega$　(for any $k \in IN$)　[7],[8]. The *forward marking class* of　N, i.e. [$M_0 >$
$=_{df}$ { $M \in IN_\omega^P$ / $\exists \tau \in T^* (M_0[\tau >M$) }.

In the next considerations we shall assume　N　is　a pure, live and bounded net. In
the case of manufacturing systems the *net reversibility property* is also required. The
net　*P-invariants* are computed using　$\underline{N} \cdot \underline{i} = \underline{0}$, where　$\underline{N}$　is the *PN-connectivity
matrix*　of　N. The *support* of any P-invariant　$\underline{i}$　wrt　N　is defined as follows: supp($\underline{i}$)
$=_{df}$ {$p \in P / \underline{i}(p) \neq 0$ } $\subseteq$ P. Let $\mathcal{I}$ be the set of all (positive) P-invariants of　N
and　$\mathcal{J} \subseteq \mathcal{I}$　is a subset. The　*P-invariant matrix*　of　N　wrt　$\mathcal{J}$　is introduced as
follows: $\underline{\mathcal{J}} : \mathcal{J} \times P \rightarrow IN,$ where　$\underline{\mathcal{J}}(\underline{i},p) =_{df} \underline{i}(p) \in IN$. For convenience only, we
shall assume below that the P-cover　$\mathcal{J}$　of　N　is a set of all positive and minimal P-
invariants. Also we shall use the notion of the *revised P-invariant matrix* of N, de-
fined as: $\underline{\rho} : \mathcal{J} \times P \rightarrow \{0.1\}$, where　$\underline{\rho}(\underline{i},p) =_{df} 1$ iff $\underline{i}(p) \neq 0$ [2]. For simplicity, it
is assumed below　N　have a P-cover. Otherwise, this method is also applicable. In
the last case some additional test points is necessary to be introduced.

3　Net k-Distinguishability and Test Points

Let　[$M_0 >_\alpha =_{df}$ [$M_0 > \cup \{ M_\alpha \}$, where　M_0　is the initial marking　and　M_α
is a marking of　N　such that　$M_\alpha \notin$ [$M_0 >$. We shall say M_α is a *faulty marking*.
Since　$M \cdot \underline{i} = M_0 \cdot \underline{i}$ (for any $M \in$ [$M_0 >$ and　$\underline{i} \in \mathcal{J}$) [7] then $\Delta M \cdot \underline{i} =$
0, where　$\Delta M =_{df} M - M_0$. The last property is satisfied for any P-invariant $\underline{i} \in$
$\mathcal{J}$. Hence we can obtain　$\underline{\mathcal{J}} \cdot \Delta M^T = \underline{0}$. Therefore for $M \in$ [$M_0 >_\alpha$ the above equa-
tion may be violated. Thus we have: $\underline{\mathcal{J}} \cdot \Delta M^T = \underline{a} \in \{0,1\}^{|\mathcal{J}|}$ (for any $M \in$ [$M_0 >_\alpha$
, obviously　$\underline{a} = \underline{0}$　iff　$M \in$ [$M_0 >$). Without losing any generality, below　$(\underline{a})_s$
$\neq 0$　are interpreted　as　$(\underline{a})_s = 1$ ($s \in \{1,...,| \mathcal{J} | \}$). Hence, in accordance with
[4], any $(\underline{a})_s = 1$ will correspond to some subset of places　supp($\underline{i}_s$) $\subseteq$ P having a
(potentially) faulty behaviour. Let　$P \supseteq \Omega(\underline{a}) =_{df} \cap$ supp($\underline{i}_s$) / $(\underline{a})_s = 1 \cap \cap$
supp($\underline{i}_s$)´ / $(\underline{a})_s = 0$, where supp($\underline{i}_s$)´ $=_{df}$ P $-$ supp($\underline{i}_s$)　is the corresponding set com-
plement operation. So, like　[3] the notion of D-partition can be introduced. Below
are used some basic notions given in [9].

Definition 1
By a *D-partition of the set of places*　P　of a given place-transition net　N　wrt　the P-
cover　$\mathcal{J}$　of　N, denoted by　$\Omega(N,\mathcal{J})$, or　Ω　if　N　and　$\mathcal{J}$　are understood, we shall
mean the (multi) family　$\Omega =_{df} \{ \Omega(\underline{a}) / \underline{a} \in \{0,1\}^{|\mathcal{J}|} \}$.

Proposition 1 [9]

(a) $\Omega(\underline{0}) = \varnothing$,
(b) $\forall \underline{a},\underline{b} \neq \underline{0}$ ($\underline{a} \neq \underline{b} \Rightarrow \Omega(\underline{a}) \cap \Omega(\underline{b}) = \varnothing$), and
(c) $\cup \Omega(\underline{a}) / \underline{a} \in \{0,1\}^{|\mathcal{J}|} = P$.

The notion of a k-distinguishable place-transition net under a D-partition of the set
of places　P　of　N　is given in the next definition.

Definition 2
The Petri net　N　is a　*k-distinguishable net　under*　Ω　iff

(i) $\exists \; \Omega(\underline{a}) \in \Omega \; (\; |\Omega(\underline{a})| \; = k \;) \;$ and

(ii) $\forall \; \Omega(\underline{a}) \in \Omega \; (\; |\Omega(\underline{a})| \; \leq k \;)$.

The *support* of any D-partition is defined as follows: $\text{supp}(\Omega) \; =_{df} \{ \; \Omega(\underline{a}) \in \; \Omega \; / \; \Omega(\underline{a}) \; \neq \varnothing \; \}$. Let $\pi(P)$ be the partition generated by the set of subsets of places (i.e. classes), such that each class consists of places having identical columns in the revised P-invariant matrix $\underline{\rho}$ of N (a more formal proof was given in [12]).

Proposition 2 [2]

$\text{supp}(\Omega) \; = \; \pi(P)$.

The above notions of D-partition and net k-distinguishability can be extended to the set of all vertices, i.e. places and transitions of N [9].

Definition 3

Let N be a place-transition net. Then N is a *k-distinguishable net* iff $\exists \; N'$ (N' is a net simulator of N and N' is a k-distinguishable net under Ω').

Definition 4

Let $p_{k_0} \in P$ be a given place of N such that the pre-set $\bullet \; p_{k_0} \; =_{df} \{t_1\}$ and the post-set $p_{k_0} \bullet \; =_{df} \{t_2\}$, where t_1 and t_2 are two different transitions of N. The additional place $p_{k_0} \in \; \bullet t_1 \cap t_2 \bullet$ is said to be a *test point associated with* p_{k_0} iff the initial marking $\hat{M}_0$ of the obtained net $\hat{N}$ is specified as follows: $\hat{M}_0(p)$ $=_{df}$ if $p \; = \; p_{k_0}$ then $\max\{M(p_{k_0}) \, / \, M \; \in [\; M_0 >\} \; - \; M_0(p_{k_0})$ else $M_0(p)$ fi (for any $p \; \in \; \hat{P} \; =_{df} P \; \cup \; \{ \; p_{k_0} \} \;$).

Theorem 1[11]

Let N be live, bounded, and reversible place-transition net and p_{k_0}' be a test point for N. .Then $\hat{N}$ is also live, bounded, and reversible.

4   The Fault Isolation Method

The concurrent systems we are considering here are those that can be represented by a live and bounded place-transition Petri net. The main purpose of the proposed fault isolation method is to locate the physical fault(s) in the Petri net model of the considered system. The single faulty place model will be assumed below (called in short: *p-fault model*). Hence, it is assumed that any faulty marking M_α is a consequence of some p_α. This faulty place will implicate a violation of the firing rule. The violated firing rule will make the P-invariant assertion $\underline{\jmath} \; \cdot \Delta M^T \; = \; \underline{0}$ false. Any validation of the last equation can be interpreted as a validation of the logical value $\in$ {'true','false'} of some two-argument predicate $R(M_k, \underline{i}_s)$, obtained for a given M_k $\in [\; M_0 >_\alpha$ and $\underline{i}_s \; \in \; \underline{\jmath}.$ Any such validation can be represented as an *elementary test* (or *measurement*) $\tau_s \; \in \; \Theta$ of the considered system, where Θ is the set of all such

tests, i.e. $\Theta =_{df} \{ \tau_s / i_s \in \mathcal{J} \}$. The P-invariant matrix $\mathcal{J}$ can be interpreted as a *diagnostic matrix.* Moreover, any such matrix can be considered as an *information system* [5]. Let $\mathbb{R}^{\geq}$ be the set of all nonnegative reals and $c : \Theta \to \mathbb{R}^{\geq}$ be a *cost function* such that $c(\tau) \in \mathbb{R}^{\geq}$ be the *cost* of using the elementary test $\tau \in \Theta$. The *total cost* (in short: TC) in the case of the combinational fault diagnosis approach is given by $TC =_{df} \sum_{\tau \in \Theta} c(\tau)$. The corresponding cost in the case of the sequential fault diagnosis approach will depend on the probability $\text{Prob}\{p\} \in [0,1]$ that $p \in P$ is a faulty place. Obviously $\sum_{p \in P} \text{Prob}\{p\} = 1$. Hence, the *cost of D-tree* $CDT =_{df} \sum_{p \in P} \text{Prob}\{p\} \cdot c(p)$, where $c(p) =_{df} \sum_{\tau \in \Theta(p)} c(\tau)$. By $\Theta(p) \subseteq \Theta$ it is denoted the subset of tests isolating (or locating) fault in p. It can be observed that $CDT \leq TC$. (since $\Theta(p) \subseteq \Theta$). According to Definition 3 the above presented p-fault model can be generalised to the set of all vertices $x \in P \cup T$ of N. In particular, assuming that P is fault-free, the single faulty transition model can be obtained (called in short: *t-fault model*). Let $\mathcal{J}'$ and $\Delta M'$ be $\mathcal{J}$ and ΔM for N' (the net simulator of N). Next by $\mathcal{J}' / X$ and $\Delta M' / X$ we shall denote $\mathcal{J}'$ and $\Delta M'$ restricted to the subset of columns corresponding to X, where $X \in \{P,T\}$ (obviously, we have: $\mathcal{J} = \mathcal{J}' / P$).

Proposition 3 [12]
If a t-fault model is assumed for N then: $\mathcal{J}' \cdot \Delta M'^{T} = \mathcal{J}' / T \cdot \Delta M'^{T} / T$.
Definition 5
The Petri net N is a *p-fault k-distinguishable net* (a *t-fault k-distinguishable net*) iff N is a k-distinguishable net under Ω assuming the p-fault model (t-fault model).

Example 1 (Combinational Fault Diagnosis: P-Fault Model)
Let consider the place-transition net N of Figure 2(a) corresponding to the manufacturing system shown in Figure 1 [13].
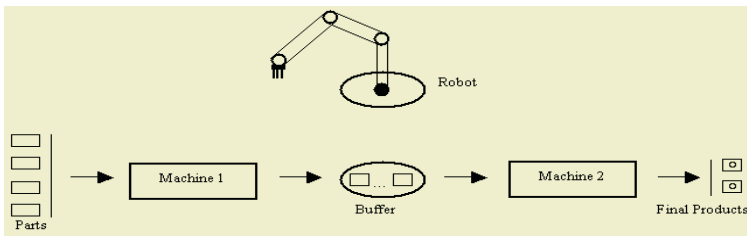
Fig. 1. A simple manufacturing system

The system consists of two different machines, a robot, and a buffer. Every part from the input storage must be processed by Machine 1 first and then by Machine 2 to produce a final product. The robot is used for unloading both machines and the buffer is used to store intermediate parts.
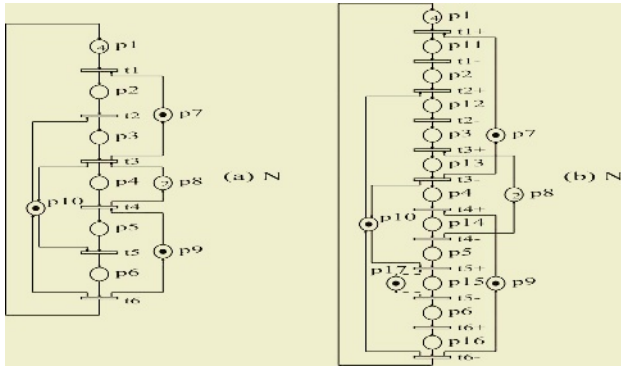
A resource state or an operation

p_1	represents pallets available
p_2	machine 1 loads, fixtures and processes a palleted raw part
p_3	robot unloads an intermediate part to the buffer
p_4	buffer stores an intermediate part
p_5	machine 2 loads and processes an intermediate part
p_6	robot unloads a final product from machine 2, defixtures and returns pallet
p_7	represents the availability of machine 1
p_8	represents buffer available
p_9	represents the availability of machine 2
p_{10}	represents robot available

Start or completion of an event

t_1	models the start of activity of p_2
t_2	the stop of activity in p_2 and the start of activity of p_3
t_3	the stop of p_3 and the start of the storage activity p_4
t_4	the stop of p_4 and the start of p_5 activity
t_5	the stop of activity p_2 and the start of p_6
t_6	models the stop of p_6 activity

(c)Places and transitions interpretation for N

Fig. 2. The Petri net model (a) corresponding to the system of Figure 1, the corresponding net simulator (b) and interpretation (c)

Since $\vartheta' = [\vartheta'/P, \vartheta'/T] =$

$$\left[\begin{array}{cccccccccc|cccccc}
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1
\end{array}\right]$$

the

maximal number of identical columns in ϑ'/P (in ϑ'/T) is 1 (is 2). According to Proposition 2, N is a p-fault 1-distinguishable net and a t-fault 2-distinguishable net. Here the revised P-invariant matrix $\varrho' = \vartheta'$. Let now $p_\alpha =_{df} p_4$ be a single faulty place and $M_\alpha =_{df} (4,0,0,1,0,0,1,2,1,1)$. And so, $\Delta M = M_\alpha - M_0 = (0,0,0,1,0,0,0,0,0,0)$. Hence we can obtain:

$$\vartheta'/P \cdot \Delta M'^T/P = \left[\begin{array}{cccccccccc}
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0
\end{array}\right] \cdot \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right] = \left[\begin{array}{c} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{array}\right].$$

The fourth column of ϑ is obtained and this corresponds to fault in p_4.

Example 2 (*Sequential Fault Diagnosis: T-Fault Model*)

Consider the same net N as in the previous example. The net simulator N´ of N is shown in Figure 2(b). Here $t_i^{\pm}$ are assumed to be fault-free (i = 1, ... ,6). Since N is a t-fault 2-distinguishable, to distinguish between faults in t_5 and t_6, an additional test point p_{17} is placed (as it is shown in Figure 2(b), using dashed line). the process of fault isolation can be carried out step by step, where each step depends on the result of the diagnostic experiment at the previous step. The graphical representation of this approach is illustrated below, where two example D-trees are shown (see Figure 3(a,b) where any $i_s´$ corresponds to $\Delta\hat{M}/T \cdot \hat{i}_s´/T = 0?$, s = 1,2,...,6). Now, to distinguish between faults in p_{15} (i.e. t_5) and p_{16} (i.e. t_6), the additional place invariant $\underline{i}_6´/T$ is used (corresponding to the above introduced test point p_{17}).

(a) D-tree 1   (b) D-tree 2
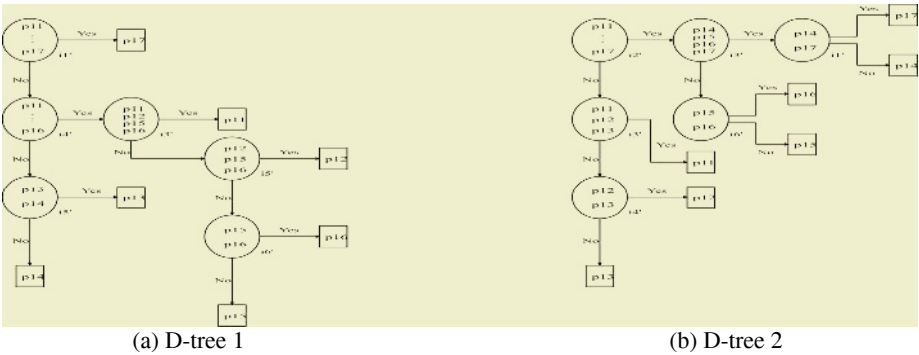
Fig. 3. Two example D-trees for N´ of Figure 2(b)

For example, by assuming an uniform distribution and homogeneous costs we can obtain: CDT / D-tree 1 $= \dfrac{24}{7}$. In a similar manner CDT / D-tree 2 $= \dfrac{20}{7}$.

5   Conclusions

The above-considered approach gives a possibility of fault isolation in concurrent systems. This process is realised by using the Petri net model of the considered system. The degree of accuracy to which faults can be located, i.e. the diagnostic resolution is given in unique way by the obtained k-distinguishability measure. The complexity of the proposed method depends on the efficiency of the existing algorithms for computation of the P-cover, i.e. the set of P-invariants covering N. The choice of diagnosis strategies, i.e. combinational or also sequential is depending on the used time requirements for testing. Moreover, an additional cost-minimisation can be obtained by assuming the considered test point set as a "hardcore". This approach can be extended for higher level Petri nets, e.g. such as coloured nets or also to design self-diagnosable circuit realisations of Boolean interpreted Petri nets.

References

1. Aghasaryan A., Fabre E., Benveniste A., Boubour R. and Jard C., *Fault detection and diagnosis in distributed systems : an approach by partially stochastic Petri nets*. Discrete Event Dynamic Systems *8*, 2 (Special issue on Hybrid Systems), (1998) 203-231.
2. Immanuel B. and Rangarajan K., *System diagnosis and k-distinguishability in Petri nets*. Private communication, India (2001) 14.
3. Mayeda W., *Graph Theory*. John Wiley & Sons,Inc.,New York (1972) 523–557.
4. Murata T., *Petri nets and their applications*. Journal Soc. Instrum. Control Eng. 22,Japan (1983)6572.
5. Pawlak Z., *Rough Sets, Theoretical Aspects of Reasoning about Data.* Kluwer Academic Publishers, Dordrecht, Boston, London (1991) 229.
6. Pietschker A. and Ulrich A., *A light-weight method for trace analysis to support fault diagnosis in concurrent systems*. Journal of Systemics, Cybernetics and Informatics,vol.1 no.6 (2003) 6.
7. Reisig W., *Petri Nets.An Introduction*. Springer-Verlag (1985) 62 – 66.
8. Reisig W., *A Primer in Petri Net Design.* Springer-Verlag (1992) 25 – 33.
9. Tabakow I.G., *Using Petri net invariants in system diagnosis*. Petri Net Newsletter 58,Germany (2000) 21 – 31.
10. Tabakow I.G., *An introduction to the place-transition nets k-distinguishability*. Concurrency, Specification and Programming. Workshop.vol.2,Humboldt-Universität zu Berlin, Germany (2002) 355 – 369.
11. Tabakow I.G., *Using Test Points to Improve the Place –Transition Net k-Distinguishability*. Proc. of the 7th World Multiconference on Systemics, Cybernetics and Informatics SCI 2003, Orlando, Florida USA, July 27-30, vol. IX: Computer Science and Engineering II (2003) 173 - 178.
12. Tabakow I.G., *Using place invariants to isolate faults in concurrent systems*. Systems Science vol.29 no.4, Poland (2003) 99 –112.
13. Zhou M.C. and DiCesare F., *Petri net synthesis for discrete event control of manufacturing systems.* Kliwer Academic, Boston (1993) 233.

A Robust-Fragile Dual Watermarking System in the DCT Domain

Moussa Habib[1], Sami Sarhan[2], and Lama Rajab[2]

[1] Department of Electronics Engineering, King Abdullah II School for Electrical Engineering
Princess Sumaya University, Amman, Jordan
`moussa@psut.edu.jo`
[2] Computer Science Department, King Abdullah II School for Information Technology
University of Jordan, Amman, Jordan

Abstract. The evolution for image manipulation tools and the wide distribution of images through the internet have led to the need for image copyright protection and forgery detection tools. In this paper a three models of a Robust-Fragile Dual watermarking system are proposed for the proof of ownership and image authentication. The proposed system has two main stages the embedding stage and the extraction stage. The embedding of the watermarks, which are the biometric signature for the image's owner, is done in the DCT domain of the cover image. The proposed system with its three models is tested against attacks and it gives satisfactory results in the three models for robustness and fragility and data payload.

Keywords: Robust watermark, Fragile watermark, Spatial Domain, Transform Domain, data payload (capacity), Discrete Cosine Transform (DCT), biometric signature.

1 Introduction

Usage of digital media has witnessed a tremendous growth through the last decades, because of the ease of manipulation and transmission. The ease by which the digital information can be duplicated and distributed has led to the need of copyright protection and proof of ownership and image authentication tools. Digital watermarking is one of the suggested solution for copyright protection and authentication for digital media such as images.

Digital watermarking is the process that embeds data called a watermark into a multimedia object such that the watermark can be detected or extracted later to make an assertion about the object [7]. Digital watermarking has many applications such as broadcast monitoring, owner identification, proof of ownership, content authentication and tamper proofing [2].

Each watermarking system has two main stages: the embedding stage in which the watermark is stored in the cover image and the extraction stage in which the watermark is retrieved from the cover image, in most cases the extraction is the reverse operation of the embedding stage [3].

General image watermarking methods can be divided into two groups according to the domain of application of watermarking. In spatial domain methods as in [5][1], where the pixel values of the cover image are altered in the original image. In transform domain in which the image pixels are altered after applying a certain transform

on the image, a watermark such as DCT [6][8], Fourier, wavelets…etc .In this paper the proposed system is a DCT domain watermarking system.

Another classification of watermarking systems is according to whether the original image exists during the extraction stage or not. If it exists it is called a still image or non-blind watermarking system otherwise it is called a blind watermarking system as in [1]. The proposed system is blind. There are few non-blind watermarking systems, the directives is toward blind watermarking systems.

2 Spotlight on the Watermark

In the early stages of digital watermarking, most of the digital watermarking techniques embed either a character string such as the name of the author   either a sequence of bits as a watermark or the name of the author but this does not guarantee the image source since any one can embed a watermark in the image. But in the most recent stages of digital watermarking biometric traits such as signature in [1] and fingerprints see [4] of the image's owner are used as a watermark this has increased the security and the utility of watermarking techniques. Anoop and Jain [1] have used the online biometric signature of the images owner, which is captured during the signing if the document, this watermark is embedded in the spatial domain of the document. In this paper the two embedded watermark are an offline signature, the signature of the owner is captured, scanned and stored as a binary image.

3 The Proposed Robust-Fragile Dual Watermarking System

3.1 The Evolution of the New Algorithm

The existing watermarking systems that embeds the watermark in the DCT domain of the cover image rounds or truncates the DCT coefficients to the nearest integers before embedding, and ignoring the decimal part. The purpose for this operation is to increase the robustness of the watermark. But there are imperceptible details are lost by this process so these techniques are lossy. In the proposed system this point is taken into account and no truncation is done to the DCT coefficients as a result the system become less lossy.

The proposed system is blind, blockwise .It embeds two watermarks in the DCT domain of the cover image. The first watermark is a robust watermark for the proof of the ownership. The second watermark is a fragile watermark for image authentication, in which any distortion is this watermark gives an indication that the image is forged. The proposed system has three models that differ in the type of DCT coefficients that are used for embedding the watermarks. There are three types as shown in Figure 1.The three models are: Model-A, Model-B and Model-C where Model-A embeds the watermarks in Type I, Model –B embeds the watermarks in Type II and Model-C embeds the watermarks in Type III of DCT coefficients respectively.

Here are some of the necessary concepts and terminologies used in the proposed watermarking system which are introduced to simplify the discussion:

- **I:** the cover image in which the two watermarks will be stored.
- I_{DCT}: the cover image in the DCT domain.
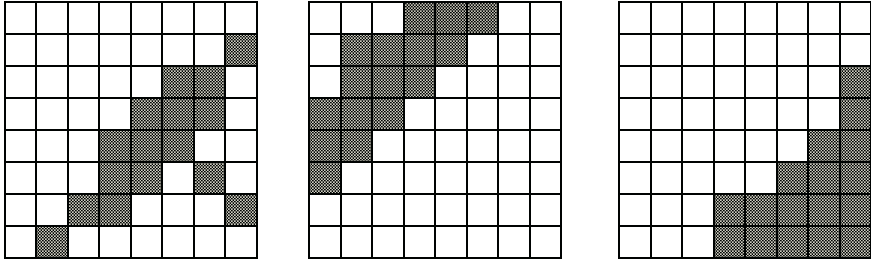
Type I of DCT coefficients.   Type II of DCT coefficients   Type III of DCT coefficients.

Fig. 1. Types of DCT coefficients used for embedding the watermarks

- I_w: the watermarked image after embedding the two watermarks.
- I_a: the image that has the integer part from each DCT coefficient.
- I_b: the image that has the decimal part from each DCT coefficient.
- W_a: the first watermark image (author signature) that will be embedded in I_a.
- W_b: the second watermark image (author signature) that will be embedded in I_b.
- S_i: the sub image with size =8x8.
- I_{wa}: the image of the integer part with the watermark W_a embedded in it.
- I_{wb}: the image of the integer part with the watermark W_b embedded in it.
- I_{wDCT}: the watermarked image in the DCT domain before splitting.

3.2   The Embedding Stage

The input for the embedder is the two watermarks and the cover image. The output for the embedder is the watermarked image.

Before the embedding of the two watermarks the model should be determined Model-A, Model-B or Model-C. The complete steps for the embedding stage are as follows:

Step 1: Partition the cover image **I** into a set of sub images with size =8x8.

Step 2: Apply DCT on each subimage S_i producing a cover image in the DCT domain I_{DCT}.

Step 3: Split the image I_{DCT} into two images, this is done by splitting each DCT coefficient value into two values each stored in an image, the first image I_a has the integer part for each DCT coefficient and the other image I_b has the decimal part rounded to one decimal place. each of the two images I_a and I_b has the same size of the cover image.

Step 4: The image I_a is partitioned into sub images S_i.

Step 5: Watermark W_a is embedded in the I_a image in the 5th bit of each selected DCT coefficient in the chosen model. Each subimage S_i takes 16 bits from the watermark. Note that only the foreground pixels of the watermark are embedded. This process is repeated for each subimage S_i until all the foreground bits in the watermark are embedded within the image I_a to produce I_{wa}.

Step 6: Watermark W_b is embedded in the I_b image in the LSB of each selected DCT coefficient in the chosen model. Each sub image takes 16 bits from the watermark. This process is repeated for each sub image S_i until all the watermark bits are embedded in the image I_b to produce I_{wb}.

Step 7: The two images I_{wa} and I_{wb} are merged into one image to give the watermarked image in the DCT domain I_{wDCT}.

Step 8: Apply the Inverse DCT on the watermarked image I_{wDCT}. The resulting image is the image I_w in the spatial domain.

3.3 The Extraction Stage

The extraction function takes as an input the watermarked Image I_w. The output of the extractor is the two extracted watermarks. The complete steps for the extraction stage are as follows:

Step 1: Partition the watermarked image I_w into a set of sub images with size= 8x8.

Step 2: for each subimage S_i the DCT is applied producing a watermarked image in the DCT domain I_{wDCT}.

Step 3: Split the watermarked image I_{wDCT} into two images I_{wa} and I_{wb} as done in the embedding process.

Step 4: Partition each of the images I_{wa} and image I_{wb} into sub images with size =8x8.

Step 5: Scan the image I_{wa} sub image by sub image and retrieve the 5th bit from each DCT coefficient in the set of DCT coefficients used for embedding the watermark. The extracted watermark here is the robust watermark.

Step 6: scan the image I_{wb} sub image by sub image and retrieve the 1st bit from each DCT coefficient in the set of DCT coefficients used for embedding the watermark. The extracted watermark here is the fragile watermark.

4 Experimental Results

The experimental results for the proposed Robust-Fragile Dual watermarking system with its three models are discussed here. The benchmark Lena 256x256 is used for this test. The two watermarks are binary images for the signature of the images' owner, the first watermark is 128x128 and the second is 64x64. First one is robust watermark, the correlation (similarity measure) between the first extracted watermark and the original one is computed to show its robustness, the correlation is shown in equation 1. On the other hand the second extracted watermark is fragile watermark, any distortion in it gives an indication that the image is forged and inauthentic. The two watermarks are shown in Figure 2.

$$Correlation = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}I'(i,j)\times I(i,j)}{\sum_{i=1}^{N}\sum_{J=1}^{N}I'(i,j)^2} \qquad (1)$$

In the normal test the embedding effectiveness is 100% for both watermarks in the three models. The proposed with its three models system is tested against a set of attacks and the robustness is measured in each case the proposed system gives a satis-

factory results in all cases. Also the fragile watermark in all cases detects that the attacked image is inauthentic. The results are shown in Table 1.
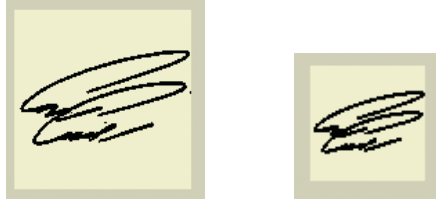
Fig. 2. The two watermarks used in the three models for the proposed system

Table 1. The correlation between the original watermark and the extracted robust watermark after attacks for the three models

The attack	Model A	Model B	Model C
Median Filter	0.999	0.997	0.999
Rotate 90	0.999	0.992	0.999
Rotate 180	0.996	0.994	0.999
Crop &Resize	0.996	0.994	0.999
Salt &pepper (2%)	0.973	0.970	0.970
Gaussian noise	0.960	0.960	0.960
JPEG-25	1	1	1
JPEG-50	0.998	0.994	0.994
JPEG-75	1	0.999	1

5   Conclusion

The proposed system gets benefit from the decimal part of the DCT coefficients, which is found that it is so sensitive to attacks, by embedding a second watermark for the authenticity of the image, and do not neglect this good feature in this part. Also neglecting the decimal part may cause minor changes in the image, which may be Lossy when returning to the spatial domain of the image, so in this system is less Lossy since only one decimal place remains but it is better than truncating it at all. This has also led to the increase of the image payload also selecting only the foreground pixels of the first watermark has increased the payload with the robustness.

References

1. Anoop N., Jain A.: Multimedia Document Authentication using On-line Signatures as Watermarks. To appear in Proc. SPIE-EI 2004, Security, Steganography and Watermarking of Multimedia Contents VI, San Jose, CA, January, (2004)18- 22
2. Cox J., Miller L., Bloom A.: Digital Watermarking.(1st edition), USA: Morgan Kaufman Publishers,(2002)

3. Cox J., Miller L., Bloom. A.: Watermarking applications and their properties. Las Vegas. Nevada, March, (2000)27-29
4. Jain A., Pankanti S., Bolle R. (eds.): BIOMETRICS: Personal Identification in Networked Society. Kluwer, (1999)
5. Lim Y., Xu C., Feng D.: Web Based Image Authentication using Fragile Invisible Watermark. Pan- Sydney Area Workshop on Visual Information Processing. Vol 11, (2002)
6. Saraju M., Ramakrishnan R., Kankanhalli K, Mohan. S.: An Adaptive DCT domain visible watermarking technique for protection of publicly Available image
7. Saraju M.: Digital Watermarking: A Tutorial Review, (1999)
8. Swanson M., Zhu B., Tawfik A.: Transparent Robust Image Watermarking. Proc.of the Int.Conf. on Image Processing,Vol. 111. (1996) 211-214

A Data Hiding Scheme to Reconstruct Missing Blocks for JPEG Image Transmission

Jia Hong Lee[1], Jyh-Wei Chen[1], and Mei-Yi Wu[2]

[1] Department of Information Management, Kun Shan University of Technology
`jhlee@mail.ksut.edu.tw`
[2] Department of Information Management, Chang Jung University
Tainan, Taiwan, R.O.C.

Abstract. This paper presents a new approach for recovering blocks of missing data in the mobile radio environment. When block based compression algorithm such as JPEG are used as part of the wireless transmission process and such images are transmitted over fading channels, the noise effects will usually cause block missing or damaging. The proposed method is based on quantizing the difference of coefficients between two distinct blocks and carefully encoding and embedding them in the least significant bits of other blocks. The proposed scheme can recover high-quality JPEG images from the corresponding corrupted images up to a block loss rate of 20% without carrying any side information. Experimental results show the proposed method outperforms than the method of Shao et al.[1].

1 Introduction

In common wireless scenarios, the image is transmitted over the wireless channel block by block. Due to heavy fading, we may lose an entire block, or even several consecutive blocks of an image. Missing block reconstruction is aimed at masking the missing blocks to create subjectively acceptable images. There are two major methodologies to deal with the missing block recovery. The first one reconstructs blocks using correlation between the lost block and its neighbors. Basically, the techniques are based on a spatial interpolation or temporal interpolation and replacement using a picture correlation property[2-5]. The other one reconstruct missing blocks using self-mbedding technique. Fridrich and Goljan[6] firstly proposed the concept to reconstruct images with self-correcting capabilities. Their method embeds important data of an image in itself as a means for reconstruct the image content. However, their method is applied in spatial domain which cannot be used in the JPEG image transission. There were some researchers used the vector quantization(VQ) techniques to embed referenced block index into its corresponding image block [7-10]. One can reconstruct the damaged block by extracting the embedded index from its correponding masking block and then repair it using the content of indexed code block from the code book. VQ based method can obtain very high quality of reconstructed images, however, the transmission sender and receiver must own the same code book will make the kind of methods unpractical in many real applications.

Shao et al.[1] proposed the self-embedding concept in JPEG images to reconstruct the missing blocks. Their method directly extracting and embedding the DCT coefficients into other blocks can obtain high quality reconstructed images. However, using their method still requires extra side information to reconstruct missing blocks.

In this paper, instead of directly embedding coded DCT coefficients, we embed the difference of coefficients between two distinct blocks and carefully encoding and embedding them in the least significant bits of other blocks. Less bit information is required to embed by using the new embedding scheme since coding the difference of DCT coefficients of two image blocks is more effectively than an entire block of DCT coefficients. In addition, there is no requirement of extra data in the recovery process for missing image blocks.

2   The Proposed Method

Within the proposed scheme, the following issues are addressed: (1) what kind of data should be extracted and embedding into other blocks, (2) where and how should the important data be embedded, (3) how to extract the embedded data to reconstruct damaged blocks.

The proposed scheme diagram is illustrated in Figure 1. For each block b_k in the original image, we will extract important data and then embed the data into its corresponding masking block. To keep a better quality of embedded image, we just modify one-bit data (LSB) of DCT coefficient for each block. Therefore, there are totally 64 bits data quantum can be used to embed important data. We aim to embed the difference of DCT coefficients between b_k and b_w to reduce the embed data length. The proposed new embedding scheme is described as follows:
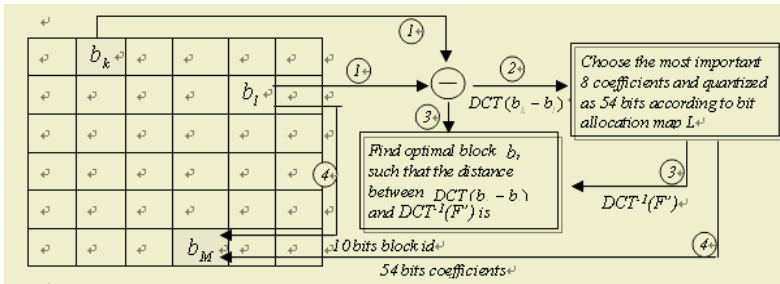
Fig. 1. The flowchart of the proposed data embedding scheme

Step 1:

Let $e_{kl}(i, j)$ represent block difference of DCT coefficient between b_k and b_l with coordinate i-th row and j-th column and is denoted as

$$e_{kl}(i, j) = b_k(i, j) - b_l(i, j) \tag{1}$$

For the block difference e_{kl}, DCT is applied to generate its corresponding frequency block F_{kl}, where

$$F_{kl}(u,v) = \frac{c(u)c(v)}{4} \sum_{i=0}^{7} \sum_{j=0}^{7} e_{kl}(i, j) \cos(\frac{(2i+1)u\pi}{16}) \cos(\frac{(2j+1)v\pi}{16})$$

$$c(x) = \begin{cases} \dfrac{1}{\sqrt{2}} & \text{if } x = 0, \\ 1 & \text{otherwise.} \end{cases} \tag{2}$$

Step2:
Choose the most important 8 coefficients and quantized as 54 bits according to bit allocation map L and the remaining 10 bits is reserved for coding block b_l's id. Only the DC coefficient $F_{kl}(0,0)$ and AC coefficients $F_{kl}(0,1)$, $F_{kl}(1,0)$, $F_{kl}(2,0)$, $F_{kl}(1,1)$, $F_{kl}(0,2)$, $F_{kl}(0,3)$ and $F_{kl}(1,2)$ are reserved. Figure 2 shows the content of bit allocation map L. The quantized coefficients $F'_{kl}(u,v)$ is then defined as

$$F'_{kl}(u,v) = \begin{cases} 0, & if\ F(u,v) = 0, \\ 2^{L(u,v)-1} - 1 & if\ F_{kl}(u,v)/\alpha > 2^{L(u,v)-1}, \\ -(2^{L(u,v)-1} - 1) & if\ F_{kl}(u,v)/\alpha < -(2^{L(u,v)-1} - 1), \\ F_{kl}(u,v)/\alpha & otherwise. \end{cases} \tag{3}$$

Where α is a quantization factor to keep good image quality.

Step3:
Let $\hat{e}_{kl}(i,j)$ be the inverse DCT result of the quantized coefficients $F'_{kl}(u,v)$, where

$$\hat{e}_{kl}(i,j) = \frac{1}{4}\sum_{u=0}^{7}\sum_{v=0}^{7} c(u)c(v)\alpha F'_{kl}(u,v)\cos(\frac{(2i+1)u\pi}{16})\cos(\frac{(2j+1)v\pi}{16}) \tag{4}$$

For each block b_k, we require to compute the minimum distortion between b_k and b_l to maintain a higher reconstructed image quality. An optimal block b_w will be found if a test block b_l satisfies the following condition:

$$\min_{l}\ \sum_{i=0}^{7}\sum_{j=0}^{7}\left|e_{kl}(i,j) - \hat{e}_{kl}(i,j)\right| \tag{5}$$

A block and its masking block should be as far as possible so that the two corresponding blocks will be seldom corrupted simultaneously. Kang and Leou's[8] method is applied in this paper to determine proper positions of masking blocks. The masking block selection rule is described as

$$(p,q) = \begin{cases} (i\times 2 + j + 2, i)\,mod\,32 & if\ 0 \le i \le 15, \\ ((i-16)\times 2 + j + 2, i)\,mod\,32 & if\ 16 \le i \le 30, \\ ((i-31)\times 2 + j + 2, i)\,mod\,32 & if\ 31 \le i \le 32. \end{cases} \tag{6}$$

where (i, j) is the index of considered block b_k and (p, q) is the corresponding index of masking block in the original image with size 256×256.

Step 4:
There are 54 bits data extracted from Step 2 and 10 bits of block id for the optimal block b_w which are embedded into the LSB of the DCT coefficients in the selected masking block b_M.

$$L = \begin{bmatrix} 8 & 7 & 6 & 6 & 0 & 0 & 0 & 0 \\ 7 & 7 & 6 & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Fig. 2. The used bit lengths allocation matrix

3   Experimental Results

Four 256×256 images "Lena," "Baboon," "Jet," and "Peppers" with different block loss rates are used to evaluate the performance of the proposed method. There are some overhead using the data embedding methods to reconstruct damaged image blocks. The image quality will be reduced and the file size will be increased simultaneously. Table 1 shows the cost of overhead using different data embedding methods. However, this kind of data embedding methods embed important data of an image in itself as a means for reconstruct the image content can obtain a better recovery result than traditional interpolation method. Table 2 shows a poor reconstructed result using a neighborhood interpolation method proposed by Ancis et al.[2].

Table 1. The overhead of different data embedding methods (PSNR[dB]/File Size[KB])

	Lena	Baboon	Peppers	Jet
Test jpeg image	40.5 / 21.9	36.7 / 36.8	39.4 / 22.5	40.8 / 20.7
Shao's	31.8 / 26.9	29.7 / 40.2	31.9 / 27.0	32.3 / 24.7
Proposed	34.7 / 30.9	28.6 / 40.1	34.3 / 31.0	34.8 / 28.6

Table 2. PSNR(db) of reconstructed images using the interpolation method by Ancis et al.[2]

Image	Block loss rate				
	3%	5%	10%	15%	20%
Lena	33.6	29.1	26.5	24.5	23.8
Baboon	31.3	29.8	27.1	25.3	24.0
Pepper	32.7	31.3	28.1	25.7	24.1
Jet	30.1	29.1	26.2	24.2	23.1

Figure 3 shows these images corrupted with 20% block loss rate. Figure 4 shows the reconstructed images using the proposed method. We also compare with the method proposed by Shao et al.[1]. Table 3 shows the experimental results with 5%, 10% and 20% block loss rates, respectively. The proposed method performs better than the method of Shao et al. Table 4 shows the experimental results with one, two, and three row of image blocks losing respectively. Our method obviously outperforms than Shao's method.

(a)   (b)   (c)   (d)
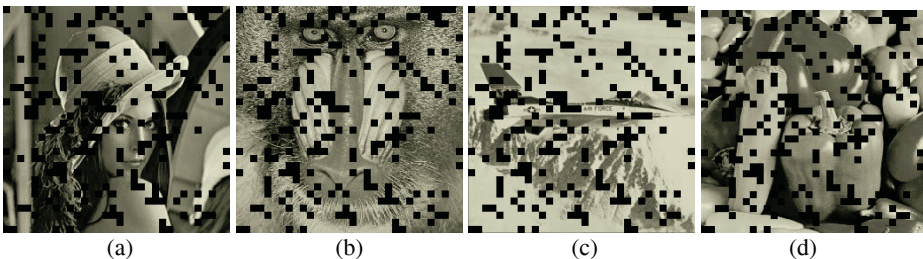
Fig. 3. Damaged images (20% loss) (a) Lena (b)Baboon (c)Airplane (d)Peppers

Figure 5(a) illustrates the corrupted image with three rows of blocks missing. Figure 5(b) displays the reconstructed images using the proposed method and there is no visible artifact in the image at the normal view resolution. Figure 5(c) displays the poor visible artifacts on the place of data losing using Shao's method.
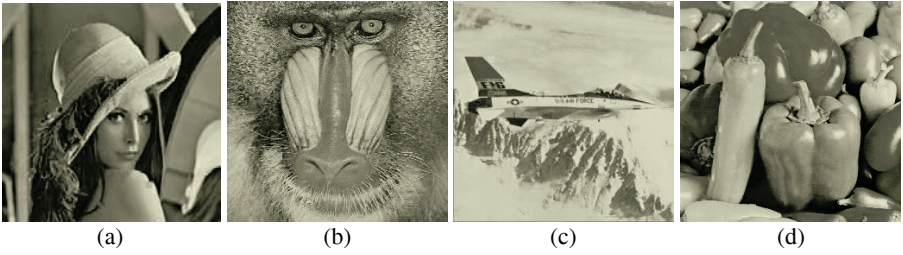
Fig. 4. The images reconstructed from Figure 3 using the proposed method.(a) Lenna (b)Baboon (c)Airplane (d)Peppers

Table 3. PSNR(db) of reconstructed images using different methods

Image	Block loss rate					
	5%		10%		20%	
	Shao's	Proposed	Shao's	Proposed	Shao's	Proposed
Lena	30.50	33.28	29.86	32.43	27.82	29.79
Baboon	28.71	27.76	28.01	26.91	26.71	25.47
Peppers	32.21	34.30	31.72	32.72	29.82	31.60
Jet	31.41	33.97	29.28	33.15	28.02	30.24

Table 4. PSNR(db) of reconstructed images using different methods

Image	Number of losing block Rows					
	one		*two*		*three*	
	Shao's	Proposed	Shao's	Proposed	Shao's	Proposed
Lena	31.34	34.32	30.59	33.40	28.89	32.17
Baboon	29.17	28.19	28.69	27.65	28.26	27.14
Peppers	31.07	34.12	30.34	33.64	29.16	33.34
Jet	32.35	34.79	32.21	33.92	32.10	32.72

Fig. 5. The performance of different recovery methods; (a) the image damaged with lossing three rows of blocks, (b) reconstructed result using the proposed method, (c) reconstructed result using Shao's method

4   Conclusions

This paper presents a new approach for recovering blocks of missing data for jpeg image transmission. The proposed method is based on quantizing the difference of coefficients between two distinct blocks and carefully encoding and embedding them in the least significant bits of other blocks. Although it will cost some overhead in image quality and file size using the proposed scheme, it can recover a higher quality JPEG images than traditional interpolation methods. In addition, the experimental results show the proposed method outperforms than the method of Shao et al.[1].

References

1. Y. Shao, L. Zhang, G. Wu and X. Lin, "Reconstruction of missing blocks in image transmission by using self-embedding," Proceedings of International Symposium on Intelligent Multimedia, Video and Speech Processing, May, 2001, pp.535-538.
2. M. Ancis and D.D. Giusto, "Reconstruction of Missing Blocks in JPEG Picture Transmission," IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Aug. 1999, pp. 288-291.
3. S. S. Hemami and T. H.Y. Meng, "Transform coded image reconstruction exploiting inter-block correlation," IEEE Transactions on image processing, 1995, vol. 4, PP. 1023-1027.
4. S.D. Rand, G. Sapiro and M. Bertalmio, "Structure and texture filling-in of missing image blocks in wireless transmission and compressing applications," IEEE Transactions on Image Processing, 2003, 12(3), pp.296-303.
5. W.Y.F. Wong, A.K.Y. Cheng and H.H.S. Ip, "Concealment of damaged blocks by Neighborhood Regions Partitioned Matching." Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2001, Vol. 2, pp.45-48.
6. J. Fridrich and M. Goljan, "Images with self-correcting capabilities," Proceedings of IEEE International Conference on Image Processing, 1999, Vol. 3, pp.792-796.
7. K. L. Hung and C. C. Chang, "Reconstruction of lost blocks using codeword estimation," IEEE Transaction on Consumer Electronic, 1999, vol. 45, PP.1190-1199.
8. L.W. Kang and J.J. Leou, "A new error resilient coding scheme for JPEG image transmission based on data embedding and vector quantization," Proceedings of the International Symposium on Circuits and Systems, 2003, Vol. 2, pp.532-535
9. N. Shacham and P. McKenney, "Packet recovery in high-speed networks using coding and buffer management," Proceedings of IEEE Computer and Communication Societies, June 1990, vol. 1, pp.124-131.
10. H.C. Wei, P.C. Tsai and J.S. Wang, "Three-sided side match finite-state vector quantization," IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(1), pp.51-58.

Adaptive Selection of Coefficient's Portions in the Transform Domain Watermarking

Yoon-Ho Kim[1], Hag-hyun Song[2], and Heau Jo Kang[3]

[1] Div. of Computer & Multimedia Content Eng. 800, Doan-Dong
Mokwon University, Taejon, 305-729, Korea
`yhkim@mokwon.ac.kr`
[2] IITA, Taejon, Korea
`hhsong@iita.re.kr`
[3] Div. of Computer & Multimedia Content Eng. 800, Doan-Dong
Mokwon University, Taejon, 305-729, Korea
`hjkang@mokwon.ac.kr`

Abstract. This paper investigates the transform domain-based adaptive watermark algorithm based on fuzzy inference. In addition, our approach has double-aim in first to enhance the robustness of the watermarking and in second to reduce the complexity of the fuzzy inference rule. Not only statistical characteristics (SC) but also human visual system (HVS) parameters, such as contrast sensitivity, texture degree and average value are involved to select the optimal coefficients region. Rule based fuzzy inference map is designed by utilizing a possible relation between HVS and SC. The performance analysis of the proposed approach is evaluated by the experiments of imperceptibility and correctness of watermark. Experimental results showed that proposed scheme has a better PSNR and has stronger response than previous results of Cox et al.

1 Introduction

In the knowledge-based information society, all most of all the multimedia productions are digital form. These types of information are easy to be stolen by way of transmitting on networks. Therefore, the protection and enforcement of intellectual property rights for digital media has become an important issue. A commonly used scheme is to insert digital watermark into the media such as audio, image and video, which can be detected and used as an evidence of copyright. Watermarking strategies divided into two major categories: transform-domain and spatial-domain approach [1], [2]. Transform-domain watermarking methods, also called multiplicative watermarks, such as Fourier transform, discrete cosine transform, discrete wavelet transform(DWT), are generally considered to be robust against attacks [3], [4]. The watermark is hidden in the middle or lower frequency band is more liable to be suppressed by compression. Therefore, for watermarking, how to select the best frequency band of the image is more important than any other procedures [5], [6].

In this paper, we decompose the original image into the frequency domain with hierarchical sub-bands using DWT. The main goal of this method is to utilize a fuzzy inference system (FIS) in order to select optimal coefficients region. Further more, HVS and SC are also used to generate an effective fuzzy inference rules. These make it possible to enhance the robustness in watermarking with respect to the different types of attacks.

2 Analysis of HVS by DWT Coefficients

There are two most common requirements for effective watermarking. It should be perceptually invisible, which means the watermark is not visible under typical viewing conditions. It should also be robust to common signal processing and intentional attacks. In general, transform domain based approached are superior to that of spatial domain in preserving contents fidelity and robustness under attacks. The main issue, in these applications, is to select the optimal range of frequency coefficients to hide a watermark. DWT produced watermarked images which are visually better due to the absence of blocking artifacts. The DWT coefficients are organized into wavelet sub-tree. It provides hierarchical multi-resolution: three part of multi-resolution representation and a part of multi-resolution approximation. Namely, consider a 2-D image, wavelet decomposition will result a pyramid structure and it can be expressed as

$$A_{2^{j+1}}f = \sum_{k}\sum_{l} h(2m-k)h(2n-l)A_{2^{j}}f$$

$$H_{2^{j+1}}f = \sum_{k}\sum_{l} h(2m-k)h(2n-l)A_{2^{j}}f$$

$$V_{2^{j+1}}f = \sum_{k}\sum_{l} h(2m-k)h(2n-l)A_{2^{j}}f \qquad (1)$$

$$D_{2^{j+1}}f = \sum_{k}\sum_{l} h(2m-k)h(2n-l)A_{2^{j}}f$$

where the sub-bands $A_{2^{j+1}}f$, $H_{2^{j+1}}f$, $V_{2^{j+1}}f$ and $D_{2^{j+1}}f$ represent the low-low, low-high, high-low and high-high sub-bands in resolution level n. Generally, the HVS is not sensitive to detect changes in such bands as $H_{2^{j+1}}f$, $V_{2^{j+1}}f$ and $D_{2^{j+1}}f$.

The HVS has identified by several phenomena which is related with spatial resolution, intensity resolution, and intensity sensitivity and so on. The visual properties of HVS can be described in many terms. Contrast sensitivity and texture sensitivity are main parameters of HVS model: the one means variance or difference of the pixel's brightness. The more contrast property the coefficients region has, the stronger the embedded watermark could be. The other means the perceptivity of sine wave. The more complex the background is, the larger the watermark could be. Contrast sensitivity function can be expressed as

$$CSF = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad (2)$$

Where I_{max} and I_{min} denotes the max. and min. brightness for selected region respectively. The formulation of the texture degree $c(B_k)$ for given region can be summarized as follows:

$$c(B_k) = \frac{1}{n^2} \sum_{(i,j)\in B_K} p(Ag_k) \cdot \frac{|x(i,j) - Ag_k|}{Ag_k} \qquad (3)$$

$$p(Ag_k) = (1/Ag_k)^a$$

Where Ag_k is the average gray value for each block region; degree $p(Ag_k)$ is the weighted coefficient.

3   Proposed Watermarking Algorithm

3.1   Watermark Embedding and Extraction

After transformation, the watermark is hidden in the low- or middle-frequency coefficients region. Therefore, it is important for us to select the best frequency regions of the image. Fig. 1 describes the block diagram of the proposed watermarking algorithm. In order to embed the generated watermark into the selected wavelet coefficients, an area of image size of N*N at random location has been chosen from the wavelet transform image. The procedure of watermark extraction is a just inverse that of embedding, and the watermark embedding process can be express as follows:

  Step 1: Set the original image size of 256*256
  Step 2: Perform the 1-level DWT
  Step 3: Create the random number sequence matrix with value of {-1, 0, 1}
  Step 4: Calculate the HVS and SC values in sub band of cH, cV using Step 2
  Step 5: Generate the fuzzy production rules based on Step 4
  Step 6: Find the suitable coefficients range in sub-band using fuzzy rule
  Step 7: Combine the fuzzy inference engine with these RNS
  Step 8: Embed the additional energy to suitable coefficients range
  Step 9: Save the embedded image

3.2   Fuzzy Inference Rules by HVS and SC

One important feature of our algorithm is it's capability to adaptively calculate the number of region to be watermarked. Therefore, the fuzzy inference rules and membership functions were developed based on both characteristics of HVS and statistical parameters. These characteristic variables can be represented by a number of fuzzy-set values [7].

  Suppose that the following two rules with two-input and single-output variable are in the model for HVS and SC:

$$
\begin{aligned}
&If\ \ C\ \ is\ \ ZE\ \ then\ \ HVS\ \ is\ \ NM \\
&If\ \ T\ \ is\ \ NB\ \ then\ \ HVS\ \ is\ \ NM
\end{aligned}
\tag{4}
$$

Where C and T are the input variables with respect to the contrast sensitivity and texture degree; HVS is the out variable; and ZE, NB, and NM are the triangular type membership functions in the fuzzy set. These two can be formulated by a single, combined rule

$$
If\ \ C\ \ is\ \ ZE\ \ or\ \ T\ \ then\ \ HVS\ \ is\ \ NM
\tag{5}
$$

  Consequently, we can expand or compress a fuzzy rule into a new model both in considering input and output variables. How to design the membership function and input variables which determine the number of rules is another important task, which mainly depends on the human experts. Designed fuzzy inference rule is shown as in Table 1. This fuzzy association map means that the bigger the C and T, the less sensitive the human visual system, so the stronger the strength of watermark embedded possibly. By using the fuzzy rules in Table 1, we can insert the watermark into the DWT coefficients to obtain the watermarked image

Watermark encryption

Image watermark → W → Plaintext sequence → W_p → ⊕ → Encrypted watermark → W_e

Watermark embedding

Original image → 1- Level DWT → Selecting optimal block in every sub-block → Watermark insertion → Watermark image

Fuzzy Inference Engine — Outputs of fuzzy controller

Watermark extracting

Watermark extraction → IDWT → Extracting encrypted watermark

Watermark decryption

Extracted image watermark ← W' ← Decrypted code sequence ← W'_p ← Extracting encrypted watermark

Fig. 1. Block diagram of the proposed algorithm

Table 1. Fuzzy association map for the proposed method

HVS		C					SC		Ave.				
		NB	NM	ZE	PM	PB			NB	NM	ZE	PM	PB
	NB	NB	NB	NM	ZE	ZE		NB	NB	NB	NM	ZE	ZE
	NM	NB	NM	NM	ZE	ZE		NM	NB	NM	NM	ZE	ZE
T	ZE	NM	NM	ZE	PM	PM	SD	ZE	NM	NM	ZE	PM	PM
	PM	ZE	ZE	PM	PM	PB		PM	ZE	ZE	PM	PM	PB
	PB	ZE	ZE	PM	PB	PB		PB	ZE	ZE	PM	PB	PB

4   Experimental Results

In order to evaluate the proposed scheme, some numerical experiments are performed. Experimental images are shown in Fig. 2. Each has a size of 256*256 and the gray level is 256, respectively. The watermark image is a binary image of size 32*32.

To evaluate the imperceptibility of the watermarked image, the peak signal-to-noise ratio (PSNR) is used. Assuming that the original image W and the watermarked image W` both have image size N*M. The mean square error between W and W` is defined as follow:

(a)   (b)   (c)
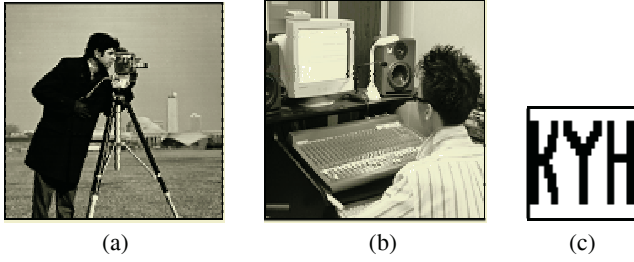
Fig. 2. Experimental images: (a) "cameraman", (b) "Musician", (c) "watermark"

$$A = 255^{2}$$

$$B = \frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \left[W(x,y) - W(x,y) \right]^{2} \tag{6}$$

Consequently, PSNR can be calculated by Eq. (7)

$$PSNR = 10 \log_{10}(A/B)[dB] \tag{7}$$

For detection or verification, it needs to verify if a specific watermarking pattern exists or not. In this case, a correlation is often used for full extraction of the watermark. The correlation r between the embedded image and the original image, can be calculated by

$$r = \frac{\sum_{m} \sum_{n} (A_{mn} - \overline{A})(B_{mn} - \overline{B})}{\sqrt{\sum_{m} \sum_{n} (A_{mn} - \overline{A})^{2} \sum_{m} \sum_{n} (B_{mn} - \overline{B})^{2}}} \tag{8}$$

Where A and B stand for the original image and embedded image, respectively, and $\overline{A}$ and $\overline{B}$ are the mean of the element of the each matrix (A) and matrix (B). We use Eq. (2), (3) to calculate the contrast sensitivity and text degree. Figures 3(a), 3(b) showing these HVS distribution corresponding to the each blocks in each sub-band named cH and cV, Table 2.

(a)   (b)
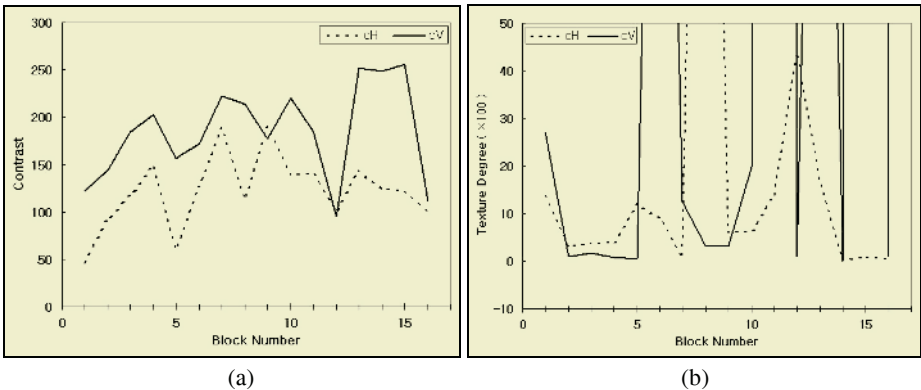
Fig. 3. HVS distributions for each sub-band: (a) contrast sensitivity, (b) texture degree

The performance analysis of the proposed approach can be evaluated based on the experiments of imperceptibility and correlation. It is also used destruction attack of JPEG compression and noise to measure the robustness. Table 2 shows the PSNR and correlation results after different factors of JPEG compression. With regard to noise attack, it is added in 20, 40, 60 and 80% to the watermarked image. The results listed in Table 3.

Table 2. Detection results of the JPEG compression attack

Method	Cox et al's method				Proposed method			
	Cameraman		Musician		Cameraman		Musician	
JPEG Ratio	PSNR [dB]	Corr (%)	PSNR [dB]	Corr (%)	PSNR [dB]	Corr (%)	PSNR [dB]	Corr (%)
10 %	11.47	82	11.29	82	15.79	94	15.79	94
30 %	10.70	79	10.81	80	14.31	91	14.79	92
50 %	10.51	78	10.56	79	12.39	86	12.25	85
70 %	9.97	76	10.23	76	9.73	73	9.69	73
90 %	9.53	73	9.46	72	6.30	40	6.90	49

Table 3. Detection results of the noise attack

Method	Cox et al's method				Proposed method			
	Cameraman		Musician		Cameraman		Musician	
Noise	PSNR [dB]	Corr (%)	PSNR [dB]	Corr (%)	PSNR [dB]	Corr (%)	PSNR [dB]	Corr (%)
80 : 1	11.02	80	10.89	79	14.66	92	14.19	91
60 : 1	10.56	78	10.42	78	12.47	86	11.47	83
40 : 1	8.70	66	8.96	67	9.69	74	8.83	68
20 : 1	6.56	44	7.03	51	7.29	53	7.01	50

5   Conclusions

In this paper, transform domain-based watermarking algorithm which can adaptive select the watermark region is presented. Our approach has double-aim in first to enhance the robustness of the watermarking and in second to reduce the complexity of the fuzzy inference. Fuzzy inference algorithm is designed to select optimal coefficients region, and SC are also utilized as well as HVS parameters.

The performance analysis of the proposed approach is evaluated by the experiments of imperceptibility and correctness of watermark. Experimental results showed that proposed scheme has a better PSNR and has stronger response than previous results of Cox et al.

References

1. C .Busch, W. Funk and S. Wolthusen,: Digital Watermarking: From Concepts to Realtime Video Applications. IEEE Computer Graphics and Applications, Vol.19 (1999) 25-35
2. M. Barni, F. Bartolini, V. Cappellini, N, Checcacci,: Object Watermarking for MPEG4 Video Streams Copyright Protection. Proc. Of the SPIE In Security and watermarking of Multimedia Contents, Vol.3971. (2000) 465-476

3. A. Piva, R. Caldelli, A. Rosa,: A DWT-based Object Watermarking System for MPEG4 Video Streams, Proc. of IEEE International Conference Image Processing. (2000) 5-8
4. C. Hsu and J. Wu,: Digital Watermarking for Video, Proc IEEE International Conference on Digital Signal Processing, Vol.1. (1997)
5. I. J. Cox and M. L. Miller,: A review of watermarking and the importance of perceptual modeling, Proc. SPIE Conf. on Human Vision and Electronic Imaging , Vol. 3016. (1997) 92-99
6. D. C. Lou, T. L. Yin,: Adaptive digital watermarking using fuzzy logic techniques, the journal of the Society of Photo-optical Instrumentation Engineers, Vol. 41, (2002) 2675-2687.
7. Y. H. Kim and H. H. Song,: An Adaptive Digital Watermarking using DWT and FIS, Journal of Digital Contents Society, Vol. 5. (2004) 128-131

Copyright Authentication Enhancement of Digital Watermarking Based on Intelligent Human Visual System Scheme

YangSun Lee, Heau Jo Kang, and Yoon-Ho Kim

Division of Computer & Multimedia Content Eng. 800, Doan-Dong, Taejon
Mokwon University, Taejon, 305-729, Korea
{yslee,hjkang,yhkim}@mokwon.ac.kr

Abstract. Recently, According as digital technology with knowledge base society develops, illegal copy by user's vulnerability of copyright authentication becomes frequent. Therefore, it need research about user's copyright authentication method. In this paper, a watermark inserting adaptive method according to the sensitivity of human visual system (HVS) in DCT transform domain is proposed. Also, we proposed watermark techniques using adaptive voice information(VI). It effects that get in original image than method to use conventional image is less and shows robust watermark restoration ability form outside attack. Performance analysis is achieved about still image and restoration of watermark information using OFDM/QPSK still image transmission system in wireless channel environment. Analysis result, VI watermark performance that influence in original image is very small. And it could know that show high restoration performance.

1 Introduction

While digitization phenomenon of multimedia data offers convenience in edit of multimedia data, transmission and save in the Internet recently, but it can generate problems such as copyright dispute and militarism copy etc. Also, according as demand of multimedia service through radio internet increases by technology of communications and rapid development of computer, copyright protection about data and problems for certificate were shown.. Therefore, research also about digital watermarking is consisting vigorously by method to protect copyright of digital multimedia data but yet perfect solution come out[1]-[4].

In this paper, we propose watermark techniques to use Voice Information (VI) as watermark data that is inserted on still image. And, we analyzed effect that watermark techniques to propose gets in original image. Also, we design OFDM/QPSK system for radio multimedia service and analyzed about original image and watermark restoration from outside attack in radio channel environment.

2 Watermarking Technique

2.1 Audio Watermark Scheme

The established audio watermarking is to insert and extract the additional information like copyrights in audio signals, preventing it from being heard. It would insert the

watermark on the audio signals by changing the original signal itself or putting certain properties. The ways of making this audio watermark include Low-bit Coding, Phase Coding, Spread Spectrum and Echo Hiding.

2.2   Image Watermark Scheme

By the method of inserting watermark or the applied technique, they can be divided as one for inserting the data by topological orientation and by the orientation defined with frequency range, beside these, there is invisible watermarking and watermarking by the produced key values. In this paper, watermarking technique in the field of frequency is used as a method in common recently.

2.2.1   Watermarking Scheme Using Frequency Domain in HVS System

Mostly used is recently the watermarking technology which changes the data by frequency space conversion. This has more advantages than the watermarking through space analysis. The frequency-using way converts the multimedia data into analog signals with frequency components. It also converts the watermark into analog signals and inserts it on.

Generally for the methods transforming the data, there are DCT, FFT, and Wavelet Transform. This scheme has its advantage to make hard to erase the watermarks once inserted as the watermark factors are distributed over the entire region of the data.
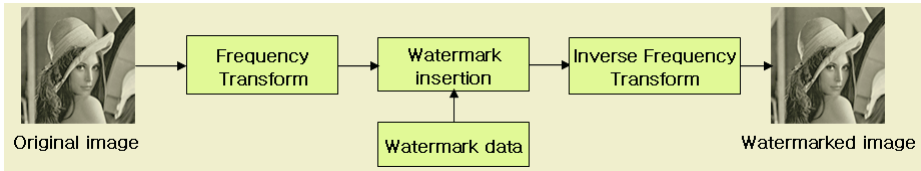
Fig. 1. Watermark insertion in frequency domain

Figure 1 illustrates the general procedure for frequency domain watermarking.

2.3   VI(Voice Information) Watermark Scheme Using Intelligent HVS Scheme

In this paper, we suggest a watermark technique which uses voice information for the watermark data to be inserted into still image. Effect that voice information deals damage in original image more than thing which inserted image such as logo image to water mark data is very small. Because of, frequency band is low than original image. And it have advantage that can have more robust performance than watermark insertion that use existent image when extracted and reconstructed watermark information. In the case of voice data, because restoration of voice information is available even if have the 10^{-3} BER. Therefore, we can improve all invisibility and robustness in Trade-Off relation in watermark technology.

Figure 2 illustrates the procedure insert PCM modulated voice data after frequency transform using DCT scheme. Figure 3 shows extract process of inserted watermark data.
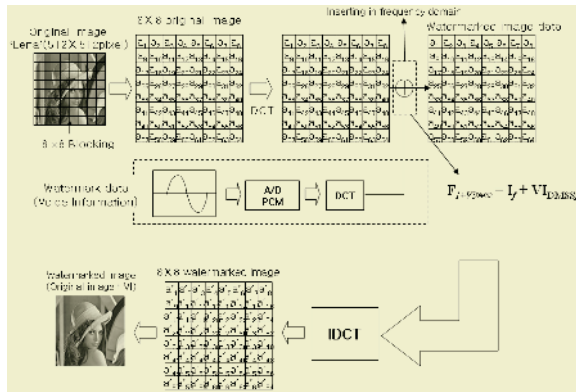
Fig. 2. Watermarking embedding process
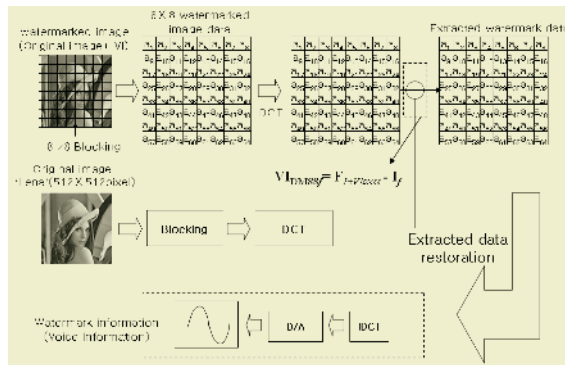
Fig. 3. Extract process of watermark data

3   OFDM/QPSK System for Image Transmission

3.1   OFDM/QPSK System

Figure 4 and figure 5 illustrate the structure of OFDM/QPSK system. The N serial data elements (spaced by $T = 1/f_s$ where f_s is the symbol rate) modulate N subcarrier frequencies, which are then frequency division multiplexed.

In the transmitter, the transmitted high-speed data is first converted into parallel data of N subchannels. Then, the transmitted data of each parallel subchannel is modulated by PSK-based modulation. The transmitted signal is

$$s'(t) = s(t)e^{j2\pi f_c t} \tag{1}$$

Figure 5 illustrates the general procedure for receiver of OFDM/QPSK system. The received signal can be expressed as

$$r'(t) = s'(t) + bs'(t-\tau)e^{j\theta} + n(t) \tag{2}$$

Where, n(t) is AWGN with two-sided power spectral density $N_o/2$.
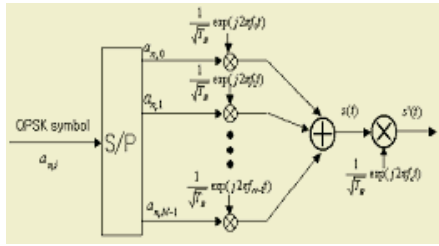
Fig. 4. The transmitter of OFDM/QPSK system
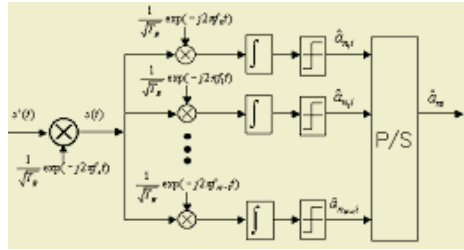
Fig. 5. The receiver of OFDM/QPSK system

4   Simulation and Analysis

In this paper, we analyzed performance of watermark information using OFDM/QPSK still image transmission system for wireless multimedia service.

The simulation parameters are used to performance evaluation that is presented in table 1.

Table 1. Parameter for simulation

Title	Value
Original image	'Lena512_std.bmp', 512*512 pixel
VI Watermark	'Audio.wav' Fs=8000Hz, Time=3sec
Image Watermark	'Mokwon64.bmp', 64*64 pixel
Frequency Transform	DCT
Transmission system	OFDM/QPSK
Noise Environment	AWGN

Original image uses 'Lena512_st.bmp' image, and watermark information used 'Audio.wav' information that have F_s(frequency sampling)=8000Hz, Time=3second. And, frequency transformation method used DCT scheme and OFDM/QPSK system for image transmission in AWGN environment.

Figure 6 shows performance of OFDM/QPSK system in AWGN channel environment.

Table 2 shows extracted VI watermark data, BER, and PSNR of restored image for proposed VI watermarking in reference SNR=7dB. The extracted VI watermark data can know that reconstruct by same level with original VI data although some noise occurs. And, table 3 shows extracted watermark data using logo image in condition

such as VI system. According to results, we can confirm than original image visually that performance decreases greatly. Thus, we could know that performance of proposed VI watermarking scheme is superior more than watermark scheme using logo image.
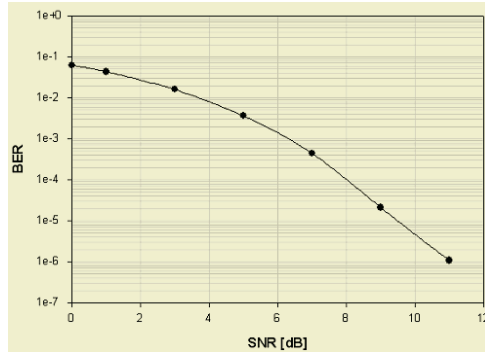
Fig. 6. BER of OFDM/QPSK system in AWGN

Table 2. Performance results of OFDM/QPSK system(SNR=7dB, proposed VI watermark)
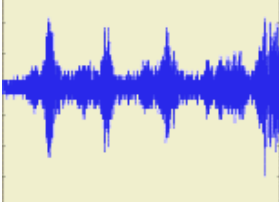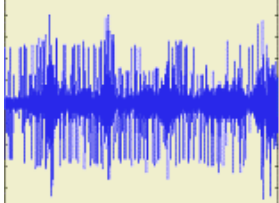
Results of Simulation			
Original Image		**Original Voice Information**	
Restoration Image		**Extracted Voice Information**	
PSNR	34.7982		
BER	4.4632e-4		
Transmission System	OFDM/QPSK		

Table 3. Performance results of OFDM/QPSK system(SNR=7dB, conventional logo image watermark)

Original Watermark Image	Extracted watermark image

5   Conclusion

In this paper, a watermark inserting method according to the sensitivity of intelligent human visual system (HVS) in DCT transform domain was proposed. Also, we proposed watermarking techniques using adaptive voice information (VI).

As a result of simulation, we could know that performance of proposed VI watermarking scheme using intelligent HVS scheme is superior more than watermark scheme using logo image. In the case of voice data, because restoration of voice information is available even if have the 10^{-3} BER. Therefore, we can improve all invisibility and robustness in Trade-Off relation in watermark technology. We found problem about data size of voice information when we insert voice information in image in VI watermark techniques. But, we confirmed that must apply voice information data below than F_s=8000Hz, Time=2sec about image more than 256*256 Pixel. And, when size of voice information is F_s=8000Hz and Time=8sec, we could know that it does not cause big effect about image more than 512*512 Pixel.

References

1. A. Z. Tirkel, R. G. Schyndel and C. F. Osborne, "A digital watermark," *In Proc. of IEEE ICIP'94*, vol. 2, pp. 86-90, Nov. 1994.
2. A. Z. Tirkel, C. F. Osborne and R. G. Schyndel, "Image watermarking – a spread spectrum application," *In Proc. of IEEE ICIP'96*, vol. 2, pp. 785-789, Sep. 1996.
3. W. Bender, D. Gruhl, N. Morimoto and A. Lu, "Techniques for data hiding," *IBM Systems Journal*, vol. 35(3/4), pp. 313-336, 1996.
4. I. Pitas, "A method for signature casting on digital images," *In Proc. of IEEE ICIP'96*, vol. 3, pp. 215-218, 1996.
5. http://www.watermarking.net/watermarking/watermarking.htm
6. R. Prasad, "*Universal Wireless Personal Communications*," Artech House Publishers, 1998.
7. L. J. Cimini Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Comm.*, vol. com-33, pp. 565-575, July 1985.

Image Retrieval
Based on a Multipurpose Watermarking Scheme

Zhe-Ming Lu[1,2,*], Henrik Skibbe[1], and Hans Burkhardt[1,**]

[1] Institute for Computer Science, University of Freiburg
Georges-Koehler-Allee 052, 79110 Freiburg i.Br., Germany
`{Zheminglu,Hans.Burkhardt}@informatik.uni-freiburg.de`
`henrik.skibbe@onlinehome.de`
[2] Department of Automatic Test and Control, Harbin Institute of Technology
P.O. Box 339, 150001 Harbin, China
`zhemingl@yahoo.com`

Abstract. The rapid development of Internet and multimedia technologies has made copyright protection and multimedia retrieval be the two most important issues in the digital world. To solve these problems simultaneously, this paper presents a multipurpose watermarking scheme for image retrieval. First, several important features are computed offline for each image in the database. Then, the copyright, annotation and feature watermarks are offline embedded into all images in the database. During the online retrieval, the query image features are compared with the exacted features from each image in the database to find the similar images. The experimental results based on a database with 1000 images in 10 classes demonstrate the effectiveness of the proposed scheme.

1 Introduction

With the development of computer, multimedia, and network technologies, two important issues have arisen nowadays. The first problem is that the amount of audio-visual information available in digital format has grown exponentially recently, which has resulted in information explosion and has exceeded the limit of human acceptability. Digitization, compression, and archival of multimedia information has become popular, inexpensive and straightforward. Subsequent retrieval of the stored information, however, might require considerable additional work in order to be effective and efficient. The second problem is that there is almost no limit for anyone to make lossless and unlimited copies of digital contents distributed over Internet and via CD-ROM, which is a major obstacle from the owner's viewpoint for entering the digital world. Copyright protection has therefore been one of the most important issues in the digital world. Here we concern with the problems in image retrieval and copyright protection. On the first issue, we should use some effective retrieval methods to get the desired images. In typical content-based image retrieval (CBIR) systems [1], the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. The feature vectors form a feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The similarities /distances between the feature vectors of the query example or sketch

---

* Room 01-046
** Room 01-030

and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing scheme. Recent retrieval systems have incorporated users' relevance feedback to modify the retrieval process. The existing image retrieval algorithms can be classified into two classes, i.e., low-level based and high-level (semantic) based. The low-level features include color [2], texture [3], shape [4] and spatial relationships [5]. Some retrieval methods based on semantic-level [6] have been proposed recently. On the latter issue, encryption may be one of solutions. However, conventional cryptographic systems permit only valid keyholders access to encrypted data, but once such data is decrypted there is no way to track its reproduction or retransmission. Over the last decade, digital watermarking has been presented to complement cryptographic processes. Invisible watermarks can be broadly classified into two types, robust and fragile (or semi-fragile) watermarks. Robust watermarks [7] are generally used for copyright protection and ownership verification because they are robust to nearly all kinds of image processing operations. In comparison, fragile or semi-fragile watermarks [8] are mainly applied to content authentication because they are fragile to most modifications. To fulfill copyright protection and content authentication together, multipurpose watermarking algorithms based on wavelet transform [9] and Vector Quantization [10] have been presented.

In general, the above two issues are taken into account separately. This paper presents a simple multipurpose watermarking scheme to solve these two problems simultaneously. The main idea is to offline embed three watermarks, i.e. the copyright, the denotation, and features into each image in the database. During the online retrieval, we can query based on the copyright, the denotation and the features.

2   Feature Extraction

Generally, we should extract the available features as many as possible. However, because the embedding of watermarks will affect the image quality, we should select the best representative features as less as possible. Here, we use three kinds of features, i.e., global invariant feature based on Haar Integral [11], statistical moments of the color histogram and Hu moments, which can be described in detail as follows.

2.1   Global Invariant Feature

Invariant features remain unchanged when the image is transformed according to a group of transformations. In [11], a kind of global feature invariant to rotations and translations are presented. Given a gray-scale image

$$\mathbf{M} = \{\mathbf{M}(\mathbf{x}), \mathbf{x} = (x_0, x_1), 0 \le x_0 < N_0, 0 \le x_1 < N_1\} \tag{1}$$

and an element $g \in G$ of the group of image translations and rotations, the transformation can be expressed as:

$$(g\mathbf{M})(\mathbf{x}) = \mathbf{M}(\mathbf{x}'), \text{ with } \mathbf{x}' = \left\{ \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \mathbf{x} + \begin{pmatrix} t_0 \\ t_1 \end{pmatrix} \right\} \mod \begin{pmatrix} N_0 \\ N_1 \end{pmatrix} \tag{2}$$

Based on above definition, an invariant transformation T must satisfy

$$T(g\mathbf{M}) = T(\mathbf{M}), \ \forall g \in G \tag{3}$$

For a given gray-scale image $\mathbf{M}$ and an arbitrary complex-valued function $f(\mathbf{M})$, it is possible to construct such an invariant transformation T by the following Haar integral:

$$T[f](\mathbf{M}) := \frac{1}{|G|}\int_G f(g\mathbf{M})dg = \frac{1}{2\pi N_0 N_1}\int_{t_0=0}^{N_0}\int_{t_1=0}^{N_1}\int_{\varphi=0}^{2\pi} f(g(t_0,t_1,\varphi)\mathbf{M})d\varphi dt_1 dt_0 \qquad (4)$$

For discrete images, because we choose integers for (t_0,t_1) and we use K steps for φ, we can obtain the following formula:

$$T[f](\mathbf{M}) \approx \frac{1}{KN_0 N_1}\sum_{t_0=0}^{N_0-1}\sum_{t_1=0}^{N_1-1}\sum_{k=0}^{K-1} f(g(t_0,t_1,\varphi=\frac{2\pi k}{K})\mathbf{M}) \qquad (5)$$

We can apply the Monte-Carlo method for the calculation of multi-dimensional integrals. In this paper, we only use the global feature (i.e. one float-typed value) calculated from luminance component for a color image.

2.2 Statistical Moments of the Color Histogram

Generally speaking, texture feature extraction methods can be classified into three major categories, namely, statistical, structural and spectral. In statistical approaches, texture statistics such as the moments of the gray-level histogram, or statistics based on gray-level co-occurrence matrix are computed to discriminate different textures. In this paper, statistical moments of the gray-level histogram are used to describe texture. Let z be a discreet random variable representing gray-levels in the range $[0, L-1]$, where L is the maximum gray value. Let $p(z_i), i = 0,1,...,L-1$ be a normalized histogram. Then the n-th moment with respect to the mean is given by:

$$\mu_n(z) = \sum_{i=0}^{L-1}(z_i - m)^n p(z_i) \qquad (6)$$

where m is the mean value of z. The second-order moment, variance, is a measure of gray-level contrast. And the third-order moment is a measure of skewness of the histogram and the fourth-order moment is a measure of its relative flatness. In this paper, we use the mean value m and three moments $\mu_2(z), \mu_3(z), \mu_4(z)$ of each color component's histogram in RGB color space, i.e., four float-typed values per color-component, as the features to be embedded.

2.3 Hu Moments

Hu moments are a set of algebraic invariants that combine regular moments [12]. They are invariant under change of size, translation, and rotation. Hu moments have been widely used in pattern recognition and proved successful in various applications. These moments can be used to describe the shape information of the image. For the digital image $\mathbf{M}$, in this paper, we only use first 4 moments for luminance component (i.e. four float-typed values) as follows:

$$\phi_1 = \eta_{20} + \eta_{02}, \phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2, \phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \qquad (7)$$

where $\eta_{pq} = \dfrac{\mu_{pq}}{\mu_{00}^{\gamma}}$, $\mu_{pq} = \displaystyle\sum_{x_0}\sum_{x_1}(x_0 - \bar{x}_0)^p(x_1 - \bar{x}_1)^q \mathbf{M}(x_0, x_1)$, $\bar{x}_0 = m_{10}/m_{00}$,

$\bar{x}_1 = m_{01}/m_{00}$ and $m_{pq} = \displaystyle\sum_{x_0}\sum_{x_1}x_0^p x_1^q \mathbf{M}(x_0, x_1)$.

3   Offline Multipurpose Watermarking

In our system, before we perform the online retrieval, we first embed three watermarks into each image in the database. Because these three watermarks possess different purposes, so we call our watermarking is a multipurpose watermark scheme. The first watermark is a copyright watermark, which is used for copyright protection. The second one is the annotation watermark, which is the name of the image or the semantic meaning of the image. The last one is the feature watermark, which is composed of the extracted features. These watermarks are all robust watermarks

In our watermarking system, the famous dither modulation method [13] is used to embed three watermarks into DCT block of each image's luminance component (we transform from RGB space to YUV space, only the Y component is used, and UV components remaining unchanged). The watermark's bits are distributed image orientated, so that cropping the image will crop the watermark the same way as the host-image. The dither modulation algorithm encodes each watermark bit in a middle-frequency DCT coefficient. Assume a DCT coefficient DCT_i in an even interval representing the bit '0' and in an odd interval representing the bit '1'. An interval is even, if, by a given Δ, $\lfloor DCT_i/\Delta \rfloor$=even, else odd. If a DCT coefficient is already in an interval which represents its bit, we move it into the middle of the interval, else we move it into the middle of the nearest interval next to its current interval. In our system, we embed twelve bits into each 8×8 DCT block.

The key problem is how to embed a float-typed value into the image. For simplicity, we embed the 4 bytes (i.e., 32 bits) which represents the float value directly into the image. However, a problem arising is the error in bits, even only one bit, may make the extracted value very different from the embedded one. To overcome this problem, we embed each float value multiple times into the image, and during the extraction, we determine the extracted bits by the majority cases. The second problem is how to guarantee the security of the watermarks, especially for the copyright watermark. Here we use the interleaving embedding technique. We randomly select the columns of DCT blocks to embed one of the watermarks, maybe the first column is used to embed the feature watermark, and the second column may be used to embed the annotation watermark, and so on. The selection scheme can be viewed as an embedding key.

4   Online Image Retrieval with Various Query Strategies

After embedding the watermarks into all the images in the database, we can perform the online retrieval now. Here we can use several kinds of queries. The most normal one is feature-based. And we can also use copyright-based and annotation-based, or we can use the combination query schemes.

For the feature-based query scheme, we first extract the features from the query image. Note that if the query image is without a feature watermark, we should compute the feature online for it. Then we compare the query feature vector with the feature watermark in each database image to find the most desired similar images. For the feature-and-annotation-based query scheme, we can only search the desired images with the same annotation (in general, it is the semantic of the image). For the feature-and-copyright-based query scheme, we can only search the desired images with the same copyright. For the feature-annotation-copyright-based query scheme, we can only search the desired images with the same copyright and with the same annotation.

5   Experimental Results

We perform the experiments based on a test database with 1000 images in 10 classes, each class including 100 images. They are with size of 384×256 or 256×384. During the offline process, we extract and embed three kinds of features. For the global invariant feature, we use n=10000 and $f(\mathbf{M}) = \mathbf{M}(0,1) \cdot \mathbf{M}(2,0)$. We embed a 48×48 sized binary copyright watermark, 17 float-typed feature values and 16-bytes annotation text. In the experiment, we view the annotation text as part of the feature watermark. We combine the feature watermark and the copyright watermark to construct the watermark to be embedded for each database image. Then, we embed the watermark to the corresponding image. After we obtain a watermarked image database, we can perform the online retrieval with various queries.

For the query based on features, we show an example of retrieval results in Fig. 1. The average precision and recall for each class is shown in Table 1. The average precision and recall for each class can be obtained as follows: First, we randomly select ten images from the class. Then, we use each image to be the query image. For each query image, we get the "recall" by obtaining the ratio of returned images in this class in the first 100 returned images, and find the position of the first returned image which is not in this class and divide it by 100 to obtain the "precision". After getting ten recalls and ten precisions, we average them to get the average recall and precision.

For image retrieval system, the most important operation to which our system should resist is the high-quality compression, while other attacks are not so important. Under the condition that we can extract the watermark more than 80% similar to the original embedded information. In previous sections, we have mentioned that the extracted feature watermark should be no bit loss. With regard to this, in the case of 100% recovery, the experimental results show that the average lowest JPEG compression quality factor to which the feature watermark can resist is 90.

6   Conclusions

This paper present an image retrieval system based on multipurpose watermarking scheme. The advantages of this scheme lie in three aspects. First, the system embeds the features in the images, and we need no extra space to save the feature data. Secondly, the image file can be copied to other database without recomputing corre-

sponding features. Thirdly, the image is also copyrighted and annotated, and the retrieval system can give the originality of the image and the semantic of the image.
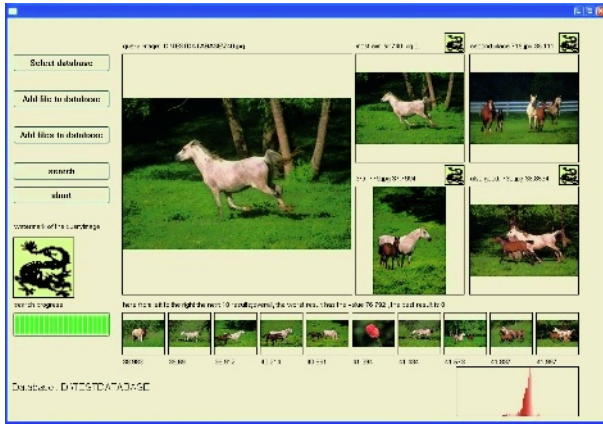
Fig. 1. The retrieval interface of our system

Table 1. The average recall and precision for each class

Class number	Semantic	Recall	Precision
1	People	0.27	0.02
2	Beach	0.14	0.01
3	Building	0.20	0.01
4	Bus	0.24	0.02
5	Dinosaur	0.67	0.29
6	Elephant	0.35	0.02
7	Flower	0.24	0.04
8	Horse	0.48	0.12
9	Mountain	0.18	0.02
10	Food	0.20	0.01

Acknowledgement

References

1. Long, F. H., Zhang, H. J., Feng, D. D.: Fundamentals of Content-based Image Retrieval, in Multimedia Information Retrieval and Management - Technological Fundamentals and Applications. Editor: Feng, D. D., Siu, W.C., Zhang, H. J., Springer-Verlag, Berlin Heidelberg New York (2003)
2. Smith, J. R.: Color for Image Retrieval. In: Image Databases-Search and Retrieval of Digital Imagery, Castelli V, Beagman L D, eds. John Wiley & Sons, Inc. Ch. 11, (2002)285-311

3. Sebe, N., Lew, M. S. Texture Features for Content Based Retrieval, In: Principles of Visual Information Retrieval, Lew M S, ed. Springer, Ch. 3, (2001) 51-85

4. Lee, K. M., Street, W. N.: Incremental Feature Weight Learning and Its Application to A Shape Based Query System. PRL, Vol. 23, No.7, (2002)265-274

5. Sciascio, E. D., Donini, F. M., Mongiello, M.: Spatial Layout Representation for Query by Sketch Content-based Image Retrieval, PRL, Vol. 23, No. 13, (2002)1599-1612

6. Lew, M. S.: Features Selection and Visual Learning, In: Principles of Visual Information Retrieval. Lew, M. S., ed. Springer, Ch. 12, (2001)297-318

7. Wang, Y., Doherty, J. F., Van Dyck, R. E.: A Wavelet-based Watermarking Algorithm for Ownership Verification of Digital Images. IEEE Trans. Image Processing, Vol. 11, No.2, (2002)77-88

8. Jaejin, L., Chee, S. W.: A Watermarking Sequence Using Parities of Error Control Coding for Image Authentication and Correction. IEEE Trans. Consumer Electronics, Vol.46, No. 2, (2000) 313-317

9. Lu, C. S., Liao, H.Y.M.: Multipurpose Watermarking for Image Authentication and Protection. IEEE Trans. Image Processing, Vol. 10, No. 10, (2001)1579-1592

10. Lu, Z. M., Xu, D. G. and Sun, S. H.: Multipurpose Image Watermarking Algorithm Based on Multistage Vector Quantization. IEEE Transactions on Image Processing. Vol. 14, No. 6, (To be published).

11. Siggelkow, S., Burkhardt, H. Fast Invariant Feature Extraction for Image Retrieval, in State-of-the-Art in Content-Based Image and Video Retrieval. Veltkamp, R. C., Burkhardt H., Kriegel, H. P. ed. Kluwer Acadenic Publishers (2001)

12. S.O. Belkasim, M. Shridhar, and M. Ahmadi. Pattern recognition with moment invariants: a comparative study. Pattern Recognition, 24(12):1117–1138, 1991.

13. Chen, B., Wornell, G. W.: Digital Watermarking and Information Embedding Using Dither Modulation. IEEE Second Workshop on Multimedia Signal Processing, (1998)273-278

Compressed Domain Video Watermarking in Motion Vector

Hao-Xian Wang, Yue-Nan Li, Zhe-Ming Lu, and Sheng-He Sun

Harbin Institute of Technology, Department of Automatic Test and Control
P.O. Box 339, 150001 Harbin, China
haoxianwang@yahoo.com.cn, liyuenan@dsp.hit.edu.cn
zhemingl@yahoo.com

Abstract. In this paper, a novel watermarking scheme in the compressed domain is proposed for video copyright protection. The robust watermark is embedded in motion vectors of the bitstream based on the relationships between one-pixel accuracy and half-pixel accuracy motion vectors at the encoder. Watermark can be blindly extracted by parsing the properties of motion vectors at the decoder. Simulation results show that the proposed algorithm can perform the real-time watermark embedding process without affecting visual quality of the video sequence and it is also standard-compliant. The proposed algorithm can be applied to any video compression standard with half-pixel accuracy motion vectors such as MPEG-2 and H.263.

1 Introduction

With the rapid development of Internet, the multimedia distributions have been extremely easy and common. Those multimedia data can be duplicated without significant degradations [1]. As a result, multimedia data are suffering from being illegally duplicated. Thus, multimedia copyright protection has received more attentions. Digital watermarking which was proposed in the end of the 20th century provides a prospective way to protect copyrights. In recent years, many watermarking algorithms have been proposed and those algorithms embed watermarks in various kinds of multimedia data, including still images, audio, video, 3D models and so on.

   Digital video watermarking serves a number of purposes including copy control, broadcast monitoring, video authentication, copyright protection, enhanced video encoding and so on [2]. Since video sequences should be stored and transmitted in compressed formats, so it is reasonable to embed watermarks in compressed video bitstreams. Compressed domain watermarking algorithms should meet several demands. Firstly, watermarks should be detected blindly. Secondly, algorithms should be standard-compliant and real-time implemented. Thirdly, visual quality and bit rate should also be maintained. Compared with the research in the uncompressed video watermarking branch, those in the compressed video watermarking branch seem to be quite rare. In this paper, a novel video watermarking algorithm in the compressed domain based on motion vectors is proposed. The rest of this paper is organized as follows: The proposed watermark embedding and extracting schemes are described in Section 2. In Section 3, we give the simulation results of the proposed algorithm. Finally, the conclusions are given in Section 4.

2 Watermarking in Motion Vector

Motion estimation and compensation are widely used in video compression standards. Temporal redundancies are significantly reduced by using motion estimation in encoding inter frames. Motion vectors take a large amount of the compressed video bitstream. Thus, watermark can be embedded in the bitstream by modifying the components of motion vectors. Several watermarking algorithms based on motion vectors have been proposed. In a proposal for MPEG, Jordan et al. proposed to regularize the motion vectors into a modified data sequence for easily watermark retrieving [3]. Zhang et al. proposed to embed watermark in the less significant components of motion vectors [4]. In the algorithm proposed in [5], parities of motion vectors are modified to embed watermark.

In hybrid video compression schemes, motion estimation is used to predict the blocks to be coded. Compression is archived by quantizing the DCT coefficients of residual blocks after motion compensated prediction (MCP) and entropy coding are performed to further reduce the bit rate. Half-pixel accuracy motion estimation is adopted in compression standards like MPEG-2 [6] and H.263 [7] to get higher visual qualities.

2.1 Watermark Embedding in Motion Vector

The motion vector to be processed is denoted by (MVX, MVY) in this paper. MVX and MVY are the horizontal and vertical components respectively. MVX_{int} and MVX_{half} denote one pixel and half pixel values of MVX ($MVX = MVX_{int} + MVX_{half}$), and it is the same case for MVY. MVX_{half} and MVY_{half} take values from $\{-0.5, 0, 0.5\}$. θ_{int} and θ_{half} are the phase angles of (MVX_{int}, MVY_{int}) and (MVX_{half}, MVY_{half}) in the coordinates respectively, where $\theta_{int}, \theta_{half} \in [0, 2\pi)$. $Angle_{int}$ and $Angle_{half}$ are the quantized values of θ_{int} and θ_{half} respectively. The quantization step is $\dfrac{\pi}{4}$ and the quantization rule of θ_{int} is described in (1). It can be easily seen that $Angle_{int} = \theta_{int}$ if θ_{int} is a multiple of $\dfrac{\pi}{4}$.

$$Angle_{int} = \begin{cases} \theta_{int}, & \text{if } \theta_{int} = \dfrac{k\pi}{2} \ (k = 0,1,2,3) \\[2mm] \dfrac{\pi}{4}, & \text{if } 0 < \theta_{int} < \dfrac{\pi}{2} \\[2mm] \dfrac{3\pi}{4}, & \text{if } \dfrac{\pi}{2} < \theta_{int} < \pi \\[2mm] \dfrac{5\pi}{4}, & \text{if } \pi < \theta_{int} < \dfrac{3\pi}{2} \\[2mm] \dfrac{7\pi}{4}, & \text{if } \dfrac{3\pi}{2} < \theta_{int} < 2\pi \end{cases} \tag{1}$$

The original watermark is a binary sequence W ($W = \{w_i\}$), where $w_i = 0, 1$. In order to improve the robustness of the proposed algorithm, W is permuted via a key to generate the sequence W^* ($W^* = \{w^*_i\}$):

$$w_i^* = Permute(w_i, key).$$ (2)

As the proposed algorithm can embed 3 bits in each motion vector, we transform W^* into an integer sequence G by combining the neighboring three bits, where $G=\{g_k \mid 0 \leq g_k \leq 7\}$. For example, if the three binary bits are "100", then g=4.

In the proposed algorithm, only the motion vectors that with large horizontal and vertical components are selected to embed watermarks as described in (3). The reason is that human visual system is less sensitive to fast movements.

$$|MVX_{int}| \geq T \ \& \ |MVY_{int}| \geq T$$ (3)

The proposed algorithm embeds watermark by modifying the relationship between $Angle_{int}$ and $Angle_{half}$. Changing MVX_{half} and MVY_{half} can adjust $Angle_{half}$ easily and the impact on the visual quality is negligible. $Angle^*_{half}$ is the modified value of $Angle_{half}$ and it is modified according to the following formula.

$$Angle^*_{half} = (Angle_{int} + g_k \times \frac{\pi}{4}) \bmod 2\pi$$ (4)

Then, MVX_{half} and MVY_{half} are modified according to $Angle^*_{half}$. An example is shown in Fig.1. It should be noted that the motion vector data are written into the bitstream in the form of $(MVX_{int} + MVX^*_{half}, MVY_{int} + MVY^*_{half})$. As a result, we can only get the absolute values of MVX^*_{half} and MVY^*_{half} without knowing their signs at the decoder if they are nonzero. To solve this problem, we use the parities of the integer parts of MVX and MVY to denote the signs of MVX^*_{half} and MVY^*_{half}. For example, an even $[MVX]$ indicates MVX^*_{half} is positive and an odd one means a negative MVX^*_{half}. The two functions $Parity(x)$ and $Sign(x)$ are used to indicate the parity and sign of a number respectively, the definitions are shown in (5) and (6).

$$Parity(x) = \begin{cases} 1, & \text{if } x \bmod 2 = 0 \\ -1, & \text{if } x \bmod 2 = 1 \end{cases}$$ (5)

$$Sign(x) = \begin{cases} 1, & \text{if } x > 0 \\ -1, & \text{if } x < 0 \end{cases}$$ (6)

MVX_{int} and MVX^*_{half} are then modified as follows if $MVX^*_{half} \neq 0$. MVY_{int} and MVY^*_{half} are modified in the same manner if $MVY^*_{half} \neq 0$.

if $Parity(MVX_{int}) \neq Sign(MVX^*_{half})$ &$MVX_{int}>0$, then $MVX^*_{int} = MVX_{int}-1$, $MVX^*_{half}=0.5$;

if $Parity(MVX_{int})=Sign(MVX^*_{half})$ &$MVX_{int}>0$, then $MVX^*_{half}=0.5$;

if $Parity(MVX_{int}) \neq Sign(MVX^*_{half})$&$MVX_{int}<0$, then $MVX^*_{int} = MVX_{int}+1$, $MVX^*_{half}=-0.5$;

if $Parity(MVX_{int})=Sign(MVX^*_{half})$ &$MVX_{int}<0$, then $MVX^*_{half}=-0.5$;

It can be seen from the above discussions that $|MVX^*-MVX_{int}| \leq 0.5$ and $|MVY^*-MVY_{int}| \leq 0.5$, so the modifications on motion vectors are negligible. The watermarked motion vectors satisfy the following inequations:

$$|MVX^*| \geq T-0.5 \ \& \ |MVY^*| \geq T-0.5.$$ (7)

If an unwatermarked motion vector satisfies $|MVX_{int}|=T-1$ or $|MVY_{int}|=T-1$, its MVX_{half} or MVY_{half} will be set to be 0. In this way, we can easily determine whether a motion vector has been watermarked by comparing the absolute values of its horizontal and vertical components with $(T-0.5)$.
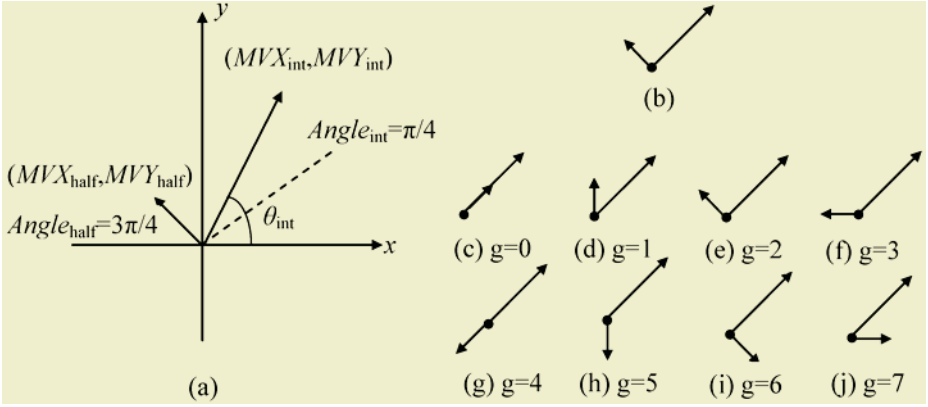
Fig. 1. An example of the proposed embedding scheme. (a) Original motion vectors in the coordinates. (b) Original directions of $Angle_{int}$ and $Angle_{half}$, the longer vector denotes the direction of $Angle_{int}$ and the shorter one denotes that of $Angle_{half}$. (c)-(j) Directions of $Angle^*_{half}$ when the watermark bits vary from 0 to 7

2.2 Watermark Extraction

The decoded motion vector is denoted by (MVX^*, MVY^*), $MVX^* = MVX^*_{int} + MVX^*_{half}$ and $MVY^* = MVY^*_{int} + MVY^*_{half}$. Then the process of watermark extraction in each motion vector can be described as follows.

Step1: Determine whether a motion vector has been watermarked according to (7). If it is watermarked, go to Step2.

Step2: Calculate $(MVX^*_{int}, MVY^*_{int})$ and $(MVX^*_{half}, MVY^*_{half})$. MVX^*_{int} and MVX^*_{half} can be calculated according to (8) and (9), where operator $[\cdot]$ in (8) denotes truncation towards zero. MVY^*_{int} and MVY^*_{half} can be calculated in the same manner.

$$MVX^*_{int} = [MVX^*] \tag{8}$$

$$MVX^*_{half} = \begin{cases} 0, & \text{if } MVX^* = MVX^*_{int} \\ 0.5, & \text{if } |MVX^* - MVX^*_{int}| = 0.5 \ \& \ MVX^*_{int} \bmod 2 = 0 \\ -0.5, & \text{if } |MVX^* - MVX^*_{int}| = 0.5 \ \& \ MVX^*_{int} \bmod 2 = 1 \end{cases} \tag{9}$$

Step3: Calculate $Angle^*_{int}$ and $Angle^*_{half}$ according to (1).

Step4: Extract watermark bits as follows by comparing $Angle^*_{int}$ and $Angle^*_{half}$.

$$g^*_k = \begin{cases} (Angle^*_{half} - Angle^*_{int}) / \dfrac{\pi}{4}, & \text{if } Angle^*_{half} \geq Angle^*_{int} \\ (Angle^*_{half} - Angle^*_{int} + 2\pi) / \dfrac{\pi}{4}, & \text{if } Angle^*_{half} < Angle^*_{int} \end{cases} \tag{10}$$

Step5: Go to the next motion vector and repeat Steps 1-4. If all watermark bits have been extracted, then go to Step 6.

Step6: Transform the extracted integer sequence G^* ($G^* = \{g^*_k\}$) into a binary sequence W^* ($W^* = \{w^*_i\}$). Then W^* is restored via the same key as used in watermark embedding process to get the final extracted watermark sequence.

3   Simulation Results

We use $PSNR_{\mathrm{aver}}$ and $PSNR^{*}_{\mathrm{aver}}$ to denote the average $PSNR$s of the unwatermarked and watermarked video sequences respectively, and they are calculated according to (11). Only the luminance component $PSNR$ of each frame is concerned.

$$PSNR_{\mathrm{aver}} = \frac{\sum_{i=0}^{N-1} PSNR_i}{N}, PSNR^{*}_{\mathrm{aver}} = \frac{\sum_{i=0}^{N-1} PSNR_i^{*}}{N} \tag{11}$$

$\Delta PSNR$ describes the difference between $PSNR_{\mathrm{aver}}$ and $PSNR^{*}_{\mathrm{aver}}$:

$$\Delta PSNR = PSNR_{\mathrm{aver}} - PSNR^{*}_{\mathrm{aver}}. \tag{12}$$

NC is used to denote the similarity measurement between the original watermark and the extracted watermark. NC is calculated as follows:

$$NC = \frac{N_{\mathrm{crt}} - N_{\mathrm{wrg}}}{N_{\mathrm{crt}} + N_{\mathrm{wrg}}}, \tag{13}$$

where N_{crt} is the number of the correctly extracted watermark bits and N_{wrg} is the number of wrongly extracted bits. It can be seen from (13) that NC=1 if all watermark bits are correctly extracted.

The watermark embedded in each video sequence is a 64×64 binary image. And it is embedded into H.263 bit-streams. As differential motion vectors are adopted in H.263, so watermark is embedded in differential motion vectors [7].

The *foreman* sequence (120 frames, 352×288, 4:2:0) and *tennis* sequence (200 frames, 352×240, 4:2:0) are used in simulations. In both *foreman* and *tennis* sequences, each group-of-pictures (GOP) contains 20 frames. The *bird* image is embedded in *foreman* sequence and *frog* image is embedded in *tennis* sequence. Quantization parameters (QP) are set to be 4, 8 and 16 respectively. The threshold T in (3) is set to be 3 for both sequences. Simulation results are listed in Table 1 and Table 2. The watermarked frames and extracted watermarks are shown in Fig.2 and Fig.3.

Table 1. Simulation results of *foreman* sequence when QP=4, 8 and 16 respectively

QP	$PSNR_{\mathrm{aver}}$ (dB)	$PSNR^{*}_{\mathrm{aver}}$ (dB)	$\Delta PSNR$ (dB)	NC
4	38.792	38.778	0.014	1.0
8	35.351	35.298	0.053	1.0
16	32.774	32.592	0.182	1.0

Table 2. Simulation results of *tennis* sequence when QP=4, 8 and 16 respectively

QP	$PSNR_{\mathrm{aver}}$ (dB)	$PSNR^{*}_{\mathrm{aver}}$ (dB)	$\Delta PSNR$ (dB)	NC
4	37.069	37.029	0.040	1.0
8	32.765	32.711	0.054	1.0
16	29.911	29.785	0.126	1.0

Fig. 2. Simulation results of *foreman* sequence. (a) Original 30th frame; (b) Watermarked 30th frame when *QP*=4; (c) Original watermark; (d) Extracted watermark
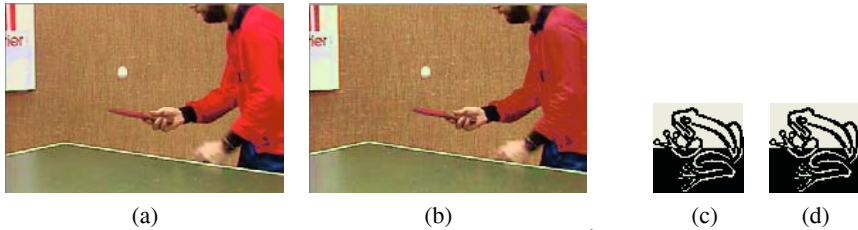
Fig. 3. Simulation results of *tennis* sequence. (a) Original 45th frame; (b) Watermarked 45th frame when *QP*=4; (c) Original watermark; (d) Extracted watermark

4   Conclusions

In this paper, we propose a new video copyright protection scheme by embedding watermark in motion vectors. Motion vectors are modified to meet specific relationships depending on watermark bits. The proposed compressed domain watermarking algorithm can effectively embed and extract watermarks without complex computations. The distortions caused by watermark embedding are imperceptible. The proposed algorithm can also be applied in other video compression standards that adopt half-pixel accuracy motion estimation for copyright protection and other applications.

Acknowledgement

References

1. Frank Hartung, Bernd Girod: Watermarking of Uncompressed and Compressed Video. Singal Processing.Vol.66. 283-301.1998
2. Gwenaël Doërr, Jean-Luc Dugelay: A Guide Tour of Video Watermarking. Signal Processing: Image Communication. Vol.18. 263-282. 2003.

3. F. Jordan, M. Kutter, T. Ebrahimi: Proposal of Watermarking Technique for Hiding/Retrieving Data in Compressed and Decompressed Video. Technical report M2281, ISO/IEC document, JTC1/SC29/WG11, 1997.
4. J. ZHANG, J. G. LI, L. ZHANG: Video Watermark Technique in Motion Vector. Proceedings of XIV Brazilian Symposium on Computer Graphics and Image Processing. Florianópolis, Brazil (2001) 179-182.
5. Z. J. ZHU, G. Y. JIANG, X. W. YU: New Algorithm for Video Watermarking. 6th International Conference on Signal Processing, Vol.1. Beijing, China (2002) 760-763.
6. MPEG: Information Technology--Generic Coding of Moving Pictures and Associated Audio Information: Part 2: Video. ISO/IEC: 13818-2, 1995.
7. Standardization Sector of ITU: Video Coding for Low Bit Rate Communication. ITU-T Recommendation H.263, 1998.

Analysis of Quantization-Based Audio Watermarking in DA/AD Conversions

Shijun Xiang, Jiwu Huang, and Xiaoyun Feng

School of Information Science and Technology, Sun Yat-Sen University
Guangzhou 510275, P.R. China
`isshjw@zsu.edu.cn`

Abstract. Digital audio watermarking may suffer from the different attacks simultaneously in DA/AD conversions, which may be modeled to noise corruption, modification of signal energy, and phase change. In this paper, we conduct many experiments to show that the phase change of the audio signals in DA/AD conversions may be represented as the temporal scaling, which results in the desynchronization of the watermark. Furthermore, we analyze the performance of quantization-based audio watermarking against Gaussian noises and the modification of signal amplitude. We present the corresponding calculation expression of *BER*. Finally, it is concluded that quantization-based audio watermarking is very susceptible to DA/AD conversions.

1 Introduction

Most of the audio watermarking algorithms can be grouped into two categories, additive-based and quantization-based watermarking schemes. Additive-based scheme [1] embeds the watermark by adding a PN (pseudo noise) sequence in time domain [2] or frequency domain [3]. While in quantization-based scheme [4~9], the original signals are quantized by different thresholds to embed watermarks. In the extracting, the hidden information can be recovered according to distances between to-be tested data and different quantization results. Due to its merits, such as blind extraction, quantization techniques become predominant methods in watermarking.

In many applications, audio watermarking will encounter DA/AD conversions. The previous works on quantization-based watermarking mainly considered a certain attack, such as amplitude modification [6], cropping [7], noise corruption and synchronization attacks [8], and temporal scaling [9]. In DA/AD conversions [10, 11], however, audio watermarking may suffer from various attacks simultaneously, such as the modification of the signal and phase, which present more challenges.

This paper discusses the possible attacks in DA/AD conversions and models these attacks as noise corruption, amplitude distortion, and phrase change. Then we analyze the performance of quantization audio watermarking against these attacks. Finally, we conclude that quantization audio watermarking is susceptible to DA/AD conversions.

2 Quantization Audio Watermarking Model

Let $X = \{x_1, x_2, \ldots, x_N\}^T$ denote the cover-signal, then perform linear transform on X:

$$Y = \{y_1, y_2, \cdots, y_N\}^T = UX \tag{1}$$

where U is $N \times N$ dimensions unitary matrix, Y the transformed coefficients.

Choose M coefficients $\{c_1, c_2, \ldots, c_M\}^T$ from Y by the key, then embed 1 bit information into each coefficient. Let $W = \{w_i\}$, a $\{-1, 1\}$ sequence, denote the to-be-hidden data. We use the embedding equation in reference [4] and assign $T_1 = 3S/4, T_2 = S/4$ to discuss:

$$
\begin{cases}
c_i' = c_i - Mod\,(c_i, S) + T_1\,, & if \quad c_i \geq 0 \quad and \quad w_i = 1 \\
c_i' = c_i - Mod\,(c_i, S) + T_2\,, & if \quad c_i \geq 0 \quad and \quad w_i = -1 \\
c_i' = c_i + Mod\,(c_i, S) - T_1\,, & if \quad c_i < 0 \quad and \quad w_i = 1 \\
c_i' = c_i + Mod\,(c_i, S) - T_2\,, & if \quad c_i < 0 \quad and \quad w_i = -1
\end{cases}
\tag{2}
$$

where S is quantization step size, c_i and c_i' are the transformed coefficients of original audio and watermarked audio respectively, $Mod(c_i, S)$ denotes the remainder of c_i divided by S. The different values of the thresholds of T_1 and T_2 indicate the differences of embedding "1" or "-1" into the audio.

When extracting the hidden data, we can make the same unitary transformation and figure out c_i' by the key, then extract watermark bit w_i' by the following corresponding detection formula (5) in reference [4]. $(T_1+T_2)/2 = S/2$ here.

$$
w_i' = \begin{cases}
1 & if \quad Mod(c_i', S) \geq (T_1 + T_2)/2 \\
-1 & if \quad Mod(c_i', S) < (T_1 + T_2)/2
\end{cases}
\tag{3}
$$

3   The Effects of DA/AD Conversions on Audio Watermarking

To better investigate the effects of DA/AD conversions, we design the following test model to simulate the environment in Fig.1. The analog line is considered clear here.
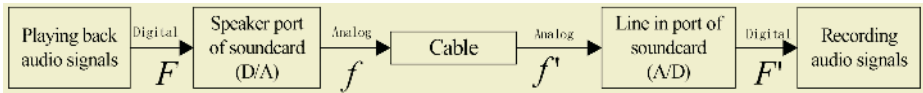
Fig. 1. Test Model for DA/AD Conversions

Table 1. Four different 16-bit signed mono audio files in WAVE format are used for testing

File name	Sample rate(kHz)	Length(s)	Properties
march.wav	8 and 44.1	56	Both low-frequency and high-frequency
drum.wav	8 and 44.1	56	Mainly composed of low- frequency
flute.wav	8 and 44.1	56	Mainly composed of low- frequency
dialog.wav	8 and 44.1	56	Daily dialog

By many experiments, we observed audio signal will suffer from the modification of the amplitude and the phase during DA/AD conversions.

A. Modification of Signal Energy
After DA/AD conversions, wave distortion may occur in the watermarked audios, mainly represented as the modification of signal energy and the noises corruption.

Besides noise corruption, many experiments show that the audio amplitude is modified after DA/AD conversions, and the modifications rely on the volume played back, the performance of soundcards and analog channel. Fig.2 shows that energy of the recorded audio by Sound Blaster Live5.1 soundcard is obviously reduced.
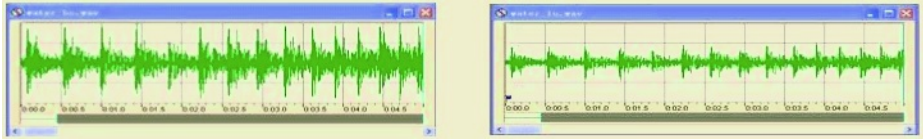
Fig. 2. The original audio (*left*) and the recorded audio (*right*)

B. Phase Change

Based on test model in Fig.1, five different soundcards are employed to test files in Table 1. Where, Sound Blaster Live5.1 is a civil sound blaster, TERRATEC AUREON Xfire 1723 and ICON StudioPro7.1 are the professional one, Audio PCI and Nightingale Pro 6 are common sound blasters. Table 2 is the test results.

From Table 2, we have the following observations during DA/AD conversions: a) The temporal linear scaling occurs, and the scaling factor is varied with different soundcards; b) The number of modified samples is related to the sampling rate of audio signals even the soundcard is unchanged.

From the view point of signal processing, watermark is weak signal embedded into a strong background, thus the variety of carriers will influence the watermark directly. So the effects of DA/AD conversions on audio watermarking are generalized as:

  i. Noise corruptions, such as the distortion of the watermark by Gaussian noise.
 ii. The modification of signal amplitude, also degrading the watermark.
iii. Temporal scaling, leading to the desynchronization of the watermark.

Table 2. Modification of audio sample size (sampled at 8 kHz and 44.1 kHz)

Soundcard / Time		Blaster Live5.1	AUREON Xfire 1723	Audio 2000 PCI	Nightingale Pro 6	Studio Pro 7.1
8 kHz	10s	Reduce: 1	Increase: 1000	Increase: 102	Increase: 1079	Reduce: 70
	30s	Reduce: 3	Increase: 3000	Increase: 306	Increase: 3237	Reduce: 210
	50s	Reduce: 5	Increase: 5000	Increase: 510	Increase: 5395	Reduce: 350
44.1 kHz	10s	Reduce: 6	Increase: 0	Increase: 0	Increase: 14	Increase: 0
	30s	Reduce: 18	Increase: 0	Increase: 0	Increase: 42	Increase: 0
	50s	Reduce: 30	Increase: 0	Increase: 0	Increase: 70	Increase: 0

4   Performance Analysis of Quantization Audio Watermarking

A. Gaussian Noise Corruption

AWGN (additive white Gaussian noise) is possibly introduced during the transmission. Let $\{n_i\}$ denote AWGN with mean 0 and variance δ. The value of $\{n_i\}$ in the transform domain is $d_i = c_i'' - c_i'$, where $\{c_i''\}$ and $\{c_i'\}$ are the transformed coefficients of watermarked audio and corrupted watermarked audio respectively.

Clearly, if $d_i \notin (kS - S/4, kS + S/4)$, the watermark can not be detected correctly by equation (3). For some unitary transformations, like DWT and DCT, $\{d_i\}$ still obey the same distribution as $\{n_i\}$. So BER against Gaussian noise corruption can be expressed as

$$BER = 1 - \sum_{k \in Z} \int_{kS-S/4}^{kS+S/4} \frac{1}{\sqrt{2\pi}\delta} \cdot e^{\frac{-x^2}{2\delta^2}} \cdot dx \approx \frac{2}{\sqrt{2\pi}\delta} \cdot \int_{S/4}^{+\infty} e^{-\frac{x^2}{2\delta^2}} \cdot dx \qquad (4)$$

Equation (4) indicates that the robustness of watermark is mostly determined by S under the same power of white noise corruption but independent on the origin audio.

B. Amplitude Change Attack

DA/AD conversions may modify the audio amplitude. Assume that the amplitude of audio $\{|f_i|\}$ changes to $\{|\beta_i \cdot f_i|\}$ due to the amplitude attacks. The amplitude of coefficient $Y = \{|y_i|\}$ will change to $\{|\lambda_i \cdot y_i|\}$, $i = 1, 2, \cdots, N$, where β_i and λ_i denote the amplitude change factors before and after transformation, respectively.

Let $C' = \{c_i'| j = 1, 2, \cdots, M\}$ denote the to-be detected coefficients selected from Y by the key. The modifications of the coefficients are $d_i = (\lambda_i - 1) \cdot c_i'$. In case of $d_i > kS + S/4$ $(k \in Z)$, errors will occur in the extracting, as shown in equation (5).

$$(\lambda_i - 1) \cdot c_i' > kS + S/4 \qquad (5)$$

For amplitude scaling attack, we have $\beta = \beta_i = \lambda_i$. Rewrite equation (5) as:

$$(\beta - 1) \cdot c_i' > kS + \frac{S}{4} \quad \Leftrightarrow \quad \begin{cases} c_i' > \dfrac{S}{(\beta - 1)}(k + 0.25), & if \quad \beta > 1 \\ c_i' < \dfrac{S}{(\beta - 1)}(k + 0.25), & if \quad 0 < \beta < 1 \end{cases} \qquad (6)$$

where $k = \lfloor (\beta - 1) \cdot c_i' / S + 0.25 \rfloor$, $\lfloor \; \rfloor$ the floor function. Note that the value of β goes to 1, k and $(\beta - 1) \cdot c_i'$ go to 0. So the smaller numbers of coefficients satisfy equation (6), indicating the lower BER. Therefore, the BER of audio watermarking against amplitude change attack can be expressed as:

$$BER = E/M \qquad (7)$$

where E denotes the number of coefficients satisfying equation (5) or (6).

By equation (5) and (6) we know the main influences on BER against amplitude change attack in quantization audio watermarking are as follows,

  i. The quantization step size, namely the embedding strength. The larger of the size, the stronger ability of the watermark to resist the attack.
 ii. The modification of audio amplitude. Consider its own property, $\beta \in (0, 1.5)$.
iii. The genres of the audio. For the same amplitude attack, $(\lambda_i - 1) \cdot c_i'$ keep unchanged, thus BER does not rely on the audio genre. As for amplitude scaling attack, $(\lambda_i - 1) = (\beta - 1)$ a constant, the stronger of the signal energy c_i', the larger $(\beta - 1) \cdot c_i'$, then the weaker ability to resist amplitude scaling attack.

C. Temporal Scaling Attack

The phase of the signal is also changed during DA/AD conversions, shown as temporal scaling. Let $F = \{f(i)\}$ and $F' = \{f'(j)\}$ denote the watermarked audio and the audio performed DA/AD processes, the length L and L' respectively. Temporal scaling occurs when $L' \neq L$, and the temporal scaling factor is $\alpha = L'/L$. We choose *SNR* to measure the effect on watermarked audios here.

$$SNR = -10 \log_{10}\left(\frac{\|F - F''\|^2}{\|F\|^2}\right), \; F'' = F'/(\sum_{i=0}^{N-1}|f'_i| / \sum_{i=0}^{N-1}|f_i|) \tag{8}$$

where F'' is the normalized audio, $N = \min(L, L')$. F' should be normalized by equation (8) at first due to the modification of audio amplitude in DA/AD.

By Sound Blaster Live5.1 soundcard, we perform DA/AD conversions with audio files sampled at 8 kHz, and calculate *SNR*s of F verse F''. Fig. 3 shows the *SNR*s decrease quickly because the temporal scaling changed the locations of some samples, and resulted in serious distortion.
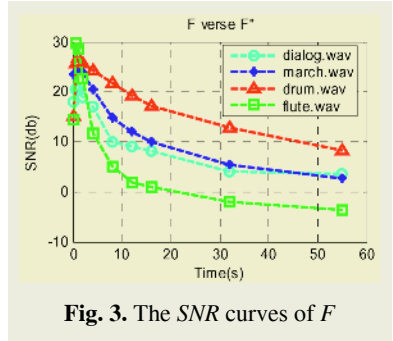
Fig. 3. The *SNR* curves of F

5   Experimental Results

To confirm the validity of results shown in Section 4.1 and 4.2, we choose *march.wav* and *dialog.wav* in Table 1 sampled at 8 kHz, with embedding strength of 9000 and 5000 respectively. We embedded the watermark into low frequency sub-band coefficients of DWT with the wavelet base db1. Fig. 4 and 5 are the test results.
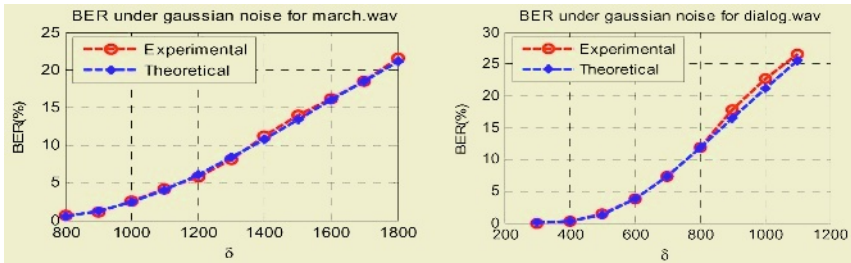
Fig. 4. *BER* against Gaussian noise
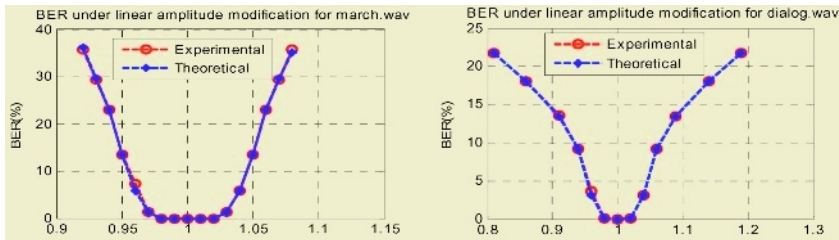
Fig. 5. *BER* against amplitude linear scaling (β)

Fig. 4 and Fig. 5 show that the experimental and theoretical results of *BER* against Gaussian noise corruption and amplitude scaling attack agree with one another very well, which indicated the correctness of equation (4) and equation (6).

6   Conclusions

In this paper, we investigated the possible attacks in DA/AD conversions to quantization-based audio watermarking, and modeled these attacks as noises, phase change, amplitude change. We analyzed the performance of quantization audio watermarking against Gaussian noises and modification of audio amplitude, and presented their calculation expression of *BER*. Many experiments show that audio phase change during DA/AD processes represents as the temporal linear scaling, which is determined by the performance of soundcards and the sampling rate of audio signals. As a conclusion, quantization-based audio watermarking is very susceptible to DA/AD conversions.

Acknowledgement

References

1. Cox I. J., Miller M. L.: A. Digital watermarking, London: Academic Press, (1999)
2. Bassia P, Pitas I, Nikolaidis N.: Robust audio watermarking in the time domain. IEEE Trans. on Multimedia, Vol. 3 (2001): 232-241
3. Cox, J. Kilian, F. Leighton and T. Shamoon: Secure Spread Spectrum Watermarking for Multimedia, IEEE Trans. on Image Processing, Vol. 6 (1997): 1673-1687
4. Tsai M J, Yu K Y, and Chen Y Z.: Joint wavelet and spatial transformation for digital watermarking. IEEE Trans. on Consumer Electronics, Vol.46, (2000):241-245
5. Chen B, Wornell G W.: Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. IEEE Trans. on Information Theory, Vol. 47 (2001): 1423-1443
6. Kim Siho, Keunsung Bae: Robust estimation of amplitude modification for scalar costa scheme based audio watermark detection. In: Proc. of IWDW (2004): 121-133
7. Huang Jiwu, Wang Yong, Shi Y. Q. A blind audio watermarking algorithm with self-synchronization. In: Proc. of ISCAS, Vol. 3 (2002): 627-630
8. Wu Shaoquan, Huang Jiwu, Huang Daren, Yun Q. Shi.: Efficiently Self-Synchronized Audio Watermarking for Assured Audio Data Transmission. IEEE Transactions. on Broadcasting, Vol 51 (2005): 69-76.
9. Mansour M and Tewfik A.: Audio watermarking by time-scale modification. In: *Proc.* of ICASSP, (2001): 1353-1356
10. Steinebach M., Lang A., Dittmann J., Neubauer C.: Audio watermarking quality evaluation: robustness to DA/AD processes. In Proc. of ITCC, (2002): 100-103.
11. Richard Popa. An Analysis of Steganographic Techniques. Ph. D Thesis, (1998): 26-27

A Lossless Watermarking Technique for Halftone Images

Ping-Sung Liao[1], Jeng-Shyang Pan[2], Yen-Hung Chen[2], and Bin-Yih Liao[2]

[1] Department of Electronic Engineering, Cheng-Shiu University
No. 840, ChengCing Rd., NiaoSong Township, Kaohsiung County 833, Taiwan
bsliao@csu.edu.tw
[2] Department of Electronic Engineering, National Kaohsiung University of Applied Sciences
415 Chien-Kung Road, Kaohsiung 807, Taiwan
{Pan,Chen,Liao,jspan@cc.kuas.edu.tw

Abstract. In this paper, a lossless watermarking technique in dealing with halftone images is proposed to achieve both copyright protection and exact recovery of the original halftone image with legal authorization. Experimental results show that pattern substitution along with an appropriate block cutting on a halftone image makes possible the positive recovery of the original halftone image after the extraction of a watermark.

1 Introduction

In recent years, computers are broadly and deeply linked to many different information processing applications that can be easily accessed from any place around the world. As a result, to protect the copyright of digital contents has become an urgent discipline. In the latter half of the 1990s, digital watermarking has been gained popularity as a research topic to embed a security message into digital documents that may be one of images, music, audio clip and video clip. Based on the characteristics of digital contents, many kinds of digital watermarking techniques [1, 2] have been proposed by researchers to protect the copyright of the diversified contents. Fundamentally, the watermarking techniques can be categorized into two approaches: one is rooted in spatial domain, another in transformed domain [2]. Watermarking based on spatial domain is, in practice, easily and fast implemented, but it is fragile without the consideration of incorporating other strategies to resist malicious modification upon the pixel's intensity. On the contrary, watermarking approach based on transform domains, for instance, like discrete Fourier transform, discrete cosine transform, wavelet transform and fractal transform, is more robust against attacks.

So far, most existing digital watermarking algorithms embed the security message into the original images in an irreversible manner so that not only there is a small distortion between the watermarked images and the original image, but also it is infeasible to recover the original image form the watermarked image. However, in the applications of medical diagnosis, press industry, precious artworks, military service, and remote sensing, it is desired to render the original image without any modification of pixels' intensity after the extraction of watermark from the watermarked image [3, 4]. In hospitals, the physician's diagnosis is strictly conscientious in relation to patient's treatment. Any modification of a medical image, such as X film, MRI, CT, mammogram and so on, might become awfully trouble. Although that the physician's interpretation of an image with an embedded signature will not change might be easy to convince an engineering, but will the same argument is tenable in a court? It is therefore

imperious demand to restore the watermarked image into its original appearance. This has led the development of lossless watermark for authentication [1].

The concept of lossless watermarking was first introduced in US patent owned by The Eastman Kodak in 1999 [4]. Watermarking with modulo addition is considered to complete the reversibility [1]. However, salt-and-pepper artifacts might appear because the underflow or overflow will occur if the result arising from the modulo addition lies outside the range from 0 to 255 for 8-bit gray level image. Reversible data embedding proposed by Tina distinguish an image into three different regions: expandable region, changeable region and not invertible region respectively. The location map about the first region is registered by lossless compression algorithm JBIG2 and it is used to achieve lossless watermark [3]. In [4], an alternative mapping of the histogram to the circle instead of the patchwork histogram rotation was proposed to implement lossless watermarking. These works are inspiring the researches of lossless watermarking.

In the application of image watermarking, there are lots of literature contributed in watermarking both color images and gray images. The reason might be that people are always more interesting in gazing at colorful/gray images than bi-level halftone image. However, for printing, the color document is composed of 4 layers of CYMK and the intensity in each ink layer is continuous. Because almost printing machines can print only binary image, halftoning process is necessary to transfer these continuous image into a binary image. Moreover, in the era of digital publishing, a team of people around the world could be cooperative to fabricate artworks of intelligent design through the Internet. In such a case, the progress of watermarking in halftone image is very important as that in color/gray image.

The paper is organized as follows. Section 2 focuses on the survey of data watermarking for halftone images. Section 3 introduces a novel lossless watermarking that utilizes pattern substitution in combination with block cutting. The experimental results are described in Section 4. Finally, Section 5 concludes with a discussion of the proposed method.

2   Data Watermarking for Halftone Images

Numerous methods of digital watermarking in halftone images have been studied. The techniques of watermarking a halftone image could be primarily categorized to be either block-based watermarking or pixel-based watermarking. As for block-based watermarking methods, the input gray image is divided into blocks. Each block is available to embed one data bit (either 0 or 1) by different embedding scheme. Apparently, the number of watermark pixels is restricted by the limitation of the number of blocks. On the other hands, in pixel-based watermarking methods, the watermarking procedure and the halftoning procedure are joined together so that one halftone pixel can embed one watermark bit per pixel.

In [5], three different classes of fragile data embedding used for authentication were addressed. The policy of the first data hiding is to modify the total number of black pixels of a block to be either odd or even, corresponding to 1 and 0, respectively. The policy of the second data hiding is to tune the total number of white pixels of a block in proportion to a given threshold. If the total number is greater than the given threshold, a data bit 1 is embedded. On the contrary, a data bit 0 is embedded. The policy of the last

one is to perform a consecutive computation of bit-wise exclusive or, pair-wise multiplication, sum and module arithmetic.

In [6], two error diffusion kernels are chosen for halftoning. If data bit corresponding to the processing block is 0 (black), the gray pixels of the processing block are processed by Jarvis kernel. If the data bit relative to the processing block is 1 (white), the gray pixels of the processing block are by Stucki kernel. In decoder, the knowledge of energy distribution of 2-D FFT and detail textures of a block is useful in discriminating the block property.

In [7], Fu and Au proposed two approaches of data hiding in halftone image. One case is the data hiding smart pair toggling when only the halftone image is available as the input source, but both the input gray image and halftoning scheme are unknown. Such a data hiding will toggle the pixel value in a manner that the resulting image is with better visual quality. Another case is data hiding error diffusion when both the input gray image is ready and halftoning scheme is known. This method integrates the data hiding into the procedure of error diffusion so that it can yield higher visual quality and forbid pepper-and-salt effect. Recently, Fu and Au proposed correlation-based watermarking error diffusion (CWED) to embed a watermark. In the case of normal error diffusion, the map of local correlation coefficients of an image seems like random [8]. That means the halftone image is with no watermark inside. In the case of CWED, the map of local correlation coefficients that is equivalent to the phantom of watermark would appear more visible if the correct security code is applied.

3   Lossless Watermarking for Halftone Images

The capability of early existing watermarking works in halftone images almost focus on embedding a watermark into halftone images with a little perceptual degradation and the extraction or detection of the embedded watermark. It is a pity that there is no attention about exactly reconstructing the original halftone image from the watermarked halftone image. In this paper, we firstly assume that the halftone image is the unique input source and it could be split into blocks with lots of items. This results in a data set with higher dimension to represent the halftone image. In practice, these data always cluster together in the higher dimensional vector space. Thus, it leaves us more null coordinates for later usage of watermarking. If the vector (pattern) relative to one cluster and one pattern from the null space are considered to stand for watermark bit 1 and 0, respectively. We can utilize pattern substitution to accomplish lossless watermarking. In next paragraph, the concept of our proposed method based on vector space and cluster is described in detail.

To embed a watermark into a halftone image and fully reconstruct the original halftone image from the watermarked halftone image, a candidate pattern that has high frequency in a halftone image will be treated as a cluster and artificially designated its twin pattern whose hamming distance against candidate pattern is 1 so as to make their difference be small. During watermarking stage, the candidate pattern and its twin pattern are used to denote watermark bit 1 and watermark bit 0, respectively. During the reconstructed stage, candidate patterns are used in place of their twin patterns. Generally speaking, the idea of pattern substitution is feasible to achieve the goal of lossless watermarking in a halftone image.

Firstly, the halftone image is split into M's non-overlaid blocks with size n. That is, each block has n bits. For each block, it is therefore 2^n possible patterns. As 2^n is far greater than M, the patterns of a halftone image cannot fill all grids in n-dimensional and binary vector space. On the other hands, for any image, lots of patterns among these M's patterns are usually grouped into p clusters ($p < M$) due to the pattern identity. These patterns in p clusters are preferentially chosen to embed the watermark bits.

Once p clusters are determined, bit toggling on p candidate patterns is performed bit by bit and from most significant bit to small significant bit. As for a specific candidate pattern, its legal twin pattern must obey three limitations: (1) it isn't identical to one of other (p -1) candidate patterns. (2) Their hamming distance is indeed required to be 1. (3) It is also not identical to the twin patterns that are determined from other (p -1) candidate patterns. Consequently, the number q of clusters for watermarking may be lesser than p ($p < q$). However, the number of pattern in these q clusters is absolutely greater than that of watermarking bits.

The lossless watermarking algorithm is described below.

In the stage of embedding a watermark:

(1) For instance, the size of original halftone image is 256 x 256. It is split into 2^{12} non-overlaid blocks with size 4 x 4.
(2) Calculate the frequency of all block patters. The top 64 patterns possessing higher frequency are chosen as the candidate patterns used for data hiding.
(3) The search of twin pattern is proceeding from the candidate pattern with high rank of frequency. Because the size of a block is 4 x 4, there are 16 possible twin patterns with hamming distance 1 in association with the candidate pattern. If one of these possible twin patterns is not appeared in the list of 64 top patterns and their twin patterns, it will be treated as a legal twin pattern. Otherwise, all 16 possible twin patterns of a candidate pattern are also the twin patterns of other candidate patterns with higher rank of frequency. If one of the twin patterns in association with these candidate patterns with higher rank of frequency could be revised, the processing candidate pattern can get its twin pattern. On the contrary, it is failure to obtain the twin pattern of the processing candidate pattern. These legal candidate patterns are collected as the watermarking area.
(4) To forbid the extraction and destruction of a watermark in a halftone image, a random seed is utilized to permute the bits of a watermark.
(5) As to a watermark, if the number of bit 1(0) of is greater than that of bit 0 (1), all legal candidate patterns are used to embed watermark bit 1(0); their twin patterns are used to embed watermark bit 0 (1). The permutated watermark is then embedded into a halftone image through the pattern substitution in the watermarking area.

Finally, all legal candidate patterns, their twin patterns and random seed are kept for watermark extraction.

In the stage of extracting a watermark:

(1) The watermarked halftone image is split into non-overlaid blocks the same as step (1) in the stage of embedding a watermark.
(2) Apply both all legal candidate patterns and their twin patterns to determine the watermarking area. Thus, the permutated watermark could be extracted bit by bit.

(3) Based on the random seed, the original watermark will be reconstructed via the permutation.

(4) All twin patterns are replaced by their candidate pattern so that the original halftone image could be recovered again.

4   Experimental Results

In this experiment, 4 halftone images shown in Figure 1 are chosen as the test image to demonstrate our proposed method. Figure 2 is the binary watermark containing 4 alphabets. The size of test images is 256 x 256, and that of the binary watermark is 25 x 25. The watermarked images are given in Figure 3. According to the watermark embedding algorithm mentioned in the previous section, the number of candidate patterns is image dependent. The frequency of these potential candidate patterns must be greater than the size (625 bits) of the watermark. Table 1 lists the number and frequency of candidate patterns, the payload versus watermark, and the payload versus image. The payload vs. watermark (image) is defined as the ratio between total numbers of candidate patterns times block size versus the pixel number of the watermark (the test image). For instance, ten candidate patterns and their twin patterns are necessarily required to hide the watermark into test image 'House' in recoverable manner. Its payload vs. watermark is obtained from the computation of (2*10*4*4)/(25*25). Although there is pepper-and-salt in the watermarked halftone images, their original halftone image could be reconstructed if users are authorize to perform reconstruction by the legal authorized key that is composed of the random seed and the pairs of candidate pattern and its twin pattern.

(1A)   (1B)
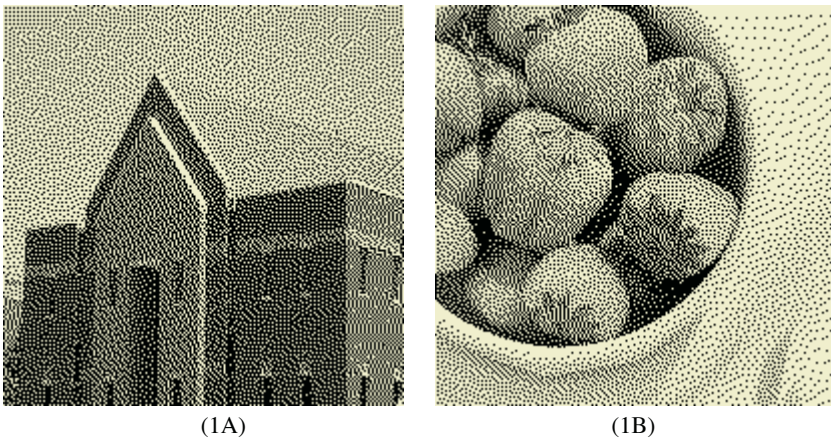
Fig. 1. Test image: (1A) House, (1B) Fruits, (Cont.)

Fig. 2. A watermark with size of 25 x 25

| (2A) | (2B) |
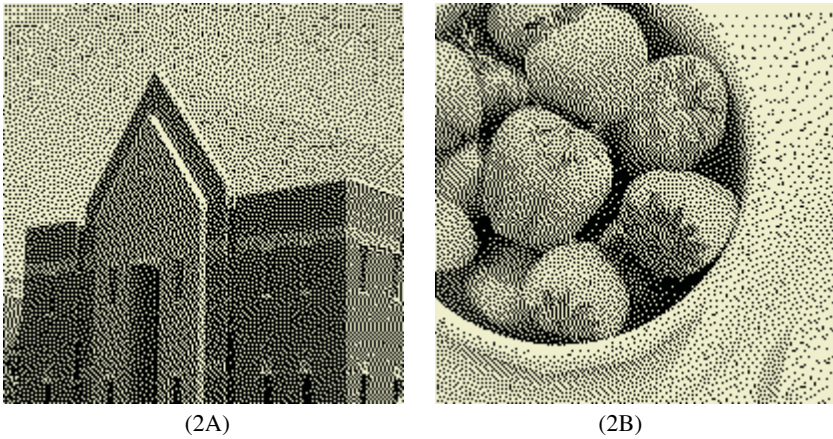
Fig. 3. Watermarked images: (2A) Watermarked House, (2B) Watermarked Fruits, (Cont.)

Table 1. The statistics of the candidate patterns for watermarking

Images	number	frequency	Payload vs. watermark	Payload vs. image
House	10	639	0.512	0.0020
Fruits	13	626	0.666	0.0026
Baby	8	639	0.410	0.0016
Girl	7	650	0.358	0.0014

5 Conclusions

In this paper, authors utilize the similarity of hamming distance between candidate pattern and its twin pattern to fulfill pattern substitution in the procedure of data embedding. Although the watermarked halftone image is with a little perceptive degradation, the original halftone image could be successfully recovered from the watermarked image. In the application of image retrieve, the queried image could be protected by our proposed lossless watermarking before display. In printing, only users with legal authority are allowed to print the quarried image in paper.

References

1. I.J. Cox, M. L. Miller and J.A. Bloom,: Digital Watermarking, Morgan Kaufmann Publishers, (2002).
2. J-S Pan, H-C Huang, L.C. Jain,: Intelligent Watermarking Techniques, World Scientific, (2004).
3. H. Heijmanns and L. Kamstra,: Reversible Data Embedding Based on the Harr Wavelet Decomposition, Proc. VII-th Digital Computing: Techniques and Applications, Sydney, (2003), pp. 5-14.
4. C. D. Vleeschouwer, J-F Delaige and B. Macq,: Circular Interpretation of Bijective Transformations in Lossless Watermarking for Media Asset Management, IEEE Trans. On Multimedia, Vol. 5, No. 1, March, (2003), pp. 97-105.

5. P.W. Wong and N.D. Memon: Image Processing for Halftones, IEEE Signal Processing Magazine, July, (2003), pp. 59-70.
6. S-C Pei and J_M Guo,: Hybrid Pixel-Based Data Hiding and Block-Based Watermarking for Error-Diffused Halftone Images, IEEE Trans. on Circuit and System for Video Technology, Vol. 13, No. 8, (2003), pp. 867-884.
7. M.S. Fu, O.C. Au,: Data Hiding Watermarking for Halftone Images, IEEE Trans. on Image Processing, Vol. 11, No. 4, (2002), pp. 477-484.
8. M.S. Fu, O.C. Au,: Correlation-Based Watermarking for Halftone Images, IEEE International Symposium on Circuits and Systems, (2004), pp. II-21-II-24.

Implementation of Current Mode Fuzzy-Tuning PI Control of Single Phase UPS Inverter Using DSP[*]

Emine Doğru Bolat[1,**] and H. Metin Ertunç[2]

[1] Kocaeli University, Faculty of Technical Education
Department of Electronic & Computer, Kocaeli, Turkey
ebolat@ieee.org
[2] Kocaeli University, Faculty of Engineering
Department of Mechatronics, Kocaeli, Turkey
hmertunc@kou.edu.tr

Abstract. Current mode fuzzy-tuning proportional-integral (PI) control of UPS (Uninterruptable Power Supply) inverter is presented in this paper. Double loop current mode and fuzzy logic controller are combined in this control scheme. Double loop current mode control scheme includes two control loops as voltage and current. Output voltage and filter inductor current are used as feedback variables in voltage control loop and current control loop respectively. In voltage control loop, voltage feedback gain is adjusted by fuzzy-logic controller (FLC). The input variables of the FLC are the output voltage error and its derivative and the output variable of the FLC is voltage feedback gain. This control scheme is simulated using C++ and PSIM and implemented using Texas Instruments TMS320LF2407 DSP (Digital Signal Processor). Simulations and implementation are realized under linear, nonlinear and fluorescent loads. The results prove that this control scheme provides low total harmonic distortion (THD) and fast dynamic response

1 Introduction

Uninterruptible Power Supplies (UPS) are used to produce clean and uninterrupted power to the critical loads such as medical/life support systems, computer systems, communication systems, etc. A typical UPS system consists of a battery (dc source), a dc-ac inverter, and an LC filter. The role of the inverter for UPS is to regulate the output voltage waveform with constant voltage and constant frequency although the variations of the line source or loads. The quality of UPS can be evaluated by THD value of output voltage and characteristic of transient response.

UPS power quality depends on the choice of PWM inverter control methods. Traditional analog control methods are generally used in PWM inverter design. However there are a number of disadvantages in an analog system, for example, temperature drifts and aging effect of the components, more component numbers for the system, necessity for making adjustment to many physical parts, and sensibility to Electro Magnetic Interference (EMI). When an analog circuit is affected by temperature drift or EMI noise, it could cause a number of problems such as dc offset in output voltage,

---

[**] Student Member, IEEE

change of output switching frequency, increase of output voltage harmonics and so on [1]. Therefore, to be able to avoid all these disadvantages the digital current mode fuzzy-tuning PI control scheme is used in this study.

Basic inverter circuit with an LC filter and load, R_L is given in Fig 1. In this Figure, the full bridge inverter, LC filter, and load are considered as the plant to be controlled. r_C is the equivalent series resistor (ESR) of the capacitor, while r_L is the ESR of the inductor. Single phase PWM inverter modulates a DC bus voltage V_{dc} into a cycle by cycle average output voltage V_a. The amplitude of V_a is directly proportional to the commanded duty cycle of the inverter and the amplitude of the dc bus voltage V_{dc}. Therefore, the range of V_a changes between $+V_{dc}$ and $-V_{dc}$ [2].
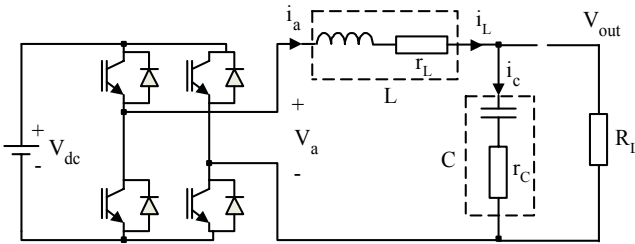
Fig. 1. Basic inverter circuit with an LC filter and R_L

2   Current Mode Fuzzy-Tuning PI Control Scheme

Digital control has such advantages as simplification of control hardware and a corresponding reduction of coast in comparison with analog control. Deadbeat-controlled PWM inverter has very fast response for load disturbances and nonlinear loads. But in deadbeat control approach, the control signal depends on a precise PWM inverter load model and the performance of the system is sensitive to parameter and load variations. Repetitive control generates high-quality sinusoidal output voltage whereas its dynamic response is very slow [3].

Current mode fuzzy-tuning PI control scheme presented in this paper is shown in Fig 2. It has double loop current mode control scheme in core and includes two control loops as inner current and outer voltage loops. The filter inductor current and output voltage are sensed as feedback. In this scheme, note that the output inductor is hidden within the inner current control loop. This simplifies the design of the outer voltage control loop and improves UPS performance in many ways, including better dynamics and a feedforward characteristic which could be used to compensate DC bus ripple and dead time effect, etc. The objective of this inner loop is to control the state-space averaged inductor current. The current mode control requires a circuit for measurement of inductor current i_L, however, in practice such a circuit is also required in voltage mode control systems, for protection of the IGBT against excessive currents during transients and fault conditions [1].

In this study, the advantages of FLC and double loop current mode control are combined in the same control scheme. The FLC can handle nonlinearity and does not need accurate mathematical model. It is represented by if-then rules and thus can

provide an understandable knowledge representation. FLC converts linguistic control strategy to an automatic control strategy. Linguistic control strategy is based on expert knowledge and experience [3].

Inverter output voltage error and change of this error are input variables of the FLC. The input variables, voltage error, e(k) and its derivative, de(k) are calculated using the equations (1) and (2). FLC has one additional input deciding the polarity of the reference sine wave and two rule tables are composed according to the polarity of the reference sine wave. FLC is used to adjust the voltage feedback gain K_f.

$$e(k) = V_{ref}(k) - V_{out}(k) \tag{1}$$

$$de(k) = [e(k) - e(k-1)]/T \tag{2}$$

In equation (1) and (2), $V_{ref}(k)$: reference sine wave, $V_{out}(k)$: inverter output voltage, T: sample time.
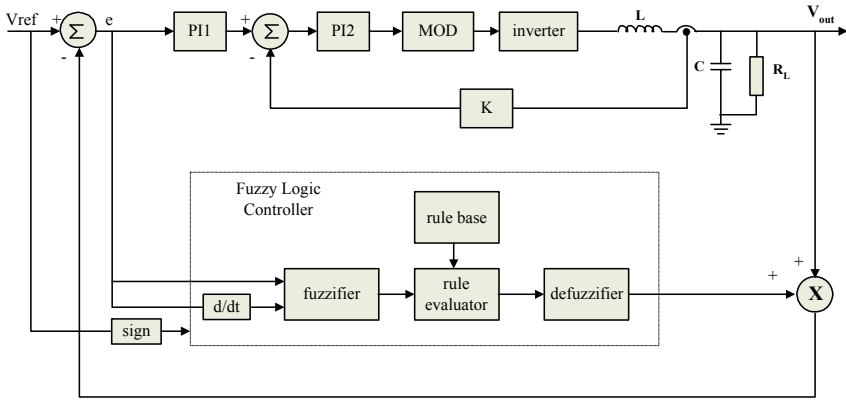
Fig. 2. Current mode fuzzy-tuning PI control scheme
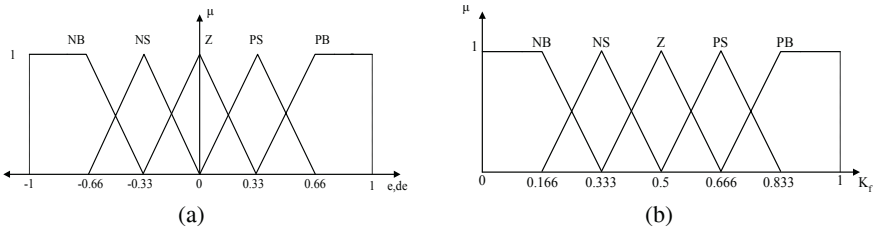
Fig. 3. (a). Membership function of fuzzy input variables, e and de. (b). Membership function of fuzzy output variable, K_f

A typical fuzzy process can be divided into three steps: the fuzzification, the inference and the defuzzification. The task of the fuzzification is to to transform the non-fuzzy measurements of the input variables into the fuzzy linguistic range [4]. In this study, each universe of discourse of two inputs and an output is divided into five fuzzy subsets. These are NB (Negative Big), NS (Negative Small), Z (Zero), PS (Positive Small), PB (Positive Big). The fuzzy subsets and the shape of membership function of FLC input and output variables are seen in Fig 3-(a) and (b). The member-

ship functions are trapezoidal and classical triangular shapes with 50% overlap. The value of each input is normalized in [-1,1] while the output variable is normalized in [0,1] by using suitable scaling factors.

The next step of fuzzy process is the inference mechanism. As fuzzy control rely on knowledge or experience, process behavior is analyzed to establish a set of rules with IF…AND…THEN statements. A typical rule example is:

> **IF** the voltage error (e) is **N**egative **B**ig (NB)
> **AND** change of voltage error (de) is **P**ositive **S**mall (PS)
> **THEN** the voltage feedback gain (K_f) is **P**ositive **S**mall (PS)

These rules define a fuzzy relationship between the observed values of the voltage error, change of voltage error and the voltage feedback gain. Fuzzy control rules are obtained from the behavior of the system and operator's expertise. The rule tables generated according to the polarity of the reference sine wave, V_{ref} are shown in Table 1-(a) and (b). Obtained two rule tables are symmetrical because of the symmetry of the positive and negative half waves of the sine wave. If the polarity of sign input is positive, the rules written in Table 1-(a), otherwise, the rules in Table 1-(b) are used. MAX-MIN method is used as the inference method. The output membership function of each rule is given by MIN operator while the combined fuzzy output is given by MAX operator.

Table 1. (a) Rule table for $V_{ref} \geq 0$, (b). Rule table for $V_{ref} < 0$

(a)
change of voltage error

voltage error		NB	NS	Z	PS	PB
	NB	PB	PB	PB	PS	Z
	NS	PB	PB	PS	Z	NS
	Z	PB	PS	Z	NS	NB
	PS	PS	Z	NS	NB	NB
	PB	Z	NS	NB	NB	NB

(b)
change of voltage error

voltage error		NB	NS	Z	PS	PB
	NB	NB	NB	NB	NS	Z
	NS	NB	NB	NS	Z	PS
	Z	NB	NS	Z	PS	PB
	PS	NS	Z	PS	PB	PB
	PB	Z	PS	PB	PB	PB

The last step is defuzzification process. In defuzzification process, the input for the defuzzification is the fuzzy set (the aggregate output fuzzy set) and the output is a crisp number. Centroidal defuzzification method (center of gravity method (COG)) is used for defuzzification in this paper. Output denormalization converts the normalized value of the control output variable into physical domain. The equation for the COG method is given in equation (3).

$$U_o = \frac{\sum_{j=1}^{n} U(U_j) U_j}{\sum_{j=1}^{n} U(U_j)} \qquad (3)$$

In equation (3), U_o is determined by means of a gravity center of the area under the membership function curve of the fuzzy output and $U(U_j)$ is a membership grade of U_j [5].

3   Simulation and Experimental Results

Current mode fuzzy-tuning PI control scheme explained above is simulated using C++ and PSIM and implemented by using a TI's TMS320LF2407 DSP. Simulations and implementation are realized under resistive, rectifier type nonlinear and fluorescent loads. System parameters used in simulations and implementation are listed in Table 2. Rectifier type nonlinear load is shown in Fig 4. In this figure, L is taken as 770μH.

In this study, the sampling frequency of the current loop is taken as 20 kHz while the sampling frequency of voltage loop is taken as 10 kHz. Simulation and experimental results of current mode fuzzy-tuning PI control scheme for different kinds of loads are shown between Fig 5-(a) and Fig 5-(f). In simulation and experiments, linear load is applied to the inverter with a firing angle of 108°.

Table 2. System parameters used in simulations and experiments

V_O (output voltage)	50Hz 220 V_{RMS}
V_{DC} (DC bus voltage)	360 V
f_S (sampling frequency)	20 kHz
L (filter inductor)	700 μH
C (filter capacitor)	30 μF
r_L (ESR of the inductor)	0.05 Ω
r_C (ESR of the cap.)	0.02 Ω
Resistive load	10 Ω (5 kVA100%)
Nonlinear load (rectifier type)	5 kVA (100%)
Fluorescent load	2.6 kVA (50%)
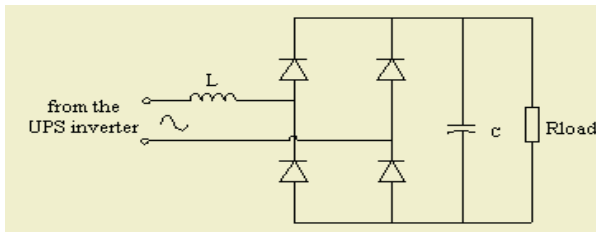
Fig. 4. Rectifier load with R_{load}-C

Table 3. THD values measured in simulations and experiments

	Linear Load (100%)		Nonlinear Load (100%)		Fluorescent Load (50%)	
	Sim.	Exp.	Sim.	Exp.	Sim.	Exp.
THD (%)	0.48	1.45	0.93	1.54	0.37	1.45

In Table 3, THD values measured in simulations and experiments for different kinds of loads are given.

(a)

(b)

(c)

(d)
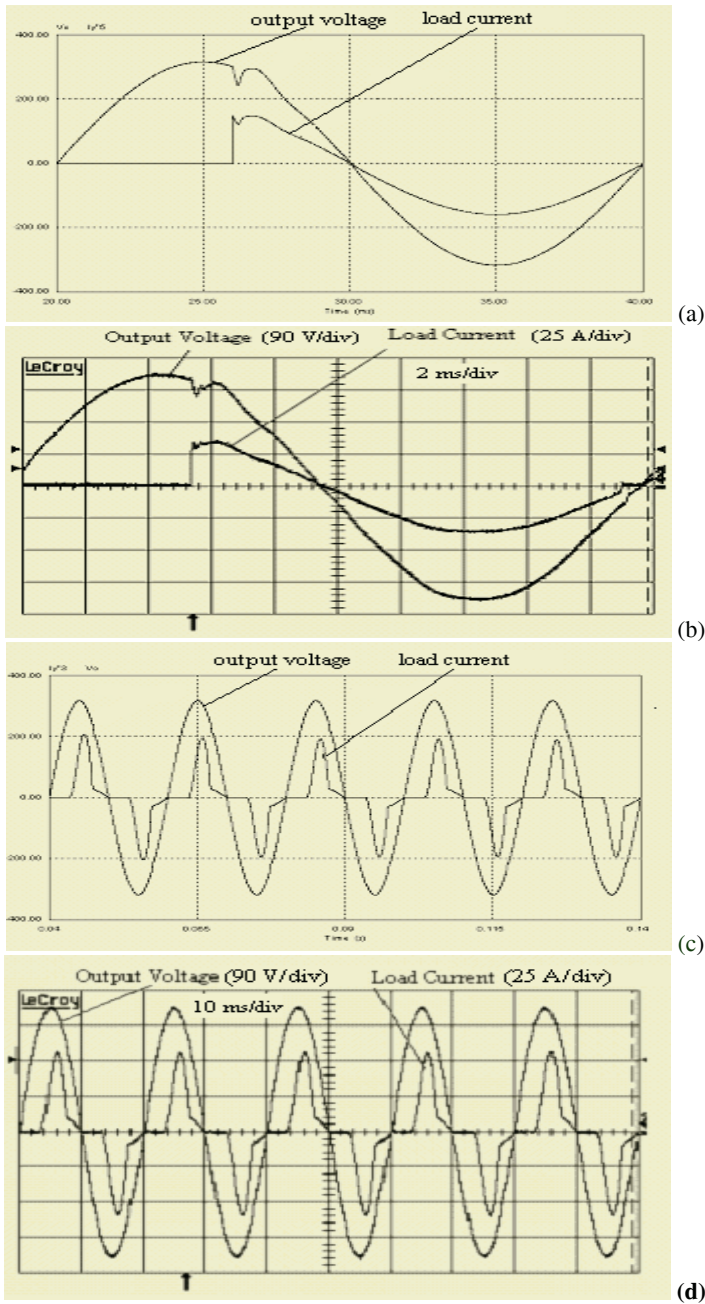
Fig. 5. Simulation and experimental results of output voltage and load current.(a). Simulation results for linear load (THD = 0.48%) (b). Experimental results for linear load (THD =1.45%). (c). Simulation results for rectifier type nonlinear load (THD = 0.93 %). (d). Experimental results for rectifier type nonlinear load (THD =1.54%). (e). Simulation results for fluorescent load (THD = 0.37%). (f). Experimental results for fluorescent load (THD =1.45%)

(e)

(f)
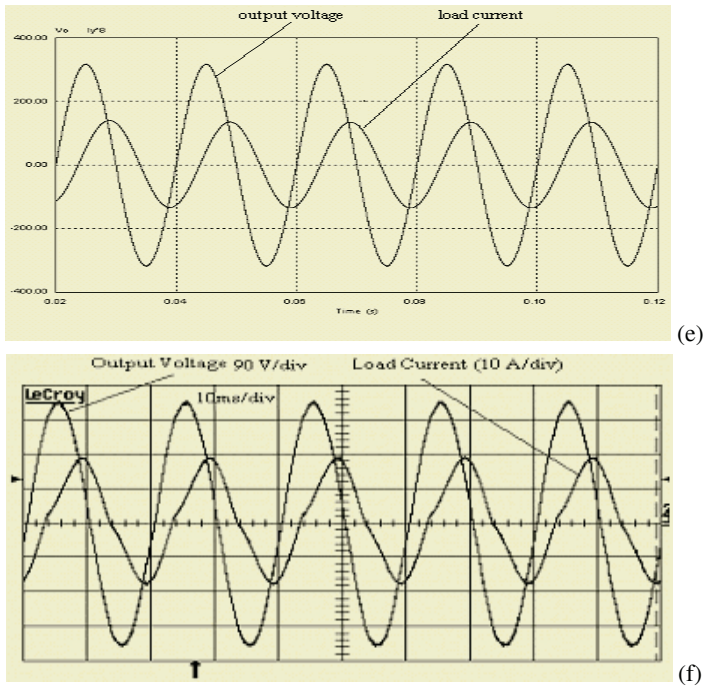
Fig. 5. Continued

4   Conclusion

In this paper, a DSP based digitally controlled PWM inverter is implemented using current mode fuzzy-tuning PI control scheme. This scheme is simulated using C++ and PSIM and implemented using TI's TMS320LF2407 DSP. Double loop current mode control scheme is used as main controller and FLC is used to adjust the feedback gain of the voltage control loop. The output voltage error and its derivative are used as input variables of the FLC. Output voltage regulations under all kinds of loads in simulations and experiments are satisfactorily good as shown in Fig 5. As shown in Table 3, THD value below 5% could be achieved under nonlinear load. And these results prove that this control scheme could enable low THD values, good voltage regulation and fast dynamic response to different kinds of loads.

References

1. Jiang, H.J., Qin, Y., Du, S.S. Yu, Z.Y., Choudhury, S.: 1998, "DSP Based Implementation of a Digitally Controlled Phase PWM Inverter for UPS", pp. 221-224,, INTELEC Twentieth International Telecommunications Energy Conference, INTELEC.
2. Choudhury S.:1999, "Implementing Triple Conversion Single-Phase On-line UPS using TMS320C240", Texas Instruments Application Report.
3. Jian L., Yong K., Jian C.:2000, "Fuzzy-Tuning PID Control of an Inverter with Rectifier-Type Nonlinear Loads", pp. 381-384 vol.1, PIEMC 2000, Power Electronics and Motion Control Conference.

4. Osterholz H.: 1995, "Simple Fuzzy Control of a PWM Inverter for a UPS System", pp. 565-570. 17th International Telecommunications Energy Conference. INTELEC '95.
5. Kelkar N.D:1997, "A Fuzzy Controller for Three Dimensional Line Following of an Unmanned Autonomus Mobile Robot", Thesis of master of science, Department of Mechanical, Industrial, and Nuclear Engineering, University of Cincinnati,Ohio.
6. Manning C D.:1994, "Control of UPS Inverters", pp. 3/1-3/5, IEE Colloquium, The Institution of Electrical Engineers, printed and published by the IEE, Savoy Place, London WC2R OBL,UK.
7. Doğru Bolat E, Bolat Y., Ertunç H. M.:2004, "Double Loop Current Mode Control Scheme for Single Phase UPS Inverter", Mechatronics 2004, The 9th Mechatronics Forum International Conference, Ankara, Turkey.

Knowledge-Based Fuzzy Control of Pilot-Scale SBR for Wastewater Treatment

Byong-Hee Jun[1], Jang-Hwan Park[2], and Myung-Geun Chun[3,*]

[1] Institute of Construction Technology, Chungbuk National University, Cheongju, Korea
[2] School of Electrical and Electronic Engineering
Chungju National University, Chungju, Korea
[3] Dept. of Electrical and Computer Engineering
Chungbuk National University, Cheongju, Korea
mgchun@chungbuk.ac.kr

Abstract. A fuzzy controller to optimize oxic phase of sub-cycle in pilot scale SBR (working volume, 20m^3) located at public swine wastewater treatment plant was investigated. The operation mode of intermittent feeding of raw water and sub-cycle with repeating anoxic-aeration conditions were adapted to avoid the high-strength nitrogen inhibition. In sub-cycle, aeration time for ammonium oxidation was tried to control by fuzzy system using DO value and its differential values for input membership function. In previous study, a lag time of DO profile was proved to be useful for inference of nitrogen loading rates and abrupt raise of DO showed potential ability to detect ending point of ammonia oxidation, however, it also needs proper safety factor for high level DO in lag time or noisy system. Here, the fuzzy system could reduce the complexity of application. As results, the fuzzy system with the simultaneous application of DO value and its differential values was appropriate for the stable control of SBR.

1 Introduction

As the most of large scale stalls are located at the up stream of drinking water intake and scattered around the broad area, their effective management and control have been very difficult. For these reasons, much effort has been provided to solve the livestock waste management problems, especially swine wastewater, through wastewater treatment along with resources recovery like composting.

To accommodate various characteristics of swine wastewater, the livestock wastewater treatment system should have flexibility for the amount and nature of influent. A large number of treatment systems, such as anaerobic [1][2], intermittent aeration [3] and crystallized precipitation [4], for the livestock wastewater have been reported. Due to its flexibility in operation, SBR (sequencing batch reactor) is also known to be suitable. Easy change of operation conditions such as influent rate and simultaneous removal of organic matter, nitrogen and phosphorus in single basin are the unique advantages of SBR. In addition, no need for settling tank and ability to discharge treated wastewater after effluent quality come to be satisfied are another attractive merits. This space-saving characteristic of SBR is adaptable for the mountain area application, which many of livestock house are located [5].

---

* Corresponding author

Nutrient removal process can be monitored with auxiliary measurable variables. For example, ORP represents the overall reactions of biological and physiochemical processes, many researchers explored efficient and reliable automatic control approaches of biological nitrogen removal by monitoring ORP [6][7]. Additionally, pH and DO were also applied to detect ending point of nitrification, but its application was not extended to control parameter [8][9]. However, Jun et al. [10] revealed that DO was one of the major control parameters of the sequencing nitrogen removal process, and the control strategies using ORP or DO with threshold comparison method were proposed. This threshold comparison method containing single input parameter such as DO value, DO differential values or ORP differential values, could be used in simple way but too dependent on each reactor characteristic to require fine-tuning of threshold value for optimal operation [11].

In this study, the ORP and DO on-line monitoring parameters were selected to develop a control strategy of aeration phase for nitrification. Because the advantages or disadvantages of each control parameter were investigated in previous studies, more improved control strategy can be extracted for fuzzy system. The aim of this work was to develop fuzzy controller with good performance using human knowledge for aeration process in sub-cycle.

2   Sequencing Batch Reactor (SBR) and Fuzzy Logic Controller

Fig. 1 shows the structure of the remote management system for several SBR plants. On site a local control system controls the operating cycle of SBR for completing nitrification and denitrification. At the remote management system, the influent loading rate and the appropriate aeration time can be managed by the fuzzy inferred system and artificial neural network (ANN). For model programming C++ Builder was used and proved that it was suitable for DB management and model implementation.
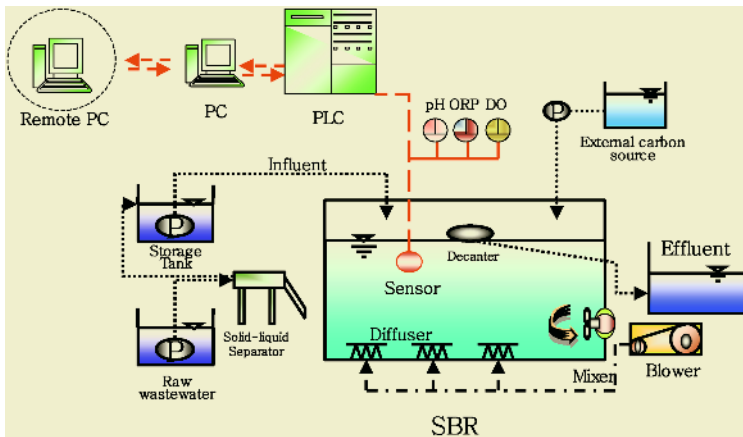
Fig. 1. Schematic diagram of the pilot scale SBR plant

A classical fuzzy logic system use the rules generated from human expert knowledge. In previous studies, Jun *et al*. [12] revealed that DO was one of the major con-

trol parameters of the sequencing nitrogen removal process, and the control strategies using ORP or DO with threshold comparison method were proposed. This threshold comparison method containing single input parameter such as DO value, DO differential values or ORP differential values, could be used in simple way but too dependent on each reactor characteristic to require fine-tuning of threshold value for optimal operation. To produce a smoother performing and fault detective system, the fuzzy controller was applied.
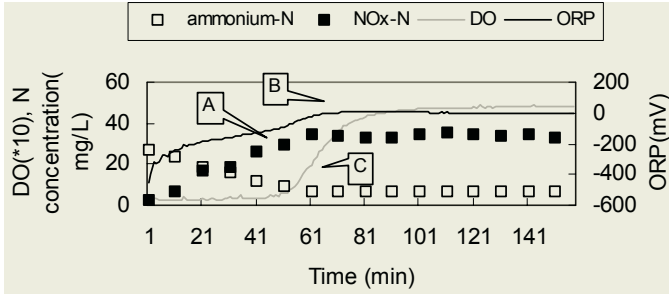
Fig. 2. The variation of DO and ORP during nitrification in SBR

Three kinds of control point relating different control strategies in the precious studies were defined as ORP bending point (A), ORP plateau (B) and DO value (C) [13]. Due to good agreement with EPAO, the bending point on ORP (A) has been widely used for EPAO detection. However, since the bending point was strongly dependent on the influent loading rates and often not appeared in relatively low loading rates, the control strategies based on the ORP bending point was found to be inappropriate with wide range of influent loading rates. The remarkable DO variation (C) from under 1 to 5mg/l at EPAO showed the applicability of more easy and clear detection of EPAO using DO profile than ORP. As shown in Fig. 2, DO profile showed duration of lag time, which means progressing of ammonium oxidation, and abrupt increasing in accordance with EPAO, proved to be good monitoring criteria. To establish more rapid, robust and reproducible control model against these disturbances, the advantages of each control strategy with DO value and the DO differential values were required to be integrated.

In this study ORP and ORP differential values along with DO values were estimated as input parameters for fuzzy system. For the generation of the membership functions of input parameter, the information such as DO and ORP value, which extracted from pilot plant were used. Table 1 showed the inference rules for aeration control. NORMAL range of DO meant the ending point of ammonium oxidation (EPAO) on DO profile and corresponded to blow off command.

Table 1. The Inference rule for aeration control

		DO differential values		
		Low	Normal	High
DO value	Low	ON	ON	OFF
	Normal	ON	OFF	Abs-OFF
	High	Abs-OFF	Abs-OFF	Abs-OFF

3   Fuzzy Control for Wastewater Treatment

3.1   Design of Fuzzy Logic Controller

A membership functions for the inputs were defined from DO and ORP profile shown in Fig. 2 with the knowledge of the human operator while the output membership function were defined as two signals: blow-on or blow-off. Fig. 3 represented the output of fuzzy controller for each input variable of DO value (A), DO differential values (B) and ORP differential values (C), respectively.

(a)   (b)   (c)
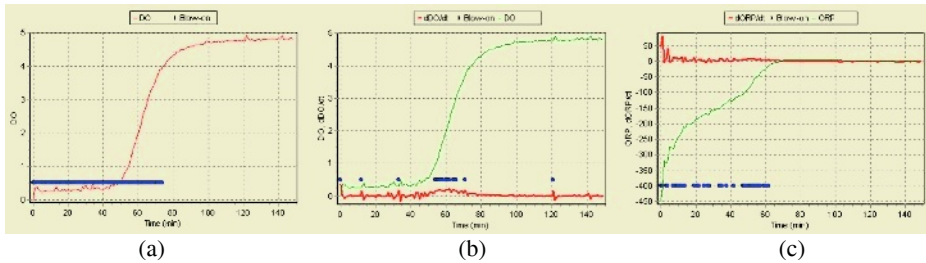
Fig. 3. The comparison of blow-on commands by different input parameters; (A) DO, (B) dDO/dt, (C) dORP/dt

The input membership function with DO value (A) generated stable and continuous blow-on signals (filled circle). However, it was thought that since high DO concentration during lag time by low MLSS or high strength aeration may cause the control failure, blow-on signal was continued to relatively high DO range of 3 to 4mg/l for operation safety. To eliminate excess aeration, a fine-tuning on membership function for each specified reactor may be required. The input membership function with DO differential values (B) showed discrete blow-on signals but was proved to have a potential to detect the EPAO specifically regardless of DO concentration during the lag time and as results, the fine-tuning was expected to be omitted. The advantages of each input parameter such as continuity of DO value and early detection of EPAO of DO differential values were expected to useful for build up the fuzzy rule set.

The input membership function with ORP differential values (C) showed less stability, less continuity and less specified control point. As results, DO value and DO differentials were chosen as an appropriate input parameter of membership function for aeration control.

3.2   Applicability of Fuzzy Logic Controller

Fig. 4 showed the control results of pilot plant under various operation conditions using single same fuzzy controller without tuning. Fig. 4 (A) showed the operation example under relatively stable and typical conditions. The solid line represents the DO profiles by threshold comparison method using only DO value and the filled circle represented the signal for aeration process with fuzzy controller using DO value and DO differentials simultaneously. Therefore the solid line without filled circle means the saved time by reduced aeration process. The fuzzy controller detected the change of DO differentials along with DO value and rapid control was realized.

Fig. 4 (B) showed the control result under temporal noisy status. During the DO lag time, a DO peak reaching to 3mg/l was appeared by unidentified reason. If the threshold comparison method using only DO value was used, the operation would stop the aeration and transferred to next anoxic phase with uncompleted nitrification. However, the fuzzy controller involving DO value and DO differential values as input variable could eliminate the abrupt temporal noise by given fuzzy set.

Fig. 4 (C) showed the application of fuzzy controller for the noisy DO value and high DO concentration during the lag time. As the undesired disturbances in system is often appeared, the control point of blow off must be set to high value not to be disturbed, and as results the aeration time was increased. However, in proposed fuzzy system could separate the real value from noise and capable to shorten the aeration within EPAO point.

Fig. 4 (D) showed the control result when wastewater was not fed in anoxic phase. In this case, DO was abruptly increased with the beginning of aeration and DO lag time was not appeared. In threshold comparison method, there was constrict aeration time of about 30 minutes to avoid process fault during lag time in previous study and excess aeration was continued until this constrict time and fault diagnosis also delayed. Because the fuzzy controller did not require this constrict time, early fault detection and aeration stopping was possible. All the results in Fig. 4, which were produced with single same fuzzy controller without fine-tuning, showed the high applicability to SBR for swine wastewater treatment.
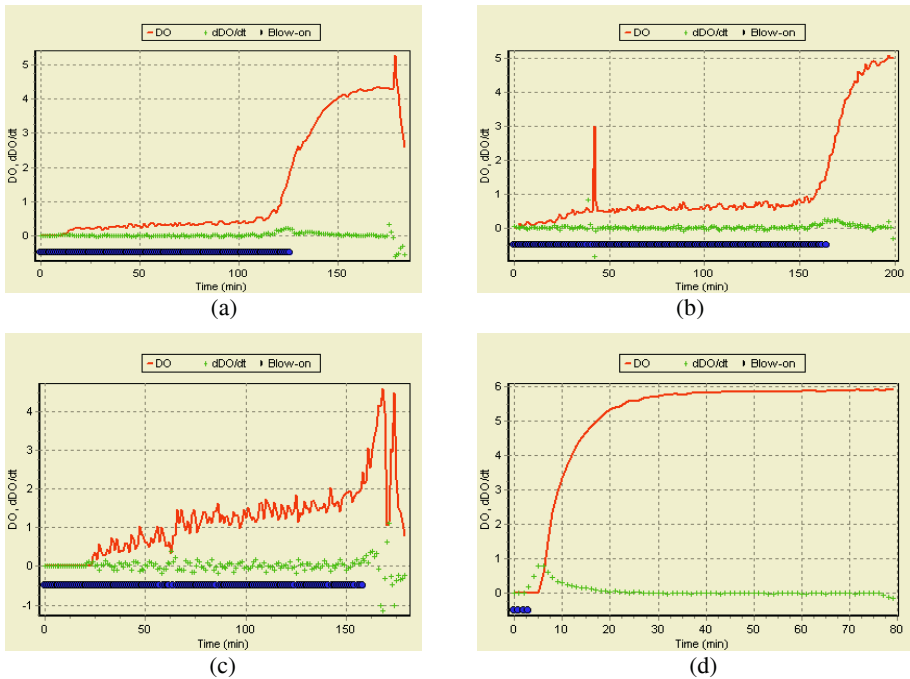
Fig. 4. Applicability of fuzzy controller in stable operation (A), noisy case (B), relatively noisy and high DO (C) and influent fault case (D)

4   Concluding Remarks

To develop the advanced control strategy for aeration process in sub-cycle of SBR operation, a fuzzy system was employed to on/off control of aeration. The relationship of on-line measurable variables and nitrogen concentration was investigated. There was a good relationship between DO, ORP and ammonium oxidation and three kinds of control point relating different control strategies in the precious studies were defined: ORP bending point, ORP plateau and DO value. The ORP bending point, which has been widely used for EPAO detection, was not appeared under relatively low loading rate. To establish more rapid, robust and reproducible control model against disturbances, the advantages of each input parameter of DO value and DO differential values were involved to fuzzy set.

The fuzzy controller detected the change of DO differential values along with DO value, so rapid and stable control was possible. Moreover, the proposed fuzzy system could minimize the aeration time within EPAO and showed the broad applicability in high DO concentration during the lag time and noisy system without additional fine-tuning.

Acknowledgements

References

1. Lee, S. M., Jung, J. Y. and Jung, Y. C., "Novel method for enhancing permeate flux of submerged membrane system in two-phase anaerobic reactor", *Wat. Res*. Vol. 35, pp.471-477, 2001.
2. Andara, A. R. and Lomas Esteban, J. M., "Kinetic study of the anaerobic digestion of the solid fraction of piggery slurries", *Biomass and Bioenergy* Vol. 17, pp.435-443, 1999.
3. Yang, P. Y. and Wang, Z., "Integration an intermittent aerator in a swine wastewater treatment system for land-limited conditions", *Bioresource Technology* Vol. 69, pp.191-198, 1999.
4. Maekawa, T., Liao, C. M. and Feng, X. D., "Nitrogen and Phosphorus removal for swine wastewater using intermittent aeration batch reactor followed by ammonium crystallization process", *Wat. Res*. Vol. 29, pp.2643-2650, 1995.
5. Norcross, K. L., "Sequencing batch reactors-An overview", *Wat. Sci. Tech*., Vol.26, pp.2523-2526, 1992.
6. Plisson, S. S., Capdeville, B., Mauret, M., Deguin, A. and Baptiste, P., "Real-time control of nitrogen removal using three ORP bending points: signification, control strategy and results", *Wat. Sci. Tech*. Vol. 33, No. 1, pp.275-280, 1996.
7. Caulet, P., Bujon, B., Philippe, J. P., Lefevre, F. and Audic, J. M., "Upgrading of wastewater treatment plants for nitrogen removal: industrial application of an automated aeration management based on ORP evolution analysis", Wat. Sci. Tech., Vol. 37, No. 9, pp.41-47, 1998.
8. Andreottola, G., Foladori, P. and Ragazzi, M., "On-line control of a SBR system for nitrogen removal from industrial wastewater", *Wat. Sci. Tech*. Vol.43, pp.93-100, 2001.

9. Paul, E., Plisson, S. S., Mauret, M. and Cantet, J., "Process state evaluation of alternation oxic-anoxic activated sludge using ORP, pH, and DO", Wat. Sci. Tech. Vol. 38, No. 3, pp.299-306, 1998.
10. Jun B. H., Kim D. H., Chio, E. H., Bae H., Kim S. S. and Kim C. W., "Control of SBR operation for piggery wastewater treatment with DO and ORP", *J. of Kor. Soc. on Water Quality*. Vol.18, pp.545-551, 2002.
11. Serralta, J., Ribes, J., Seco, A. and Ferrer, J., "A supervisory control system for optimizing nitrogen removal and aeration energy consumption in wastewater treatment plants", *Wat. Sci. Tech*. Vol. 45, pp.309-316, 2002.
12. Jun B. H., Poo K. M., Chio E. H., Lee H. I. and Kim C. W., "High-performance SBR operation by optimized feeding method of external carbon source for piggery wastewater treatment", *J. of Kor. Soc. of Environmental Eng*. Vol.24, No.11, pp.1957-1964, 2002.
13. Jun B. H., Kim D. H., Chio, E. H., Bae H., Kim S. S. and Kim C. W., "Control of SBR operation for piggery wastewater treatment with DO and ORP", *J. of Kor. Soc. on Water Quality*. Vol.18, pp.545-551, 2002.

Finely Tuned Cascaded Fuzzy Controllers with VHDL – A Case Study for Linerization of V-I Characteristics of a Convertor

Avinash G. Keskar[*], Kishor Kadbe[**], Nikhil Damle[**], and Pooja Deshpande[**]

Visvesvaraya National Institute of Technology, Nagpur, India

Abstract. This work proposes cascaded Fuzzy Logic Controllers as new hardware architecture so that it works in an accurate and fast manner. The MAX-MIN interface can be simplified as lookup table operations. Both membership function values and index numbers of input MF's are pre-stored in the lookup table and the observed crisp inputs serve as the addresses for reading them. Each output MF is treated as a singleton located at the center of the MF. So, the computational complexity of COG defuzzification is significantly reduced. As a case study for a single-phase converter, the fuzzy method of $\Delta\propto$ angle compensation is used to linearize the converter transfer characteristics in discontinuous conduction mode. On the basis of the information collected from the graphs; the fine tuning of the output of FLC1 is suggested to achieve better linearization. The required fine-tuning is achieved by using another fuzzy controller FLC2 cascaded with first fuzzy controller FLC1. All components of the FLC are simulated in VHDL using the concept of moment of equilibrium principle for MAX and MIN modules. The VHDL analysis shows faster response to conclude that dedicated cascaded hardware processor will give very accurate linear control for the converter under study.

Circuit Theory

The power circuit consists of a phase-controlled bridge converter that drives a separately excited dc motor as in figure 1. The converter may operate in either continuous or discontinuous conduction mode [1, 10]. At low speed when the counter emf is small, the conduction will be continuous. However, at high speed, the conduction will tend to be discontinuous [ref]. Conventionally,

I_a = armature current (average)
$\propto$ = Converter firing angle
V_d = converter output voltage (average)

First we assign one particular number to different fuzzy subsets of $\propto$, Ia and $\Delta\propto$ as follows.

In the traditional way, if rule is written, it is as follows. If Ia is small negative (NS) AND $\propto$ is small positive (PS) THEN $\Delta\propto$ is small negative (NS). Now, we can write this as follows: IF Ia is 3 AND $\propto$ is 4 THEN $\Delta\propto$ is 4. In this way we now store the rules in the form of numbers. This is shown below in fig. 2.

---

[*] Professor
[**] Research Scholars
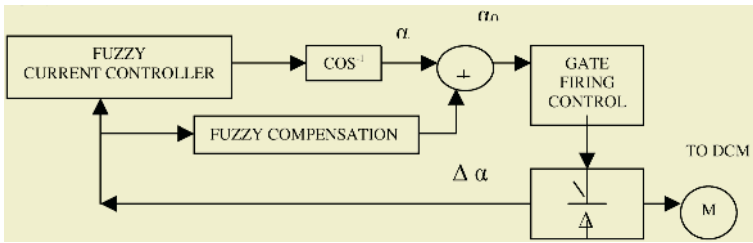
Fig. 1. Fuzzy controlled dc drive

For ∝:				NB	NS	Z	PS	PB			
				1	2	3	4	5			
For I_a		NVB	NB	NM	NS	Z	PS	PM	PB	PVB	
		1	2	3	4	5	6	7	8	9	
For Δ∝	NVB	NB	NM	NS	NVS	Z	PVS	PS	PM	PB	PVB
	1	2	3	4	5	6	7	8	9	10	11

Fig. 2. Membership Function Assignment

The rule base is developed by heuristics [2, 3] from the viewpoint of practical system operation. The current I_a is treated normalized value. The universes of discourse of the variables cover the whole discontinuous conduction region. The sensitivity of the variable determines the number of fuzzy subsets. The universe of discourse of $\propto$ is described by five fuzzy subsets, whereas I_a (pu) and $\Delta\propto$ are described by nine and eleven subsets, respectively. 50% overlap has been provided for the neighboring fuzzy subsets. Therefore, at any given point of the universe of discourse, choice of fuzzy partitioning along with the SUP-MIN composition method uses the linearization algorithm. Since the firing angle $\propto_0$ is mainly concerned with the linearity of transfer characteristic of Vd-Ia (pu) at discontinuous mode, it is necessary to linearize the $\propto_0$-$\propto$ characteristic. Hence for this purpose, fine tuning of $\propto_0$ is suggested.

Fine Tuning

A new fuzzy controller FLC2 is used to achieve fine tuning. The inputs for FLC2 are $\propto_0$ i.e. the output of FLC1 and the armature current Ia (pu) and the output of FLC2 is denoted by $\Delta^2\propto$ as shown in figure 3.

Lower Bound of $\Delta\propto=0$ and Lower Bound of $\propto=0^o$ Hence, Lower Bound of $\propto_0=0^o$ (since $\propto_0=\propto+\Delta\propto$)

Similarly, Upper Bound of $\Delta\propto=30^o$ and Upper Bound of $\propto=90^o$, Hence, Upper Bound of $\propto_0=120^o$ (since $\propto_0=\propto+\Delta\propto$)

Membership functions for $\Delta^2\propto$ is again kept same as $\Delta\propto$. The derivation is purely heuristic in nature and relies on the qualitative knowledge of process behavior [2]. The value of $\Delta^2\propto$ is found out by FLC2 in the same way as FLC1. In above method, an auxiliary compensating angle $\Delta\propto$ is generated as a function of main $\propto$ angle and armature I_a (pu) and then added with the firing angle $\propto$ to generate the $\propto_0$ angle. In this way, the linearisation of Vd - I_a (pu) characteristic mainly depends on the firing angle $\propto_0$.
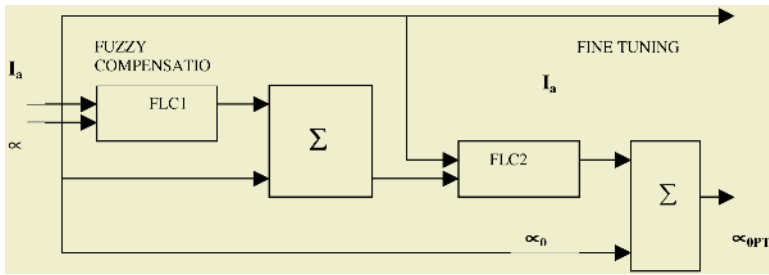
Fig. 3. Modified system for Fine tuning i.e. Fuzzy $\Delta^2\propto$ Compensation
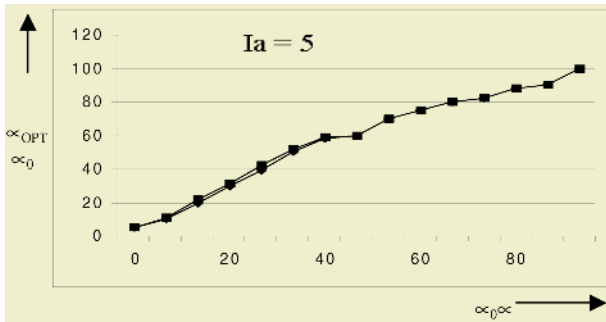
Fig. 4. Comparative Plot between $\propto_{OPT}$-$\propto$ and $\propto_0 \propto$

For the fuzzy $\Delta\propto$ compensation method, graphs are plotted between $\Delta\propto$-$\propto$, $\Delta\propto_0$-$\propto$ for different values of armature current I_a (pu). As can be seen from the graph, particularly for lower valued I_a (pu) the plot between $\propto_0$-$\propto$ is becoming nonlinear around some portion of characteristic. For comparison purpose, again the graphs are plotted between $\Delta^2\propto$- $\propto$, $\propto_{OPT}$- $\propto$ for different values of armature current Ia (pu). From the comparative study of the graphs, it is observed that Fine Tuning of firing angle $\propto_{OPT}$ has greatly reduced the nonlinearity in $\propto_0$-$\propto$ characteristics maintaining the already available linearity. From the graph shown in figure 4, it is observed that the plot between $\propto_0$-$\propto$ is nonlinear around the dotted circle After fine tuning of firing angle the linearity is achieved shown by $\propto_{OPT}$ $\propto$ plot. Both plots are shown simultaneously. Figure 5 shows the plots between $\Delta\propto$- $\propto$ and $\Delta^2\propto$- $\propto$ simultaneously to indicate that the finer compensating angle $\Delta^2\propto$ is finely tuned along and conventional compensating angle$\Delta\propto$.

For the portion for which $\Delta\propto$- $\propto$ plot is remaining constant (shown by dotted circle in figure 5) causes the nonlinearity in $\propto_0$-$\propto$ plot (because $\propto_0=\Delta\propto + \propto$). But the finer compensating angle $\Delta^2\propto$ is tuned around $\Delta\propto$ in such a way that $\Delta^2\propto$- $\propto$ plot varies linearly instead of becoming constant around that portion for which $\propto_0$-$\propto$ plot is nonlinear. While doing this, special care is taken so that already available linearity should not get disturbed [7-9]. This can be seen in figure 4, where $\propto_{OPT}$- $\propto$ plot is perfectly coinciding with $\propto_0$-$\propto$ plot for the portion where the linearity was already available. The $\propto$ values are given for different values of Ia. Figure 6 shows corresponding $\Delta\propto$- $\propto$ and $\Delta^2\propto$- plots.
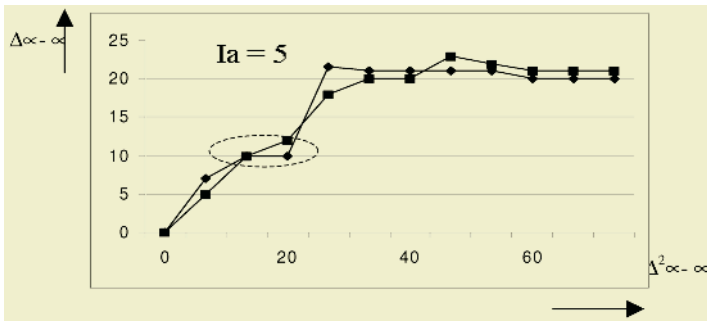
Fig. 5. Plot between $\Delta\propto-\propto$ and $\Delta^2\propto-\propto$
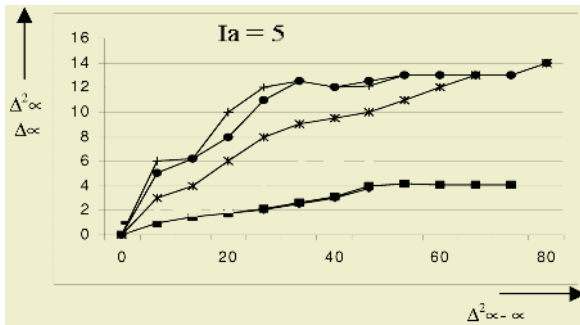
Fig. 6. Various plots between $\Delta\propto-\propto$ and $\Delta^2\propto-\propto$

VHDL Simulation

Figure 7 shows the VHDL simulation environment for fine tuning of firing angle controlled by the proposed FLC that consists of two different blocks: 1) FLC compo-nent with MIN, MAX, Inference modules based on moment of equilibrium concepts [4,5,6]. 2) Model of phase controlled converter. The VHDL simulation with ACTIVE-HDL 4.1 Aldec simulator is used for verifying that proposed FLC is cor-rectly designed and is functionally working correctly. Firstly, each module of pro-posed FLC is described with VHDL in the structural level, and knowledge base of application is described with VHDL in behavioral level. The graphs in figures 8 to 10 indicate the simulation in less clock cycles as observed.
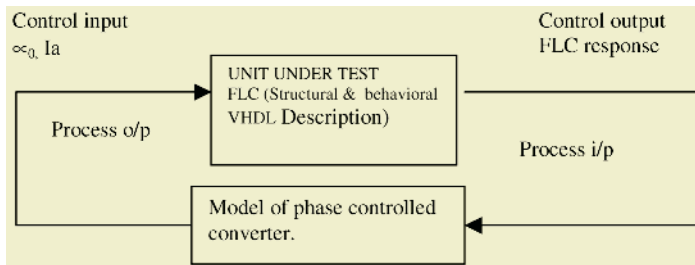
Fig. 7. VHDL simulation environments

TEST INPUT:
∝ (firing angle)   xi=27 (or 1B H in hexadecimal code format) and
Ia (armature current) qi=24 (or 18 H in hexadecimal code format)
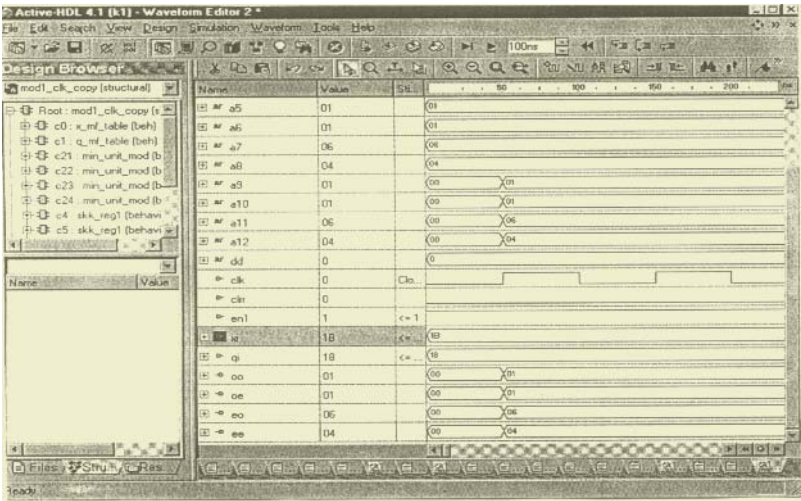
Fig. 8. The simulation result for MIN Module

RESULT:
MINoo=1, MINoe= 1, MINeo=6, MINee=4. (o==odd, e=even) clock cycles to evaluate result: 1
The simulation result for Inference Module is shown in figure 9
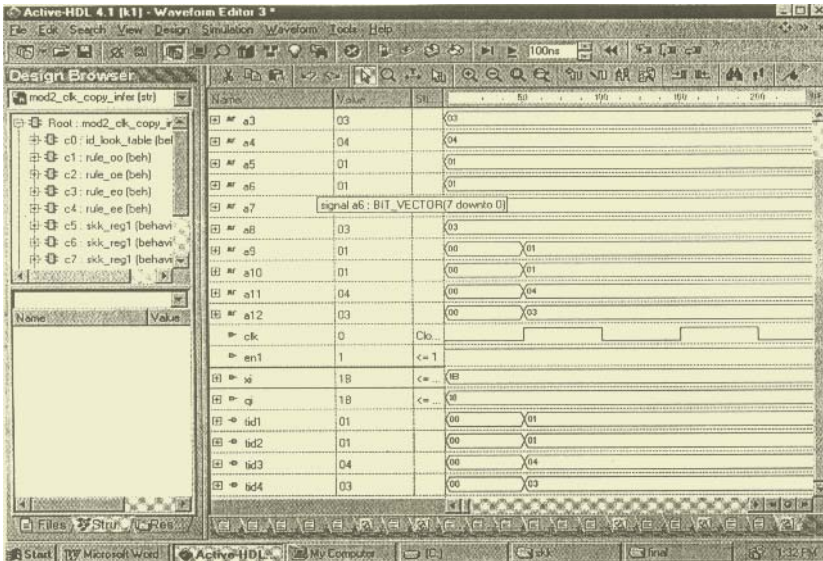
Fig. 9. Simulation result of Inference model

TEST INPUT:
xi=27 (or 1B H in hexadecimal code format) and qi=24 (or 18 H in hexadecimal code format)
RESULT:
TERM-ID1=01, TERM-ID2=01, TERM-ID3=04, TERM-ID4=03.
These results are represented as tid1, tid2, tid3, and tid4 in figure 9 .Required clock cycles:1 (ONE)
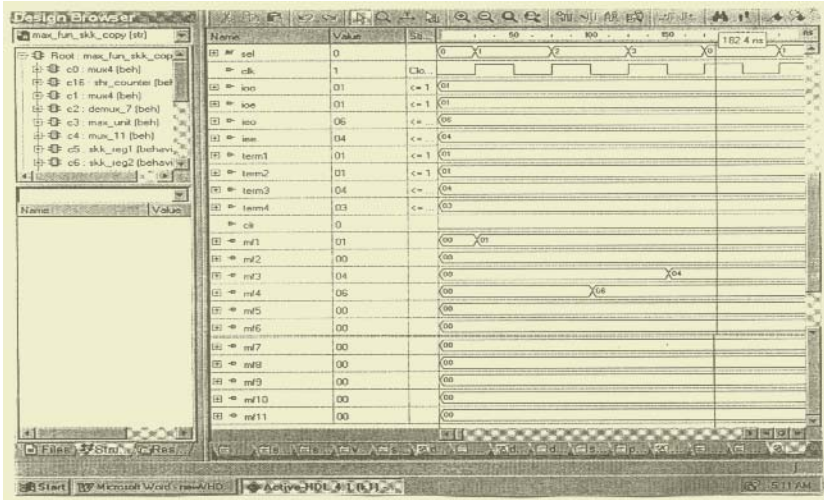
Fig. 10. The simulation result for MAX Module

TEST INPUT:
Ioo=01,Ioe=01,Ieo=06,Iee=04 ; Term1=01, Term2=01, Term3=04, Term4=03,
RESULT:
mf1=01, mf2=00, mf3=04, mf4=06, mf5=00, mf6=00, mf7=00, mf8=00, mf9=00, mf10=00, mf11=00
Required clock cycles for evaluating the result: 4 (FOUR)

Conclusion

Incorporating intelligent reasoning systems on hardware is significant because future expert systems may have to make decisions in real time. We have presented a step-by-step approach towards realizing fuzzy expert systems in VLSI circuitry. We have studied each module (MIN, MAX, Inference, and defuzzifier) considering hardware complexity and speed of operation and verified its hardware and functionality using VHDL. A very simple and quick height defuzzification method is promoted which makes the FLC hardware very simple and quick. The designed FLC is tested for the practical application of controlling firing angle compensation technique necessary for the linearization of transfer characteristic at discontinuous conduction mode. This is especially true where at low speeds linearization is required since the current is essentially discontinuous at such speeds due to the very physical nature of electrical drive.

We also studied the application and suggested fine tuning of firing angle. Cascaded fuzzy logic controllers are used to achieve fine tuning which indicates superiority of cascaded fuzzy system for fine control.

Future Expectations

This two-input-single output FLC can be extended to MISO (Multiple-Input-Single-Output) and MIMO (Multiple-Input-Multiple-Output). One can overlap the proposed FLC on the FPGA hardware in order to verify the simplicity and accuracy of proposed FLC. Other applications like robotics, pattern recognition, medical expert system or industrial process control etc. The change will be in the knowledge base of FLC. The architecture we have proposed is so simple, quick and less complex that the FLC chip will definitely be cost-effective.

References

1. Gilberto, C.D. Sousa, Bimal K.Bose, "A Fuzzy Set Theory Based Control of a Phase-Controlled Converter DC   Machine Drive" IEEE Trans. On Industry Application, Vol.30.No.1, Jan/Feb 1994
2. C.C. Lee, "Fuzzy logic in control system: Fuzzy logic controller, PartI", IEEE Trans. On Syst. Man and Cybernatics, vol.20, pp. Mar/April 1990.
3. C.C. Lee, "Fuzzy logic in control system: Fuzzy logic controller, Part II", IEEE Trans.On Syst. Man and Cybernatics, vol.20, pp. Mar/April 1990.
4. Dajjin Kim, In-Hyun Cho, "An Accurate and Cost-Effective Fuzzy Logic Controller With a Fast Searching of Moment Equilibrium Point" IEEE Trans. On Industrial Electronics, Vol.46, No.2, April pp. 452.
5. Shen Li, "Fuzzy Logic Control ASIC Chip "J. of Computer Science & Technology.
6. Meng-Hiot Lim, Yoshiyasu Takefuji, "Implementing Fuzzy Rule-Based Systems on Silicon Chips." IEEE Expert, February 1990, pp.31.
7. A. G. Keskar, "A Fuzzy Controller for extending linearity of PWM waveform," IECON, International conference on IEEE Industrial Electronics, Taiwan, 1996.
8. A.G.Keskar, "Floating membership fuzzy logic for adaptive control of AC Drive", ISIE, International symposium on Industrial Electronics, Portugal, 1997.
9. A.G.Keskar, "Modified fuzzy inference scheme for a position control system", ISIC, International symposium on integrated circuits, Singapore, 1999.
10. Mapes, Bimal K.Bose, "Linearisation of the Transfer Characteristics of a Phase-Controlled Under Discontinuous Conduction" IEEE Trans. On Industry Applications, VOL.IA-14, No.6, NOV/DEC 1978.

Control System for Optimal Flight Trajectories
for Terrain Collision Avoidance

Tapan Sharma[1], Cees Bil[1], and Andrew Eberhard[2]

[1] Sir Lawrence Wackett Centre for Aerospace Design Technology
Royal Melbourne Institute of Technology
GPO Box 2476V, Melbourne, Victoria, 3001, Australia
[2] School of Mathematics and Geospatial Sciences, Royal Melbourne Institute of Technology
GPO Box 2476V, Melbourne, Victoria, 3001, Australia
s3075886@student.rmit.edu.au

Abstract. The aim of this project is to research optimal escape trajectories for aircraft in close ground proximity using a given digital terrain model. Optimal escape trajectories will depend on pilot skills, aircraft performance and as well as other requirements, such as the need to fly low or fast. This paper presents initial results of a project to optimise for a McDonald F-4 Phantom flight trajectories in a hypothetical, digital terrain using a non-linear optimisation method. The 2 outcomes show that different controls result depending on the objectives or constraints.

Introduction

Systems currently available for terrain awareness are Ground Proximity Warning System (GPWS) [11] and Terrain Awareness Warning System (TAWS). Unfortunately they do not add to Situation Awareness (SA) and do not provide guidance to avoid a collision. Previous work has been done to generate trajectories to avoid collision such as using mixed integer linear programming [6] and the Ground and Obstacle Collision Avoidance Technique (GOCAT) [8]. The earliest methods for finding optimal trajectories were based on calculus of variation. Discretisation of the optimality conditions and solution using numerical methods made it possible to solve complicated problems [1]. Hargraves and Paris implemented a method by discretising Differential Algebraic Equations (DAE), which was done by stating optimality conditions, and applying numerical methods [5]. The most effective method of optimisation is discretisation DAEs using linear and nonlinear optimisation. Most fighter aircraft are equipped with an onboard digital terrain model. This paper presents the results of a study aimed at generating optimal escape flight trajectories within a digital terrain model and the aircraft controls necessary to fly that trajectory [7]. The study investigated two objectives: crossing the terrain in minimum time and crossing the terrain with minimum ground clearance. A MATLAB-driven software package called DIRECT [9], which utilises the Sequential Quadratic Programming (SQP) software SNOPT [3], was used to obtain solutions in this work.. The control variables considered are angle of attack, bank angle and thrust setting. The results show that escape flight trajectories can be generated within a 3D terrain model and that these trajectories are flyable based on a simple aircraft model.

Performance Model

The equations of motion below are based on the assumption of a point mass aircraft [2]. The full point mass model equations of motion are described in detail in [10]. Sideslip and unsteady aerodynamic effects were neglected at this stage of the project. The Earth is assumed to be non-rotating, and the variation of attainable thrust with altitude has been ignored because of the relatively small changes in altitude that are generated by the controller.

Trajectory Optimisation

The equations of motion from [10] can be written in brief form as:

$$\dot{x} = f(x, u) \tag{1}$$

There are also other requirements, such as constraints on the load factors n_z generated by the longitudinal and lateral motion. They are described in more detail in [10]. Another constraint applied was that, at no time, the trajectory generated should be inside the terrain. This amounts to enforcing nonlinear path constraints as a function of the terrain height and aircraft height. This is easily accomplished within the framework presented in the DIRECT software [9]. The solution method used by DIRECT is direct transcription. In this approach, the nonlinear differential equations, path constraints, and cost function, which are provided by the user, are discretised into a Nonlinear Programming (NLP) Problem. This is solved using the solver SNOPT, which implements a sparse NLP algorithm. For minimum time problems, the final time becomes a free variable. The accuracy of the solution depends on the level of discretisation selected by the user. The strategy used in this work is to begin with a relatively low discretisation level so that a feasible trajectory (satisfying the path constraints generated by the differential equations and path constraints) is generated. However, this does not ensure that the aircraft does not collide with terrain between the discrete time points. Hence, this solution is used as initial guess for a more refined solution with a higher level of discretisation.

Cost Function

In this study, minimum time, where the aircraft has to clear the terrain in the shortest possible time possible and minimum distance to the terrain (terrain following), where the aircraft has to clear the terrain by maintaining a distance to the terrain of 50 ft or less was used. The Mayer and Bolza components of the cost function, shown in [10], were used in the performance index.

Discretisation Method and Terrain Modeling

For this project, the Hermite-Simpson discretisation [6] option in DIRECT proved to be the most effective method. In this approach, the state variables **x** are approximated by piecewise cubic polynomials as a function of non-dimensional time. The control variables **u** are approximated using piecewise linear functions as a function of non-

dimensional time. More information is given in [10]. There are two different types of terrain modeling. One is constructed by using a curve fitting method, and the other is using parametric equations for cones [10]. All the simulations were run on a Pentium 4 computer running MATLAB® 6.5 using DIRECT 1.13 [9].

Discussion of Results

All the results shown have been subjected to constraints on load factors in both lateral and transverse movement. Table 1 (where?) shows additional constraints such as rate of change put upon angle of attack and bank angle. The plots shown in Fig. 1 are the plots for the 2D scenario for Terrain Following (TF). The thrust and angle of attack show very less deflection in Fig. 2 and they were rather smooth controls. There was a variation in the throttle because the aircraft needs thrust initially to clear the terrain. After clearing the terrain, the thrust is reduced gradually. Hence, the required thrust is easily within achievable limits. The plot shown in Fig. 3 is for clearing the terrain in the least possible time. Looking closely at Fig. 3, we can see that the aircraft uses momentum to increase its velocity to clear the terrain. The throttle is set to 100% so that the aircraft can accelerate as quickly as possible. The plot shown in Fig. 4 exhibits that the optimal solution for this case is for the aircraft to maneuver around the terrain, rather than attempting to pull up and fly over it.
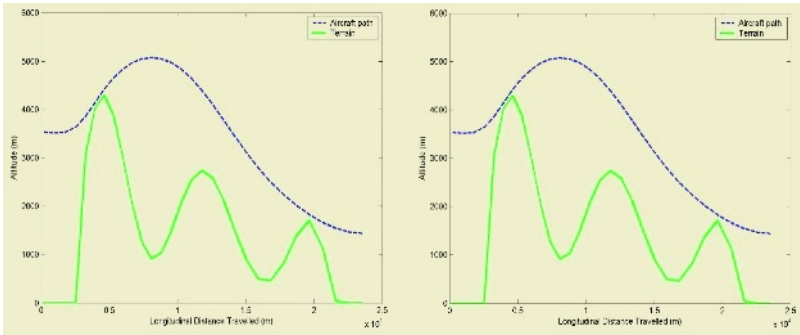
Fig. 1. Optimal Flight Path with Terrain Following pull-up and Minimum Time (2 – Dimensional)
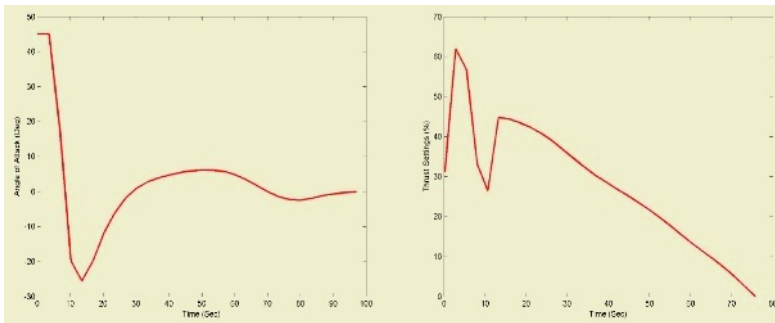
Fig. 2. Thrust Settings and Angle of Attack for Terrain Following
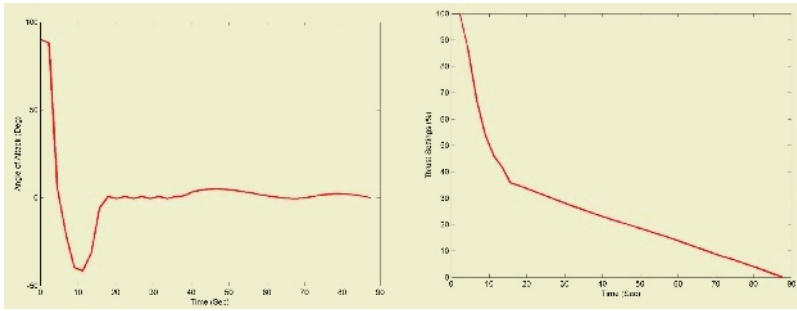
Fig. 3. Thrust Settings and Angle of Attack for Minimum Time

Conclusion

The objective of this paper was to show that generation of trajectories around a 3 dimensional obstacle is possible using numerical methods. The pull up maneuver has traditionally been sought as a viable solution. However, in more severe scenarios, it is necessary to incorporate lateral aircraft control. This result may invoke a new era for aircraft collision avoidance.

The results in this paper are based on a simple aircraft model but the point remains that the trajectories generated are flyable. In reality, all these simulations have to be run in real-time. The next step would be to use a digital terrain model consisting of wire-mesh and re-run these experiments using real-time scenario. Another improvement could also be to use actual aircraft controls such as elevators, ailerons and the thrust settings and treating the aircraft as a rigid body.
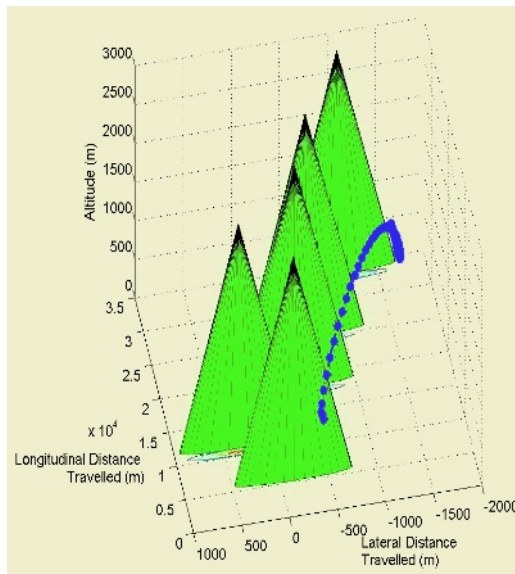
Fig. 4. Optimal Flight Path with Terrain following (3 – dimensional)

The plots shown in below in Fig. 5 are the control variables for the terrain avoidance maneuver shown in Fig. 4. We can tell that that aircraft was able to complete the go-round scenario because the control variables were rather smooth, therefore if need be, this trajectory can be flown by a pilot. The angle of attack shown below shows that the aircraft had climbed a certain height and then climbed down to gain momentum and continued that path until it reached its destination. The thrust settings reduced to zero gradually. The angle of attack and the bank angle rate were also not very abrupt. These all evidence lead to the fact that the trajectory generated was flyable.
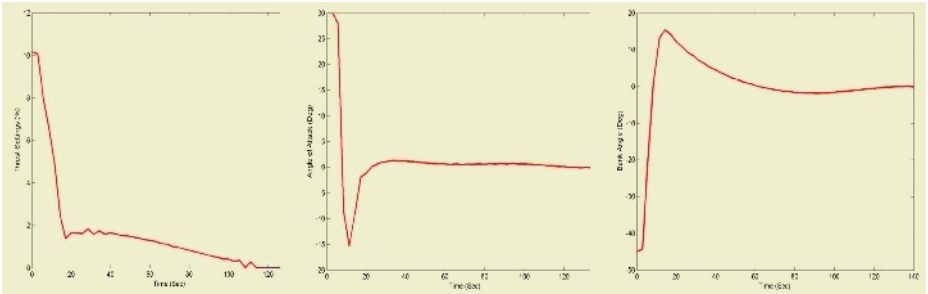
Fig. 5. Thrust Settings, Angle of Attack and Bank Angle

The plot in Fig. 6 shows the 3-Dimensional view for the Terrain Following scenario. We can see that the trajectory produced is within proximity to the terrain. If there were no limit on the load factor for longitudinal and lateral movement, the optimiser would have generated a trajectory that would have tried to get close to each and every cone. Additional constraints such as limiting the rate of change in the control variable make this plot interesting because, the control results shown in Fig. 7 proves that it this trajectory is flyable in real life case.
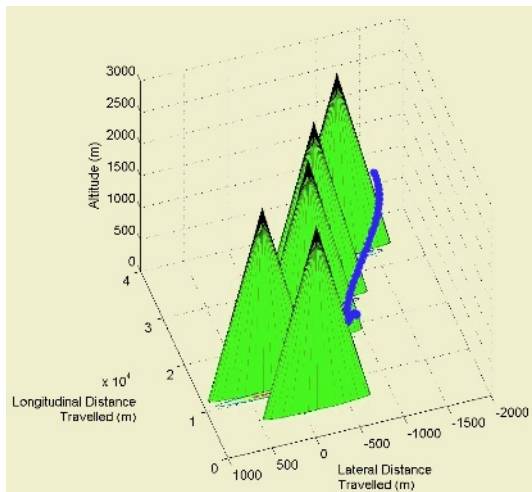
Fig. 6. Optimal Flight Path with Minimum Time (3 – dimensional)
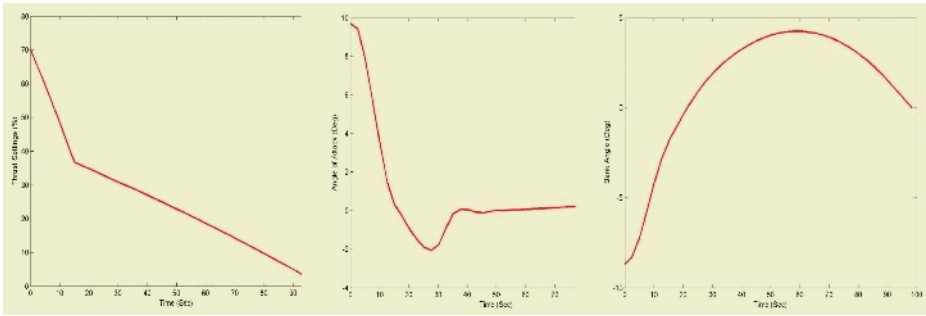
Fig. 7. Thrust Settings, Angle of Attack and Bank Angle

References

1. Hargraves, C. R. and Paris, S. W. "Direct Trajectory Optimization Using Nonlinear Programming and Colloration", Journal of Guidance, Control and Dynamics, Vol. 10, No. 4, 1987, pp. 338 – 342.
2. Bryson, A. E., Jr. and Desai, M. N., "Energy-State Approximation in Performance Optimization of Supersonic Aircraft." Journal of Aircraft, Vol. 6, No. 6, 1969, pp. 481 – 488
3. Gill, P.E., Murray, W., and Saunders, M. A., "Users Guide for SNOPT 6.0: A FORTRAN Package for Large-Scale Nonlinear Programming." Dept of Mathematical. Rept. NA 97-X, Univ. of California, San Diego, CA, 1997.
4. Menon, P. K., and Cheung, V. H. L.,"Minimum-Exposure Near Terrain Flight Trajectories for Rotorcraft," Control and Dynamic Systems, Edited by C.T. Leondes, Vol. 52, Academic Press, San Diego 1992, pp. 391 – 433.
5. Brenan, K. E., "Differential-Algebraic Equations Issues in the Direct Transcript of Path Constraints Optimal Control Problem," The Aerospace Corp., Aerospace Rept. ATR-94 (8489)-1, 1993
6. Arthurs R. Jonathan P. H. "Aircraft Trajectory Planning with Collision Avoidance Using Mixed Linear Programming." Proceedings of the American Control Conference Anchorage. AK May 8 – 10, 2002
7. Bors, A. G., "Terrain Modeling in Synthetic Aperture Radar Images using Shape from Shading." Dept of Computer Science, University of York, York YO10 5DD, UK, 2000
8. Hewitt, C., Hickey, A. J. and Boyes, J. D. "A Ground and Obstacle Collision Avoidance Technique (GOCAT)." IEES AES Systems Magazine, Dayton, OH, May 1991
9. William P., "Users Guide to Direct Version 1.13." School of Aerospace, Mechanical and Manufacturing Engineering, 2004
10. Sharma T., Bil C. and Eberhard A. "Optimal Flight Trajectories for Terrain Collision Avoidance" AIAC, Eleventh Australian International Aerospace Congress Melbourne Australia, March 2005
11. http://www.boeing-727.com/Data/systems/infogpws.html

Optimal Remediation Design in Groundwater Systems by Intelligent Techniques

Hone-Jay Chu[1,*], Chin-Tsai Hsiao[2,**], and Liang-Cheng Chang[1,***]

[1] Department of civil Engineering National Chiao-Tung University
Hsinchu Taiwan, 300, R.O.C.
[2] Chungchou Institute of Technology,Yuanlin, Changhwa, Taiwan, 510, R.O.C.

Abstract. This research develops an optimal planning model for pump-treat-inject based groundwater remediation systems. Optimizing the design of the pump-treat-inject system is a nonlinear, dynamic and discrete optimization problem. This study integrates the Genetic Algorithm (GA) and Differential Dynamic Programming (DDP) to solve this highly complex problem. The proposed model considers both the cost of installing wells (fixed cost) and the operating cost of pumping, injection and water treatment. Minimizing the total cost and meeting the water quality constraints, the model computes the optimal number and location of wells, as well as the associated optimal pumping and injection schemes. This work also investigates many factors that affect the optimal design of a remediation system, such as, various numerical cases revealing the time-varying pumping and injection rate, and the requirement to balance the total volume between pumping and injection that can significantly influence the optimal design.

1 Introduction

Groundwater is a valuable natural resource, and is threatened by contaminant because industrial and waste disposal activities are more serious in recent years. Groundwater remediation is associated with enormous costs and time. The pump-and-treat is one of the most commonly applied methods of groundwater remediation. By pumping out contaminated groundwater, treating the water, the method is primary useful for decontaminating groundwater with highly soluble pollutants.

Effective design of a remediation system in groundwater requires consideration of more than only the effectiveness of the technological process involved. Decision makers also want to minimize costs or cleanup time or maximum contaminant mass removal. In recent years, optimization planning models have been developed to design groundwater remediation strategies. Chang (1992) applied constrained differential dynamic programming (CDDP) to the optimal design of time-varying pumping rates for remediation. McKinney and Lin[1995] use mixed-integer programming to optimal design the air-stripping treatment process. The objective is to minimize the total cost including fixed and operating costs by the pumping and injection rates at five potential wells. The pumping and injection rates are not time-varying. Bear [1998]used the two-level hierarchical optimization model. At the basic level, well

---

locations and pumping/injection rates are sought so as to maximize mass removal of contaminants. At the upper level, the number of wells for pumping /injection is optimized, so as to minimize the cost, taking maximum contaminant level as a constraint. The model neglects operating cost. Guan and Aral [1999] used progressive genetic algorithm to optimize the remediation design, while defining the locations and pumping and injection rates of the specified wells as continuous variables. The proposed model considers only the operating cost of pumping and injection. But the pumping and injection rates are not time-varying. To considering the fixed and operating cost, Chang and Hsiao [2002] integrate Genetic Algorithm (GA) and Constrained Differential Dynamic Programming (CDDP) to optimize total remediation cost. But the decision variables only involve determining time-varying pumping rates from extraction wells and their locations.

This research develops an optimal planning model for pump-treat-inject based groundwater remediation systems. This study integrates the Genetic Algorithm (GA) and Differential Dynamic Programming (DDP) to solve this highly complex problem. The proposed model considers both the cost of installing wells (fixed cost) and the operating cost of dynamic pumping, injection and water treatment. Minimizing the total cost to meet the water quality constraints, the model computes the optimal number and location of wells, as well as the associated optimal dynamic pumping and injection rate.

2   Formulation of the Planning Model

The formulation is to minimize the total cost including both fixed and operating cost of system while determining the locations, numbers of the extraction or injection well and pumping/injection rate. The first component is the costs of pumping field, involving capital costs and extraction and treatment costs. The second component in is the cost of injection field, involving capital costs and injection costs.

Subject to: The transport model includes changes in head due to pumping or injection, and changes in the contaminant concentration owing to advection, diffusion, dispersion, and linear equilibrium sorption. The other constraints ensures the water-quality standard will be met at the specified monitoring wells by the end of the period, the capacity constraints for total pumping rate, the capacity constraints for each pumping well, injection well.

3   The Algorithm of GCDDP: Integration of a GA and CDDP

This investigation integrates GA and CDDP (GCDDP) to solve the problem. The total cost objective function is mixed-integer nonlinear time-varying as well as including discontinues variables (pumping/ injection well locations) and continues variables (time-varying pumping/injection rates). The operating (pumping, injection, and treatment) costs are continuous functions and are separable functions for each period.In this integrated approach, GA, a near global optimization algorithm, is used to locate the optimal well sites, while CDDP is employed to calculate the optimal pumping rates. According to this figure, the algorithm is a GA with CDDP embedded to compute the optimal operation costs for a potential network alternative (represented by a

chromosome). The total cost for each network alternative (chromosome) is the sum of the optimal operation costs and its fixed costs.

The concept of GCCDP as an optimal model, searches for optimal designs by comparing the fitness. Figure 1 illustrates the procedure of the algorithm which including parameter encoding, the fitness calculation of the chromosome using CDDP, and manipulate these possible solution through reproduction ,crossover, and mutation. The chromosome of designs is allowed to evolve through successive generation until a termination criterion is met.
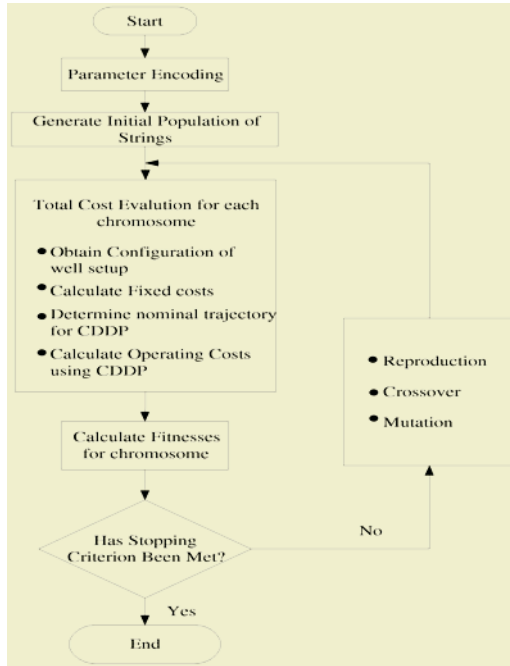
Fig. 1. The flowchart of GCDDP groundwater remediation model

The genetic operations of reproduction, crossover and mutation on population of the potential designs are essence of GAs. The reproduction operator reproducts individuals that are copied into the next generation according to fitness. The simple way to implement the solution operator is by means of the tournament selection. The crossover operation takes two individuals ,randomly selects a position that creates two segments of genetic material in each parent ,and afterwards interchanges those segments if genetic material ,thus generating two new individual. The mutation operator consists in the alternation of genetic code in random way by taking an individual and changing the value at each of its positions with another.

4   Design Example

A hypothetical, homogeneous, isotropic confined aquifer with dimensions of 600m by 1200m serves as an example. Figure 2 presents the finite-element mesh, associated

with no flow boundary conditions on north and south of aquifer, and constant-head boundaries on the west and east which are 22m and 10m.There 91 finite-element nodes, along with 24 candidate well sites, and 17 observation wells. The hydraulic head distribution prior to pumping is assumed to be steady, the initial peak concentration within the aquifer is 150 mg/L, and the water quality goal at the end of five years must be less than or equal to 0.5 mg/L at all the observation wells. The time between each period in the management model(Δt) is 91.25 days. Aquifer properties are listed in Table 1. The value of cost-related coefficients, a_1 =120 ($\$/m$), a_2 =40000($m^3/s \cdot m \cdot \Delta t$), a_3 =1000($m^3/s \cdot \Delta t$),**and** a_4 =1(case2), 2000(case3),5000(case4)($m^3/s \cdot \Delta t$)is listed.
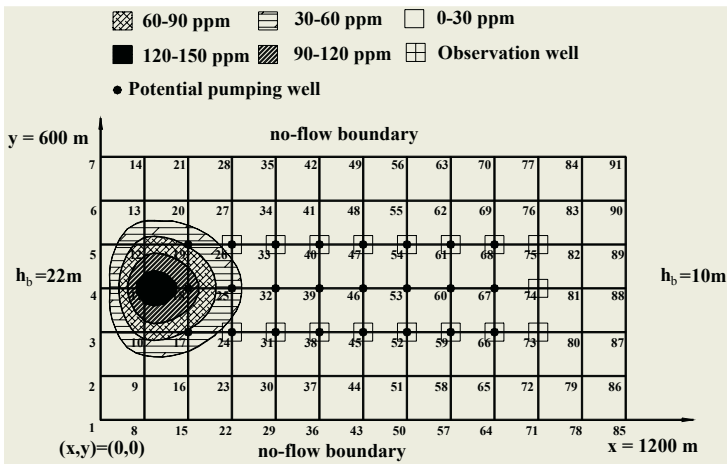
Fig. 2. Finite element mesh, Boundary Conditions, Initial Plume, and Locations of Numbered Observation and Potential Wells for all Runs of the Groundwater Reclamation Example

Table 1 lists the computational conditions under which optimization computation was implemented. In case 1, the decision variables are pumping rate and well locations. The decision variable in else cases are pumping and injection well pumping rate and their locations. In case 4, the alternative is to design a system considering rates equilibrium. Treated water was injected into the field. It doesn't need a large body of water import to the system, so the injection cost coefficient is low.

Table 1. The case list

Case	Strategy	Injection cost coefficients	Rates equilibrium
1	Pump	…	Not Consider
2	Pump&inject	5000($m^3/s \cdot \Delta t$)	Not Consider
3	Pump&inject	2000($m^3/s \cdot \Delta t$)	Not Consider
4	Pump&inject	1($m^3/s \cdot \Delta t$)	Consider

5   Results and Discussion

The results of all scenarios are summarized in Table 2. The fixed cost, operating cost, and total cost are listed from comparison. The decision variables are pumping rate and well locations while in else cases are pumping and injection well, pumping rate and their locations. Optimized total cost in case1is $63,557.Optimal solution in case 2 all type cost is same as that in case 1, the optimal strategy tends to pumping. The optimized total cost of case 3 is smallest, optimized total cost is $40,069. Because injection cost is cheaper than extraction, the optimal strategy tends to injection. It implies that the injection cost coefficient is dominating the solution because the model achieves the minimum total cost. The resulting total cost of case 4 is equal to $119,166. The system in case 4 keeps going on pumping out contaminated groundwater, treating the water and then re-injecting treated water. Figure 3 show the pumping rate at each period. Figure 4 depicts the optimal network design and pollutant concentration distribution for the finial planning period. All the concentration distributions meet water quality standard at the end of planning period at all observation.

Table 2. The result of all cases

Case	Number of pumping well	Number of injection well	Fixed cost	Operating cost	Total cost
1	2	0	2,880	60,677	63,557
2	2	0	2,880	60,677	63,557
3	0	5	7,200	32,869	40,069
4	5	5	14,400	104,766	119,166
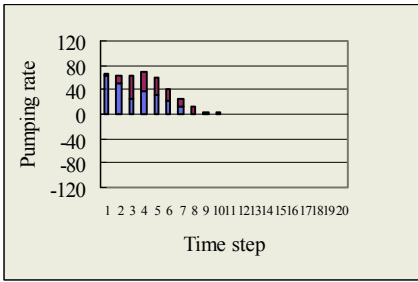
(a) case 1
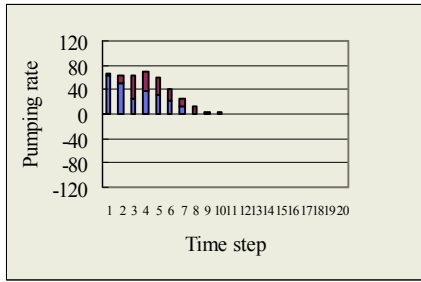
(b) case 2
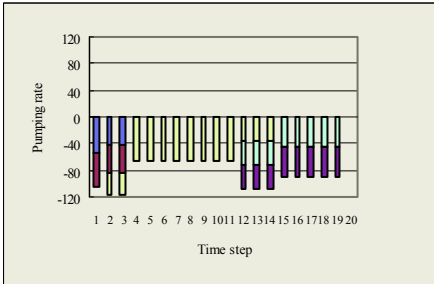
(c) case 3
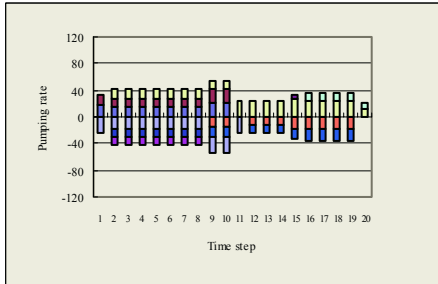
(d) case 4

Fig. 3. (a)The pumping rate at each period in case1(b)The pumping rate at each period in case2 (c)The pumping rate at each period in case3(d)The pumping rate at each period in case4

(a) case 1

(b) case 2

(c) case 3

(d) case 4
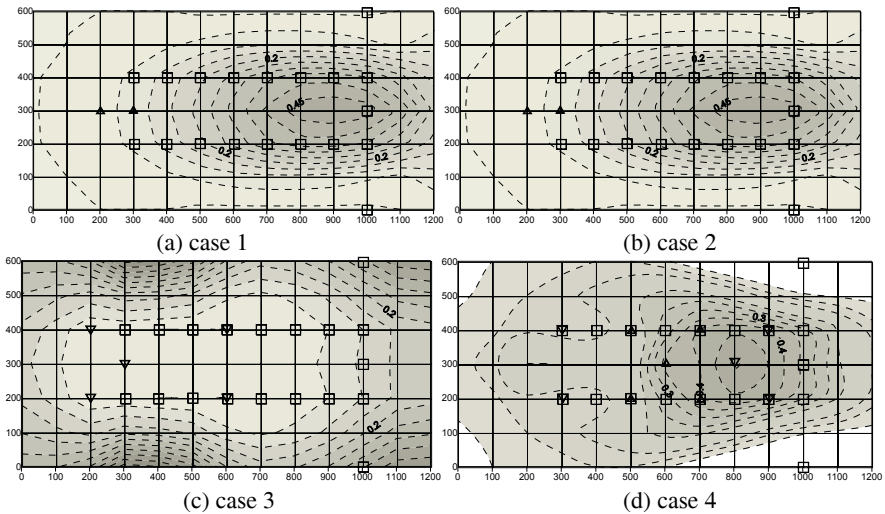
Fig. 4. The optimal number of well and concentration distribution at the final period in case1,2,3,4 (△:pumping well, ▽:injection well)

6 Conclusion

The optimization model was proposed for the optimal design of pump-treat-inject systems used to clean up the contaminated sites. The concept of PTI is pumping removal for contaminant, water injection is for dilution. In the paper, we integrated CDDP and GA to calculate the minimum total cost while simultaneously considering the fixed costs and time-varying operating costs. The methodology is applied to find the cost–effectiveness remediation strategies including well locations and pumping/injection rate.

Through a design process, the effects of injection, pumping and injection rates equilibrium, hydraulic conductivity, and water level performance were analyzed.

Numerical results performed in this paper demonstrate the significance is the parameter such as the injection coefficient affects the total remediation cost. If injection cost coefficient is higher, the optimal strategy prefers pumping. If injection cost coefficient is higher, the optimal strategy tends to injection. But if the large volume of water inject, the plume will migrate out of boundary at the final period. We suggest to be aware of the plume is not controlled.

To control the movement of contaminated ground water efficiently, the objective function is suggested to be taken into consideration of the removal efficiency. To consider more realistic and complex problem, the computational time required for the solution will increase. We suggest the computational time to be under efficiency solved by massive parallel computers.

References

1. Bear, J., and Y.W. Sun, Optimization of pump-treat-inject (PTI) design for the remediation of a contaminated aquifer: multi-stage design with chance constraints. *Journal of Contaminant Hydrology*, 29(3), 225-244,(1998)

2. Chang, L. C., and C. T. Hsiao, Dynamic optimal Groundwater Remediation including fixed and operation costs. *Ground Water*, Vol. 40, No. 5, pp 481-490. (2002)

3. Chang, L.C., and C.A. Shoemaker, Optimal time-varying pumping for groundwater remediation application of a constrained optimal control algorithm. *Water Resources Research*, 28(12), 3157-3173 (1992)

4. Culver, T.B.,and G.W. Shenk ,Dynamic optimal ground water remediation by granular activated carbon.*Journal of Water Resources Planning and Management-ASCE,* 124 (1): 59-64 ,(1998)

5. Guan, J.,and M.M. Aral ,Optimal remediation with well locations and pumping rates selected as continuous decision variables. Journal of Hydrology,221 (1): 20-42, (1999)

6. Hsiao, C. T. ,and L. C. Chang, Dynamic optimal groundwater management with inclusion of fixed costs**.** *Journal of Water Resources Planning and Management-ASCE* ,128 (1), 57-65 ,(2002).

7. McKinney, D. C., and M. D. Lin, Approximate mixed-integer nonlinear programming methods for optimal aquifer remediation design. *Water Resources Research*,31 (3): 731-740, (1995)

8. Pinder, G. F., Galerkin finite element models for aquifer simulation, *Rep. 78-WR-5*, Dept. of Civ. Eng., Princeton Univ., Princeton, N. J., (1978).

9. Wang, J.C., J.R. Booker, and J.P. Carter, Analysis of the remediation of a contaminated aquifer by a multi-well system. *Computers and Geotechnics*, 25(3), 171-189 ,(1999).

Choquet Integral-Based Decision Making Approach for Robot Selection

E. Ertugrul Karsak

Industrial Engineering Department, Galatasaray University, Ortakoy, Istanbul 80840, Turkey
ekarsak@gsu.edu.tr

Abstract. This paper addresses the robot selection problem by using a multi-criteria decision making approach based on the Choquet integral. The key feature of the Choquet integral is that it can incorporate interaction among attributes into the decision analysis, an issue that has been generally overlooked in earlier robot selection studies. The proposed decision framework makes use of the information previously elicited from experts analyzing a reference set of robot alternatives with the same selection criteria as in the current evaluation, and applies the Choquet integral to determine the most suitable robot. A comprehensive illustrative example is presented using a previously published data set, and the results obtained using the proposed approach are compared with those of another multi-criteria decision aid, which has been also employed in earlier studies addressing robot selection.

1 Introduction

Manufacturers who have been faced with intense competition in the global marketplace have invested in advanced manufacturing technologies (AMTs), such as group technology, flexible manufacturing systems, industrial robots, etc., which enable high quality and customisation in a cost effective manner. The increased concern and importance attached to AMTs by the manufacturers have consequently oriented the researchers to develop models and methodologies for evaluation and selection of AMTs. Khouja and Offodile [8] presented a survey of models for industrial robot selection, and Raafat [14] provided a comprehensive bibliography on justification of AMTs. The investment decisions for advanced manufacturing systems frequently involve multiple and conflicting criteria. This paper deals with the problem of robot selection, although the decision model presented here is not limited to this specific problem and could very well be applied to technology selection, in general.

Robots can be programmed to keep a constant speed and a predetermined quality when performing a task repetitively. Robots can manage to work under conditions hazardous to human health such as excessive heat or noise, heavy load, toxic gases, etc. Therefore, manufacturers prefer to use robots in many industrial applications where repetitive, difficult or hazardous tasks need to be performed, such as assembly, machine loading, materials handling, spray painting, and welding. Parallel to the upward trend in the use of industrial robots, researchers have focused on the robot selection problem in the past two decades [1, 4, 5, 6, 7, 9, 12, 13]. The developed approaches are based on decision making techniques including economic analysis, scoring models, statistical analysis, analytic hierarchy process (AHP), technique for order preference by similarity to ideal solution (TOPSIS), mathematical programming, fuzzy set theory, and genetic algorithms.

In general, multiple criteria decision making techniques use the weighted arithmetic mean as the aggregation tool despite its limitations. Three decades ago, Sugeno [15] introduced the concept of fuzzy integral. Although the initial research was far from being application-oriented, fuzzy integrals have later been used in decision making. Researchers generally tend to ignore the frequently encountered problem of interacting criteria by assuming that criteria are independent. Contrary to widely-used multiple criteria decision aids, the fuzzy integral enables the decision makers to incorporate the interaction between criteria into the analysis. The Sugeno integral was initially used as an aggregation tool that considers the importance of criteria expressed by a fuzzy measure for calculating an average global score. While the Sugeno integral is not stable under positive linear transformations, the Choquet integral, which is an extension of classical Lebesgue integral, fulfils this property. Hence, in this sense, the Sugeno integral appears to suit better for ordinal aggregation, whereas the Choquet integral is more suitable for cardinal aggregation [2].

This paper presents a decision framework based on the Choquet integral for evaluating and selecting industrial robots. The main feature of the Choquet integral is that it can take into account interaction among robot attributes, which has been generally ignored in previous studies focusing on robot selection. We employ 2-order fuzzy measure, which is relatively simple with only quadratic complexity and enables the modelling of interaction. The problem of finding a 2-order fuzzy measure can be handled by solving a linear program.

Expert judgments are oftentimes employed in multi-criteria decision making. However, referring to experts each time a robot is to be acquired by a manufacturing firm would definitely result in additional costs, and thus, contradict with the objective of improving manufacturing performance in terms of cost in real-world practices. Therefore, a ranking of a reference set of robot alternatives and/or ranking of criteria obtained from experts when a robot investment was evaluated, is likely to be used in subsequent robot selection practices of similar type. The decision procedure proposed in here makes use of the information previously elicited from experts analyzing a different set of robot alternatives with the same selection criteria to be used in the current evaluation, and applies the Choquet integral to determine the best robot.

2   Multi-criteria Decision Making Using the Choquet Integral

In this section, the basic material on fuzzy measures and the Choquet integral is presented. Considering the space limitations, the definitions and mathematical treatment are restricted to the minimum. The interested reader can refer to Grabisch [2] for a more profound treatment of fuzzy measures and integrals.

Definition 1. A fuzzy measure on X is a monotonic set function $\mu : P(X) \mapsto [0,1]$ with $\mu(\varnothing) = 0$, $\mu(X) = 1$, and $A \subset B \subset X$ implies $\mu(A) \leq \mu(B)$.

Definition 2. Let μ be a fuzzy measure on X. The Choquet integral of a function $f : X \mapsto [0,1]$ with respect to μ is defined as [2]

$$C_\mu(f(x_1),\ldots,f(x_n)) := \sum_{i=1}^n \left(f(x_{(i)}) - f(x_{(i-1)})\right)\mu(A_{(i)}) \tag{1}$$

where $\bullet_{(i)}$ denotes that the indices have been permuted so that $0 \le f\big(x_{(1)}\big) \le \cdots \le f\big(x_{(n)}\big) \le 1$, $A_{(i)} := \big\{x_{(i)}, \ldots, x_{(n)}\big\}$, and $f\big(x_{(0)}\big) = 0$.

Definition 3. Let μ be a fuzzy measure on X. The importance index or Shapley value for every $i \in X$ is defined as

$$\Phi_i = \sum_{A \subset X - i} \frac{(n - |A| - 1)! \, |A|!}{n!} \big[\mu(A \cup \{i\}) - \mu(A)\big] \tag{2}$$

where $|A|$ denotes the cardinal of A, and the Shapley value of μ is the vector $\Phi(\mu) = [\Phi_1, \Phi_2, \ldots, \Phi_n]$. The Shapley values possess the property that $\sum_{i=1}^{n} \Phi_i = \mu(X)$.

A problem containing n criteria requires 2^n coefficients in [0,1] to define the fuzzy measure μ on every subset, which results in a cumbersome process of obtaining huge amount of information from an expert. The concept of k-order fuzzy measure, which has been proposed to rectify this difficulty, represents a k-order approximation of the polynomial expression of a fuzzy measure in the neighbourhood of the origin [10]. The k-order fuzzy measure can be represented by at most $\sum_{i=1}^{k} \binom{n}{i}$ coefficients. The 2-order fuzzy measure, which is relatively simple with only quadratic complexity, requires only $n + \binom{n}{2}$ coefficients to define the fuzzy measure for a problem involving n criteria. Throughout this paper, we assume that interaction between more than two criteria does not exist, and thus, the model is restricted to the 2-order case in a way to alleviate the computational burden. In this case, the Choquet integral can be given as

$$C_\mu(x) = \sum_{i \in N} a(i) x_i + \sum_{\{i,j\} \subseteq N} a(ij)\big(x_i \wedge x_j\big), \quad x \in \Re^n \tag{3}$$

The problem of finding a 2-order fuzzy measure can be handled by solving a linear program. Expressed in terms of the Möbius representation, the model employed in this paper for identifying weights of the selection criteria is as follows:

$$\max z = \varepsilon \tag{4}$$

subject to

$$C(k) - C(k') \ge \delta + \varepsilon, \; if \; k \succ_A k'$$
$$-\delta \le C(k) - C(k') \le \delta, if \; k \approx_A k'$$
$$\sum_{i \in N} a(i) + \sum_{\{i,j\} \subseteq N} a(ij) = 1$$
$$a(i) \ge 0, \quad \forall i \in N$$
$$a(i) + \sum_{j \in T} a(ij) \ge 0, \quad \forall i \in N, \forall T \subseteq N|i$$
$$C(k) = \sum_{i \in N} a(i) z_{ki} + \sum_{\{i,j\} \subseteq N} a(ij)\big(z_{ki} \wedge z_{kj}\big), \quad \forall k \in A$$

where A denotes the set of alternatives, N denotes the set of criteria, $\ge_A$ indicates preferences over the alternatives, z_{ki} indicates the normalized value of the kth alternative on the ith criteria, and δ is a fixed preference threshold.

3 Application of the Choquet Integral to Robot Selection Problem

In this section, we analyze a robot selection problem using the data for the 27 robot alternatives presented in Imany and Schlesinger [4] with cost, load capacity, velocity and repeatability as the attributes considered in determining the best robot alternative. We assume that a preference structure of a reference set of robot alternatives that was obtained from experts in an earlier case of similar type is used in this robot selection study. The decision procedure proposed in this paper makes use of the information previously elicited from experts analyzing a different set of robot alternatives with the same selection criteria to be used in the current evaluation, and applies the Choquet integral to determine the most suitable robot. Table 1 shows the data for the reference set of robot alternatives.

Table 1. Related attributes for a reference set of robot alternatives

Reference robot alternative (RR_k)	Cost ($10,000)	Load capacity (kg)	Velocity (m/s)	Repeatability (mm)
RR_1	4.80	60	1.50	0.20
RR_2	5.20	50	1.40	0.25
RR_3	4.80	55	1.50	0.25
RR_4	4.90	65	1.70	0.20
RR_5	5.40	55	1.70	0.25
RR_6	4.60	50	1.40	0.20

The normalized values for benefit-related attributes ($i \in B$), i.e. attributes for which maximum possible value is desired (e.g., load capacity and velocity), and cost-related attributes ($i \in C$), i.e. attributes for which minimum possible value is more favourable (in here, cost and repeatability), are calculated using a linear normalization scheme as

$$z_{ki} = \begin{cases} \dfrac{x_{ki}}{x_i^*}, & i \in B \\[2mm] \dfrac{x_i^-}{x_{ki}}, & i \in C \end{cases} \tag{5}$$

where B and C are the set of benefit criteria and cost criteria, respectively, $x_i^* = \max_k x_{ki}$ and $x_i^- = \min_k x_{ki}$, and $0 \le z_{ki} \le 1$. The experts' preferences over the reference set of robot alternatives are expressed by a binary relation $\succ_A$ as follows:

$RR_4 \succ_A RR_1$, $RR_1 \succ_A RR_6$, $RR_6 \succ_A RR_3$, $RR_3 \succ_A RR_5$, $RR_5 \succ_A RR_2$.

By solving the linear program (4) for $\delta = 0.01$, we obtain $a(1) = 0.4274$, $a(2) = 0.1964$, $a(3) = 0.0967$, $a(4) = 0.7042$, $a(12) = 0$, $a(13) = 0$, $a(14) = -0.4274$, $a(23) = 0.2794$, $a(24) = -0.1800$, $a(34) = -0.0967$. Then, Shapley values corresponding to cost, load capacity, velocity and repeatability are calculated as 0.2137, 0.2461, 0.1881 and 0.3521, respectively. Next, the robot data presented in [4] is normalized employing (5), and then the Choquet integral ratings are calculated as given in Table 2 using the previously computed values of $a(i)$ and $a(ij)$, for $i \in N$ and $j \in N \backslash i$.

Table 2. Related robot data [4], and ratings and ranks for 27 robots (R_i) using the Choquet integral and TOPSIS

R_i	Cost ($10000)	Load Capacity (kg)	Velocity (m/s)	Repeat-ability (mm)	Choquet integral Rating	Rank	TOPSIS Rating	Rank
R_1	7.20	60.0	1.35	0.150	0.2758	10	0.2723	11
R_2	4.80	6.0	1.10	0.050	0.3656	6	0.3480	5
R_3	5.00	45.0	1.27	1.270	0.1796	17	0.2189	17
R_4	7.20	1.5	0.66	0.025	0.7064	1	0.5051	1
R_5	9.60	50.0	0.05	0.250	0.1073	26	0.1261	26
R_6	1.07	1.0	0.30	0.100	0.1775	18	0.1741	24
R_7	1.76	5.0	1.00	0.100	0.2075	14	0.2285	16
R_8	3.20	15.0	1.00	0.100	0.2219	13	0.2299	14
R_9	6.72	10.0	1.10	0.200	0.1436	21	0.1929	18
R_{10}	2.40	6.0	1.00	0.050	0.3608	7	0.3441	6
R_{11}	2.88	30.0	0.90	0.500	0.1369	25	0.1629	25
R_{12}	6.90	13.6	0.15	1.000	0.0495	27	0.0361	27
R_{13}	3.20	10.0	1.20	0.050	0.3762	4	0.3582	4
R_{14}	4.00	30.0	1.20	0.050	0.4051	3	0.3701	3
R_{15}	3.68	47.0	1.00	1.000	0.1760	19	0.1889	19
R_{16}	6.88	80.0	1.00	1.000	0.2447	11	0.2291	15
R_{17}	8.00	15.0	2.00	2.000	0.1401	22	0.2834	10
R_{18}	6.30	10.0	1.00	0.200	0.1387	23	0.1796	23
R_{19}	0.94	10.0	0.30	0.050	0.3665	5	0.3186	9
R_{20}	0.16	1.5	0.80	2.000	0.4705	2	0.3352	8
R_{21}	2.81	27.0	1.70	2.000	0.1692	20	0.2579	12
R_{22}	3.80	0.9	1.00	0.050	0.3534	9	0.3396	7
R_{23}	1.25	2.5	0.50	0.100	0.1797	15	0.1843	20
R_{24}	1.37	2.5	0.50	0.100	0.1797	16	0.1830	21
R_{25}	3.63	10.0	1.00	0.200	0.1387	24	0.1804	22
R_{26}	5.30	70.0	1.25	1.270	0.2358	12	0.2416	13
R_{27}	4.00	205.0	0.75	2.030	0.3546	8	0.3764	2

In order to demonstrate the aptness of using the Choquet integral as a decision aid, the results obtained using the Choquet integral are compared with those obtained by applying TOPSIS, a well-known decision tool. Shapley values, that are calculated using the outcomes of the linear program to determine a 2-order fuzzy measure, serve as the relative importance weights of the criteria in TOPSIS.

TOPSIS is based on the intuitive principle that the preferred alternative should have the shortest distance from the ideal solution and the farthest distance from the anti-ideal solution [3]. The traditional TOPSIS approach uses the Euclidean norm to normalize the original attribute values, and the Euclidean distance to calculate each alternative's distance from the ideal and anti-ideal solutions. A drawback of applying

the Euclidean norm is that the minimum and the maximum values of the normalized scale are not equal for each criterion, which results in difficulty in making inter-criterion comparisons. In order to rectify this difficulty and remain consistent with the normalization scheme used in the Choquet integral, we employ a linear scale transformation for normalization enabling the scale of measurement to vary precisely in the [0, 1] interval for each criterion. TOPSIS has been previously used as a decision aid in robot selection [1, 13]. One shall note that different data conversion factors, criterion weights or distance measures utilized in TOPSIS produce different results.

The results of the comparative analysis indicate that both the Choquet integral and TOPSIS determine robot 4 (R_4) as the best robot alternative. In order to analyze the ranking differences of the robot alternatives obtained by the Choquet integral versus TOPSIS, we make use of Spearman's rank correlation [11]. Using the ranks of the robot alternatives presented in Table 2, Spearman's rank correlation coefficient, r_s, is calculated as 0.871. The null hypothesis (H_0) and the alternative hypothesis (H_a) are constructed as given below to determine whether a significant positive relationship exists between the sets of rankings obtained by the Choquet integral and TOPSIS.

H_0: There is no positive correlation between the ranks of the robot alternatives.

H_a: There is a positive correlation between the ranks of the robot alternatives.

H_0 is rejected if $r_s > r_{s,\alpha}$. For $n = 27$ and the significance level of 0.05 ($\alpha = 0.05$), $r_s = 0.871 > 0.323 = r_{s,0.05}$, and thus, we reject H_0 and conclude that the Choquet integral rankings are positively correlated with the TOPSIS rankings.

4   Concluding Remarks

Real-world applications generally include interacting criteria that cannot be properly evaluated using the weighted arithmetic mean as the aggregation tool. Nevertheless, it is possible to consider interacting criteria by employing the Choquet integral in place of the weighted arithmetic mean, and thus, avoid the unrealistic independence assumption. On the basis of the knowledge of preference over a reference set of robot alternatives using cost and performance parameters such as load capacity, velocity and repeatability as decision criteria, this paper employs the Choquet integral to determine robot alternatives with the best combination of criteria values among the robots listed in a previously published data set. The 2-order model is used that enables to model interaction between criteria while preserving simplicity. The results obtained using the proposed approach are compared with those obtained by TOPSIS. The high rank correlation between the rankings obtained by the Choquet integral-based decision framework and TOPSIS is supportive of using the Choquet integral as a robust and convenient multi-criteria decision tool, which also enables the use of information elicited from experts, in robot selection.

References

1. Agrawal, V.P., Kohli, V., Gupta, S.: Computer aided robot selection: 'multiple attribute decision making' approach. Int. Journal of Production Research 29 (1991) 1629-1644
2. Grabisch, M.: The application of fuzzy integrals in multicriteria decision making. European Journal of Operational Research 89 (1996) 445-456

3. Hwang, C.-L., Yoon, K.: Multiple Attribute Decision Making: Methods and Applications. Springer, Heidelberg (1981)
4. Imany, M.M., Schlesinger, R.J.: Decision models for robot selection: a comparison of ordinary least squares and linear goal programming methods. Decision Sciences 20 (1989) 40-53
5. Jones, M.S., Malmborg, C.J., Agee, M.H.: Decision support system used for robot selection. Industrial Engineering 17 (1985) 66-73
6. Karsak, E.E.: A two-phase robot selection procedure. Production Planning & Control 9 (1998) 675-684
7. Karsak, E.E., Ahiska, S.S.: Practical common weight multi-criteria decision-making approach with an improved discriminating power for technology selection. International Journal of Production Research 43 (2005) 1537-1554
8. Khouja, M., Offodile, O.F.: The industrial robots selection problem: literature review and directions for future research. IIE Transactions 26 (1994) 50-61.
9. Liang, G.S., Wang, M.J.J.: A fuzzy multi-criteria decision-making approach for robot selection. Robotics and Computer-Integrated Manufacturing 10 (1993) 267-274
10. Marichal, J.L., Roubens, M.: Determination of weights of interacting criteria from a reference set. European Journal of Operational Research 124 (2000) 641-650
11. McClave, J.T., Benson, P.G.: Statistics for Business and Economics. San Francisco: Dellen, (1988)
12. Offodile, O.F., Lambert, B.K., Dudek, R.A.: Development of a computer aided robot selection procedure (CARSP). International Journal of Production Research 25 (1987) 1109-1121
13. Parkan, C., Wu, M.L.: Decision-making and performance measurement models with applications to robot selection. Computers & Industrial Engineering 36 (1999) 503-523
14. Raafat, F.: A comprehensive bibliography on justification of advanced manufacturing systems. International Journal of Production Economics 79 (2002) 197-208
15. Sugeno, M.: Theory of Fuzzy Integrals and Its Applications. Ph.D. Thesis, Tokyo Institute of Technology, (1974)

Fuzzy Logic and Neuro-fuzzy Modelling
of Diesel Spray Penetration

Shaun H. Lee, Bob R.J. Howlett, Simon D. Walters, and Cyril Crua

Intelligent Systems & Signal Processing Laboratories
Engineering Research Centre, University of Brighton
Moulsecoomb, Brighton, BN2 4GJ, UK
{S.H.Lee,R.J.Howlett,S.D.Walters,C.Crua}@Brighton.ac.uk

Abstract. This paper describes a comparative evaluation of two fuzzy-derived techniques for modelling fuel spray penetration in the cylinders of a diesel internal combustion engine. The first model is implemented using conventional fuzzy-based paradigm, where human expertise and operator knowledge were used to select the parameters for the system. The second model used an adaptive neuro-fuzzy inference system (ANFIS), where automatic adjustment of the system parameters is effected by a neural networks based on prior knowledge. Two engine operating parameters were used as inputs to the model, namely in-cylinder pressure and air density. Spray penetration length was modelled on the basis of these two inputs. The models derived using the two techniques were validated using test data that had not been used during training. The ANFIS model was shown to achieve an improved accuracy compared to a pure fuzzy model, based on conveniently selected parameters.

1 Introduction

In a diesel engine, the combustion and emission characteristics are influenced by fuel atomisation, nozzle geometry, injection pressure, shape of inlet port, and other factors. In order to improve air-fuel mixing, it is important to understand the fuel atomisation and spray formation processes. Researchers have investigated the characteristics of the spray behaviour, formation and structure for the high-pressure injector by experimental and theoretical approaches in order to improve the combustion performance and reduce exhaust emissions. However, further detailed studies of the atomisation characteristics and spray development processes of high-pressure diesel sprays are still relevant.

Intelligent systems, software systems incorporating artificial intelligence, have shown many advantages in engineering system control and modelling. They have the ability to rapidly model and learn characteristics of multi-variant complex systems, exhibiting advantages in performance over more conventional mathematical techniques. This has led to them being applied in diverse applications in power systems, manufacturing, optimisation, medicine, signal processing, control, robotics, and social/psychological sciences [1, 2]. Fuzzy logic is a problem-solving technique that derives its power from its ability to draw conclusions and generate responses based on vague, ambiguous, incomplete and imprecise information. To simulate this process of human reasoning it applies the mathematical theory of fuzzy sets first defined by

Zadeh, in 1965 [3]. Fuzzy inference is the process of formulating a mapping from a given input value to an output value using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned. It has been proved that the system can effectively express highly non-linear functional relationships [4]. Fuzzy inference systems (FIS) have been successfully applied in fields such as automatic control, data classification, decision analysis, expert systems and computer vision.

The Adaptive Neuro-Fuzzy Inference System (ANFIS), developed in the early 90s by Jang [5], combines the concepts of fuzzy logic and neural networks to form a hybrid intelligent system that enhances the ability to automatically learn and adapt. Hybrid systems have been used by researchers for modelling and predictions in various engineering systems. The basic idea behind these neuro-adaptive learning techniques is to provide a method for the fuzzy modelling procedure to learn information about a data set, in order to automatically compute the membership function parameters that best allow the associated FIS to track the given input/output data. The membership function parameters are tuned using a combination of least squares estimation and backpropagation algorithm for membership function parameter estimation. These parameters associated with the membership functions will change through the learning process similar to that of a neural network. Their adjustment is facilitated by a gradient vector, which provides a measure of how well the FIS is modelling the input/output data for a given set of parameters. Once the gradient vector is obtained, any of several optimisation routines could be applied in order to adjust the parameters so as to reduce error between the actual and desired outputs. This allows the fuzzy system to learn from the data it is modelling. The approach has the advantage over the pure fuzzy paradigm that the need for the human operator to tune the system by adjusting the bounds of the membership functions is removed.

Many of the combustion problems are exactly the types of problems and issues for which an AI approach appears to be most applicable and has the potential for making better, quicker and more accurate predictions than traditional methods. The increasing availability of advanced computer equipment and sensory systems, frequently results in the production of large amounts of information-rich data, and there are often inadequate means of analysing it so as to extract meaning. The aim of this investigation was to apply intelligent systems tools and techniques to achieve an improved ability to analyse large complex data sets generated during engine research in a semi-automated way. An intelligent paradigm was created based on a fuzzy logic inference system combined with conventional techniques.

2 Methods

2.1 Pure Fuzzy Logic Model

Fuzzy logic provides a practicable way to understand and manually influence the mapping behaviour. In general, fuzzy logic uses simple rules to describe the system of interest rather than analytical equations, making it easy to implement. An advantage, such as robustness and speed, fuzzy logic method is one of the best solutions for system modelling and control. A FIS contains three main components, the fuzzification stage, the rule base and the defuzzification stage. The fuzzification stage is used to

transform the so-called crisp values of the input variables into fuzzy membership values. Then, these membership values are processed within the rule-base using conditional 'if-then' statements. The outputs of the rules are summed and defuzzified into a crisp analogue output value. The effects of variations in the parameters of a FIS can be readily understood and this facilitates calibration of the model.

The system inputs, which in this case are the cylinder pressure and the air density, are called linguistic variables, whereas 'high and 'very high' are linguistic values which are characterised by the membership function. Following the evaluation of the rules, the defuzzification transforms the fuzzy membership values into a crisp output value, for example, the penetration depth. The complexity of a fuzzy logic system with a fixed input-output structure is determined by the number of membership functions used for the fuzzification and defuzzification and by the number of inference levels. A fuzzy system of this kind requires that knowledgeable human operate initialise the system parameters e.g. the membership function bounds. The operator must then optimise these parameters to achieve a required level of accuracy of mapping of the physical system by the fuzzy system. While the visual nature of a fuzzy system facilitates the optimisation of the parameters, the need for it to be accomplished manually is a disadvantage.

2.2 ANFIS Model

ANFIS largely removes the requirement for manual optimisation of the fuzzy system parameters. A neural network is used to automatically tune the system parameters, for example the membership function bounds, leading to improved performance without operator invention. In addition to a purely fuzzy approach, an ANFIS was also developed for the estimation of spray penetration because the combination of neural network and fuzzy logic enables the system to learn and improve its performance based on past data. The neuro-fuzzy system with the learning capability of neural network and with the advantages of the rule-base fuzzy system can improve the performance significantly and can provide a mechanism to incorporate past observations into the classification process. In a neural network the training essentially builds the system. However using a neuro-fuzzy scheme, the system is built by fuzzy logic definitions and then it is refined using neural network training algorithms.

3 Experimental Work

A large collection of spray data are generated using the Ricardo Proteus test engine. These data comprised images depicting the spray patterns of diesel injection processes, under selected conditions of relative pressure, nozzle size and type and in-cylinder air temperature. The images representing time-varying spray under each relative pressure condition were examined and processed using a thresholding technique whereby each image representing the instant of maximum penetration length was then determined, yielding a maximum penetration value which could be linked with its corresponding relative pressure across the injector. The collected maximum spray penetration values and corresponding relative pressures then formed a labelled data to be modelled by the FIS as shown schematically in Figure 1.
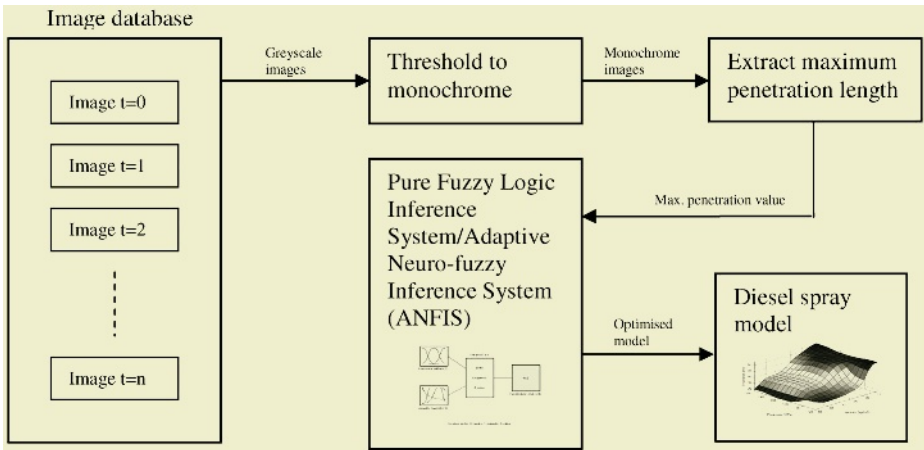
Fig. 1. Schematic diagram of FIS modelling

3.1 Pre-processing

Raw penetration lengths were plotted against time under each relative pressure and density condition. Polynomial fitting was employed to produce best fitted curves where maximum penetration values can be depicted. These were combined into a vector with which to train the ANFIS as shown in Table 1.

Table 1. Training data sets and results

Data set	Parameters		Measured penetration (mm)
	Relative pressure (MPa)	Density (kg/m^3)	
1	60	14	53
2	60	35	32
3	100	14	52
4	100	35	38
5	160	14	54
6	160	35	36

3.2 Pure Fuzzy Inference Model

Figure 2 illustrates the fuzzy sets which were used in the pure fuzzy logic inference system. There were two stages in the inference model, the in-cylinder pressure and the air density; both stages are described in detail. The pressure and density range from 60MPa - 160MPa and 14kg/m^3 - 42kg/m^3 respectively. Both in-cylinder pressure and air density fuzzy sets used generalised bell-shaped membership functions for classes low, medium and high. It was empirically selected based on the features of all data under consideration although in many cases membership functions are fixed and somewhat arbitrarily chosen. The process was carried out by examining the ranges of all data sets to determine where the majority of points were located. The functions were also created to have an approximately equal amount of overlap between each

membership curve. Experimental adjustment of the limits of the membership classes enabled the response of the model to be tailored to the experimental output from the experimental data.

The rule structure is essentially predetermined by the user's interpretation of the characteristics of the input parameters in the model. The contents of these rule-base and membership functions undertake many modifications as part of the process of heuristic optimisation and in many cases it is a continuing process. Examples of the rules initially contained in the rule-base for the pure fuzzy model are shown in Table 2.

Table 2. Fuzzy rule-base

IF Pressure = Low AND Density = Low THEN Penetration = Large
IF Pressure = Low AND Density = Med THEN Penetration = Small
IF Pressure = Low AND Density = High THEN Penetration = Small
IF Pressure = Med AND Density = Low THEN Penetration = Medium
IF Pressure = Med AND Density = Med THEN Penetration = Very Large
IF Pressure = Med AND Density = High THEN Penetration = Very Small
IF Pressure = High AND Density = Low THEN Penetration = Large
IF Pressure = High AND Density = Med THEN Penetration = Medium
IF Pressure = High AND Density = High THEN Penetration = Very Small
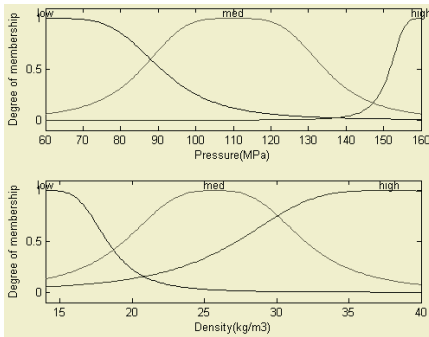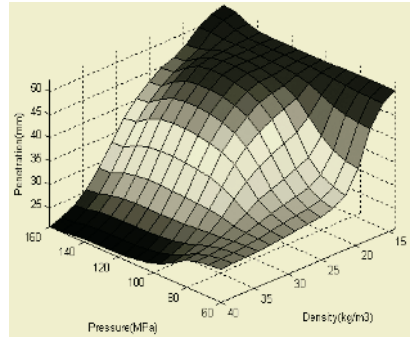
Fig. 2. Fuzzy sets   **Fig. 3.** Pure fuzzy logic model – surface plot

The fuzzifed values for the outputs of the rules were classified into membership sets similarly to the input values. While the output membership functions may be trapezoidal or triangular, in this case, an output singletons were used which has a compact form and computationally efficient representation. The fuzzy output singletons were defuzzified to a crisp value of penetration depth by means of the widely-used centre of gravity method.

The control surface in Figure 3 shows the crisp value of penetration depth at different combinations of in-cylinder pressure and air density using a pure fuzzy logic model. Each of these intersection points indicates the differing predicted value of spray penetration depth, which is determined by the design of fuzzy sets, rule-base and membership functions. The surface plot acts as a practical means of determining the output needed for each combination of input parameters.

3.3 Neuro-fuzzy Model

A FIS was devised using Matlab® based application, ANFIS. A neuro-adaptive learning technique facilitated the learning of information about a data set by the fuzzy modelling procedure, in order to compute the membership function parameters that best allow the associated FIS to track the given input/output data rather than choosing the parameters associated with a given membership function arbitrarily.

A Matlab programme was generated and compiled; The pre-processed input/output spray vector matrix which contained all the necessary representative features was used to train the FIS. Figure 4 shows the structure of the ANFIS; a Sugeno FIS was used in this investigation. Figure 5 shows the fuzzy rule architecture of the FIS which consisted of 9 fuzzy rules. During training in ANFIS, 6 sets of pre-processed data were used to conduct 180 cycles of learning. Figure 6 shows the final membership functions under two different air input conditions derived by training the generalised bell-shaped membership function.
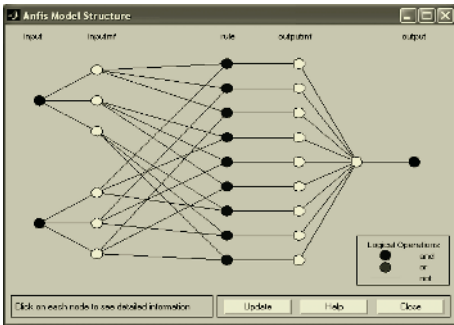
Fig. 4. The ANFIS model structure
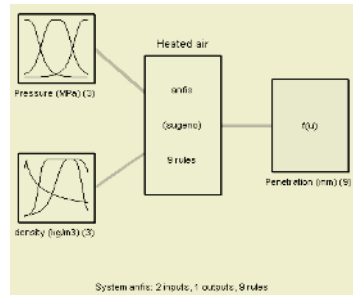
Fig. 5. Fuzzy rule architecture of the generalised bell-shaped membership function
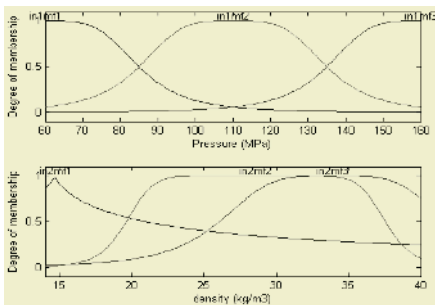
Fig. 6. Fuzzy sets
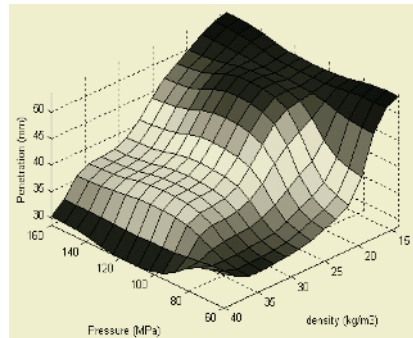
Fig. 7. Surface plot showing relationship between input and output parameters

4   Results and Discussion

Table 3 shows the predicted penetration length obtained from the ANFIS. Figure 7 depicts a three-dimensional plot that represents the mapping from relative pressure

and air density to spray penetration length. As the relative pressure and air density increases, the predicted penetration length increases in a non-linear piecewise manner, this being largely due to non-linearity of the characteristic of the input vector matrix derived from the raw image data. This assumes that these raw image data are fully representative of the features of the data that the trained FIS is intended to model. However the data are inherently noisy and training data may not always faithfully represent all the features of the data that should be presented to the model. Therefore, the accuracy of the model will be adversely affected under such circumstances.

4.1 Model Validation

The data in Table 3 was used to determine how well the FIS model could predict the penetration length corresponding to various values of pressure and density. Figure 8 shows scatter plot of the measured and FIS modelled penetration length utilising six sets of testing data. These two diagrams demonstrate that the predicted values are close to the experimentally-measured values, as many of the data points fall very close to the diagonal (dotted) line, indicating good correlation. Figure 9 shows similar comparisons between the FIS-modelled and measured values of the penetration length using the same testing data. Clearly the model created by ANFIS has a better agreement than the pure fuzzy logic model. The correlation coefficient also suggested identical findings.

Table 3. Testing data and results

Data set	Parameters		Penetration (mm)		
	Relative pressure (MPa)	Density (kg/m³)	Measured	Pure Fuzzy Paradigm	ANFIS
1	60	28	33	30	33
2	60	40	35	28	35
3	100	28	40	40	41
4	100	40	29	23	29
5	160	28	40	39	40
6	160	40	30	21	30
		Correlation coefficient		0.971	0.997
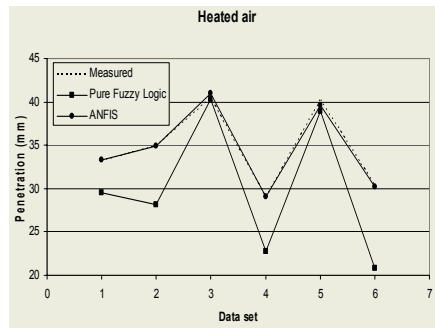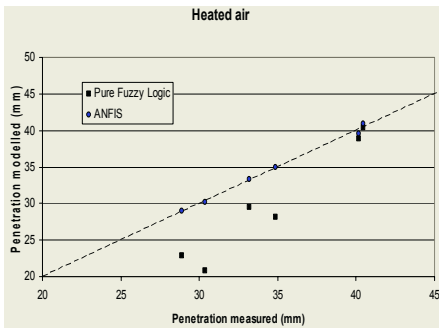
Fig. 8. Scatter plot of measured penetration and predicted penetration

Fig. 9. Comparisons between predicted and measured penetration

4.2 Discussion

The ANFIS is a non-linear computational method that has potential for modelling complex systems with unclear input to output relationships due to its ability to combine fuzzy logic and system identification techniques in a hybrid manner. This type of system has several advantages when assigned to applications in which only partial knowledge of the system characteristics are known, as is typically the case with engineering systems. Additionally, the ANFIS can rapidly identify important characteristics of the data, which is an important and useful feature of models used for estimation purposes in IC engines research. In the experiment, we have used an ANFIS to predict changes in diesel spray penetration depth as a potential means to monitor impending changes in combustion chamber and fuel injector design. As an initial step toward modelling and prediction with an ANFIS for this particular application, it has proven very useful for short-term prediction of penetration depth using engine operating parameters as the input.

The correlation coefficient reflects a model's ability to predict the output based on the input used. While both models performed fairly well and approximated the output function to a reasonable extent, the ANFIS model exhibited improved performance in this respect. Pure fuzzy logic models were conveniently constructed whilst the ANFIS performed well in cases where the input to output relationships become more complex.

5 Conclusions

This paper demonstrated that fuzzy and neuro-fuzzy techniques can be used to model diesel fuel spray penetration for an internal combustion engine, leading to convenient and quick investigation on the effect of penetration length under different operating parameters, including in-cylinder pressure, density, air temperature, etc. The pure fuzzy logic and neuro-fuzzy system, ANFIS employed in this work are quick and robust. It has been applied to sets of pre-processed raw diesel engine spray data and successfully compared. The pure fuzzy logic model employed simple calibrated membership functions and nine optimised rules to represent a diesel spray input/output mapping whilst the neuro-fuzzy model has based on a total of six sets of experimental image data which were used for training the FIS. Both devised models were validated by comparing the predicted results against the experimental data. The correlation coefficient of the penetration length estimated by ANFIS is 0.997. The pure fuzzy logic model has a smaller figure of 0.971 which suggested a poorer correlation with this model.

These fuzzy models set an example of how intelligent technique can be used in diesel spray modelling. The system is very conductive to improvement and adjustment and it can be fine-tuned and improved over time when more engine operating parameters become available. Moreover, these techniques and idea can conveniently be extended to, and be invaluable for, other combustion systems such as modelling and emission predictions in: boilers, furnaces and incinerators. Also, for internal combustion engines, potential applications include modelling and control of: spark ignition engines and gas engines.

References

1. Kalogirou S.A. Applications of artificial neural-networks for energy systems, Appl. Energy 67 (2000) pp.17–35
2. Xu K., Luxmoore A.R., Jones L.M., Deravi F., Integration of neural networks and expert systems for microscopic wear particle analysis, Knowledge-Based Systems 11 (1998) pp.213–227
3. Zadeh L.A. Fuzzy sets, Information Control 8 (1965) pp.338–353
4. Wang L.X. Fuzzy systems are universal approximators. Proceedings of the IEEE International conference on Fuzzy Systems, (1992) pp.1163-1170
5. Jang J. ANFIS: Adaptive network-based fuzzy inference systems, IEEE Transactions on Systems, Man, and Cybernetics 23, (1993) pp.665-685
6. Howlett R.J., de Zoysa M.M., Walters S.D. and Howson P.A. Neural Network Techniques for Monitoring and Control of Internal Combustion Engines, Int. Symposium on Intelligent Industrial Automation 1999
7. Baba N. and Sato K. 1998. A Consideration on the Learning Algorithm of Neural Network, Knowledge-Based Intelligent Electronic system (KES'98), South Australia, (1998) pp.7-12
8. Harris C.J., Brown M., Bossley K.M., Mills D.J., Ming Feng. Advances in Neurofuzzy Algorithms for Real-time Modelling and Control, Engineering Applications of Artificial Intelligence, Vol.9, Issue 1, (1996) pp. 1-16
9. Masters, T. Practical Neural Network Recipes in C++. Academic Press. London, 1993
10. Heywood J.B., Internal Combustion Engine Fundamentals. McGraw-Hill, 1988

Integrating Architecture of Digital Library and e-Learning Based on Intelligent Agent

Sun-Gwan Han[1] and Hee-Seop Han[2]

[1] Dept. of Computer Education, Gyeong-in National University of Education
Gyo-dae Street, Gye-yang gu, Incheon, Korea, 407-753
+81-32-540-1299
`han@gin.ac.kr`
[2] Dept of Computer Science Education, College of Education, Korea University
Anam-dong Sungbuk-ku, Seoul, 136-701, Korea
`anemon@korea.com`

Abstract. This study proposed the integrating architecture of knowledge information systems using intelligent agent. We analyzed the basic architecture of e-Learning system and digital library. We also classified the modules that can effectively integrate the each function of e-Learning and digital library. Through these analysis and classification, we designed the integrating architecture of e-Learning with digital library. We also described the roles and structures based on intelligent agent to be unified between e-Learning system and digital library. Especially, we proposed the integrating strategies of user modeling and learning metadata as the integrating architecture for contents sharing. Furthermore we provided a new integrating model for contributing the knowledge enlargement and the mutual prosperity between e-Learning and digital library.

1 Introduction

Over the past decade there has been a significant development of e-Learning systems and digital libraries. e-Learning is a new paradigm of education system using information and communication technology. Meanwhile, digital library has been developed rapidly as new information and knowledge spaces. It is emphasized to make active communities in e-Learning. On the other hand, it is emphasized to construct digital contents in digital library.

There are a lot of studies on digital libraries for educational practice. However most of these studies are viewed in the digital library systems. The researches on e-Learning have emphasized to the standard learning contents now a days. Although each system area has been made on the developments, these individual developments of e-Learning and digital library brought out a separation. Although they have a lot of relationships between Digital library and e-Learning, there were almost no studies to integrate effectively the both systems until now.

We consider that both systems have to integrate for a mutual development and new knowledge creation. The integration methods are various approaches that are organization approach, system approach, administration approach, content-based approach, user-centered approach, procedural approach, and standardization approach.

Our approach support a wide viewpoint to integrate the e-Learning and digital libraries. In order to support this approach, we presented the new integrating architecture based on an intelligent agent between two systems. The intelligent agent is able to play the important roles for integration. Furthermore, we proposed the three strategies- system integrating, a content integrating, and a user integrating strategies- for the integrating architecture of digital libraries and e-Learning systems.

2   Related Works

In the development of e-Learning systems, the research of the content standardization is studied very important areas. This standardization process includes system architectures, reference model, educational metadata, course structures, student assessment, content packaging, encapsulation, student management, runtime environments, and other specifications[8]. The standardization of learning content is like the proposed SCORM by ADL. When digital library will be integrated with SCORM-based standard, education will be able to transform by new paradigm. To implement this useful environment it is very important to be able to share the contents between digital library systems and e-Learning systems.

Wolfgang Nejdl studied on ontology and metadata for e-Learning that is intelligent agent can reason semantic relation on learning contents[9]. Andreas Schmidt presented the study on user context aware delivery of e-Learning material that effectively deliveries users LOM(learning object materials)[2]. A report on collaboration between IMS(Instructional Management System) and OKI(Open Knowledge Initiative) has been effort the study on an integration of digital repositories between e-Learning system and digital library[10].

In the digital libraries, issues are extending the role for education, sharing encountered information, building digital libraries from the web, constructing knowledge networks, etc[1][4]. Also a lot of studies proposed new models for educational practice. The study of Tamara Sumner and Mary Marlino proposed three models of digital libraries - digital libraries as cognitive tools, digital libraries as component repositories and digital libraries as knowledge networks. They presented the understating of the suite of models should refine and validate much more experiments [11].

Frank Oldenettel et al presented the project LEBONED that focuses on the integration of digital libraries and their contents into web-based learning environments. Their study described in general how the architecture of a standard LMS to enable the integration of digital libraries. They also implemented almost completely the metadata format 1xSCORM. However this paper showed there are still more researching for additional elements and the described concepts and solutions to identify and physically extract components from monolithic documents [6].

The research of Christine L. Borgman et al is how geography professors select materials for classroom lectures. Their studies are the contents searching and creation of geographic information system as well as management of personal digital libraries, space sharing and capabilities to manipulate data and images in digital libraries. This paper proposed their finding have implications for the technical design of the system, such as the importance of the library framework, annotated digital objects, and the need for searching by concept or theme. They also need more researching for implementation of their system[5].

3   Design of Integrating Architecture

3.1   Overview of Integrating Architecture

Proposed integrating areas in this paper are three fields. Figure 1 shows three fields as like system integrating, content integrating, and user model integrating.
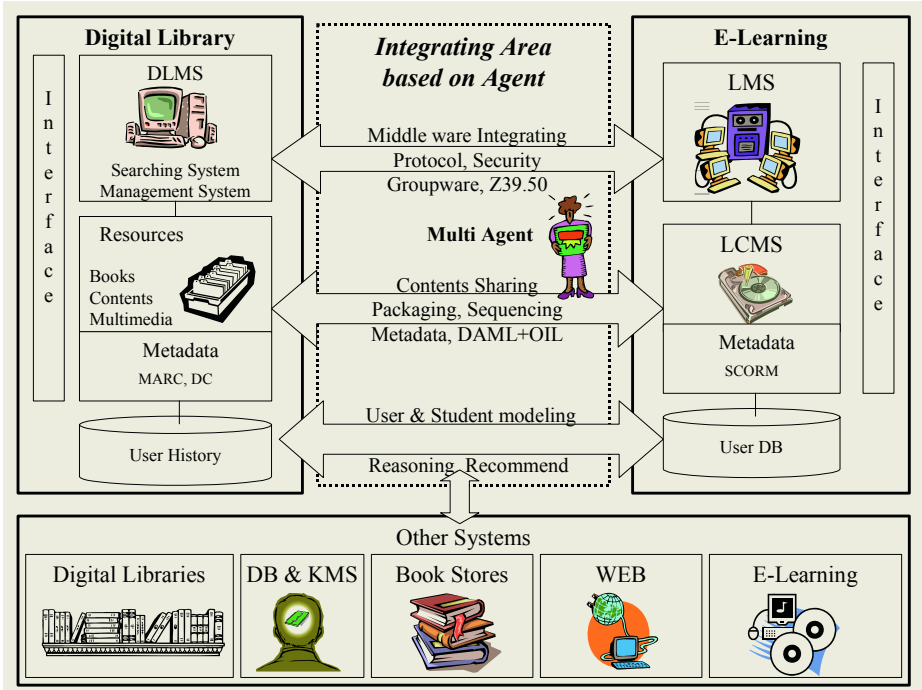
Fig. 1. Integrating Architecture Overview

As a first area, the system integrating considers the method of hardware connection, middleware development and common protocol issues. This area is also important that deals with a security problem occurred when two systems are integrated.

Second area is the strategy of content integrating. This integrating is based on the existing standard metadata on each system. Besides, this integrating requires abstract architecture for applying from content in digital library to content in LCMS.

Thirdly, this paper proposed user data integrating. User data or learner profile is used to analyze user's characteristics and activities through user modeling and two systems reason a learner's abilities and level by using the agent systems. In order to integrate each user data in two systems, we need to extract a learner's attributes and a pattern of user's action. And then, we have to get the knowledge representation of user model as procedural method or declarative method.

The kernel of these integration are applied an intelligent agent system. Especially, when multi-agent is used for these integration, two systems can work individual operation as well as resources of two systems can be integrated and used effectively.

3.2   Integrating Architecture Based on Agent

In the proposed system, the integration between both systems is able to be implemented by multi-agent system as shown figure 2. The multi agent system proposed in this study is modified UMDL agent model that studied on structure and effectiveness of agent in digital library. The types of agent are mainly divided UIA(User Interface Agent), and SA(Supervisor Agent), MA(Mediate Agent).
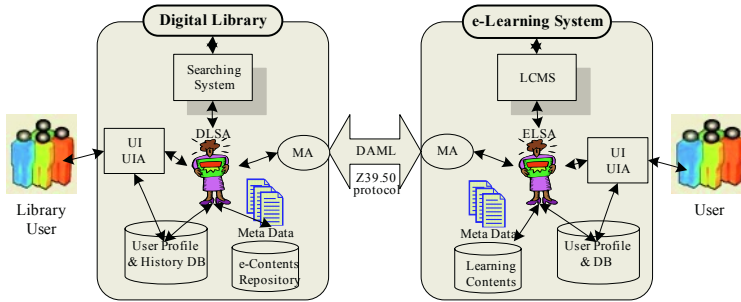
Fig. 2. Integrating architecture based on agent

In e-Learning system, UIA interprets a user's queries and monitors learning process. UIA also updates new user's information and learning history in user database. Then UIA represents a user's queries as transformed texts and forwards ELSA(E-Learning Supervisor Agent) its.

ELSA manages all agents in e-Learning systems and plays an important role as supervisor. ELSA analyzes and infers information submitted by UIA. Besides, ELSA manages controlling learning process with curriculum manage agent, collaboration with content manage agent, supervising security agent, providing learning information to user and so on.

When a learner requests new contents or the other contents in digital library, ELSA submits MA user's requests. Although learner no requests contents, ELSA analyzes a learner's abilities and find a suitable content from digital library or LCMS and offers the output results to a learner. MA works an interpreting ELSA's query, converting query by KQML type, transmitting query to MA in digital library, monitoring process of query and job of protocol security. Intercommunication language between two MAs used DAML(Darpa Agent Markup Language) in this paper. DAML can represent agent's knowledge in semantic web environment. With using OIL, DAML is possible the inference for searching contents.

In digital library, UIA connects users and digital library resources. UIA also represents user's query for searching contents and updates user's profiles for user modeling. DLSA(Digital Library Supervisor Agent) manages all agents in digital library and plays an important role as supervisor as like ELSA. DLSA works collecting, modifying and searching content, monitoring the process of query, registering information about collecting contents, and reforming contents for delivering learning content to e-Learning. MA in digital library does the same works like MA in e-Learning. While a learner processes learning and content searching, collaboration works of agents are as follows.

4   Implementation of Integrating Architecture

4.1   System Integrating

The physically integrating of hardware in both systems is difficult because both systems are managed at an each heterogeneous organization. Thus we need to work to integrate both systems using middleware and intelligent agent. Middleware can support interoperability between two digital repositories. In particular, CORBA, that is a standard of distributed object technology, can integrate distributed systems effectively.

In order to searching, retrieving, and updating for federated data-base architecture, e-Learning system have to use Z39.50 protocol that is an application protocol of information retrieval in digital library. If e-Learning system gets Z39.50 protocol, a learner can be offered various learning content from other e-Learning systems and used infinite contents from digital library easily. In addition, if searching system in two systems uses Xquery based on Z39.50 protocol, we can share different RDBMS.

4.2   Content Integrating

The roles and structures of metadata in both systems are different. Thus there are several problems to share contents in both systems. At the same time, the ontology is needed to share metadata each other but it is very hard problem to construct ontology for these all fields. In order to solve these problems, we proposed new metadata elements for digital library. We used SCORM as a basis for the final metadata format, because we believe that SCORM is the most well-engineered one of the standards.

Figure3 shows the abstract architecture for metadata and document import. The types of metadata and contents are variable at digital libraries. Importers get the desired documents as well as the corresponding metadata from the digital libraries. This intermediate metadata format is a XML representation by the XSLT processing. Creators of this system convert the metadata into SCORM format and split the monolithic Document into SCO. The rules of metadata converting and document fragment are supported from LCMS of e-Learning systems.

4.3   User Information Integrating

Integrating user information means that learner model in e-Learning combine user model based on user profile in digital library. The user modeling in e-Learning supports an adaptive learning through diagnosing learner's abilities, analyzing learner's profile, advising learning strategy, offering suitable contents and intelligent test.

The user model in both systems needs information to be represented by common attributes. Both systems have to be able to cooperate for standard model and to share the information. This transform and collaboration are not easy work because both systems have the different main roles and user's needs. Therefore, unique forms of user models are not changed. We proposed the method that extracts common attributes of user model and defines common learner's attributes by metadata for learner. Then an agent significantly infers information in learner model using this metadata. These shares of user model can make a synergy effects for both systems. A learner in e-Learning can use huge contents in digital library. So to speak, agent offers a learner
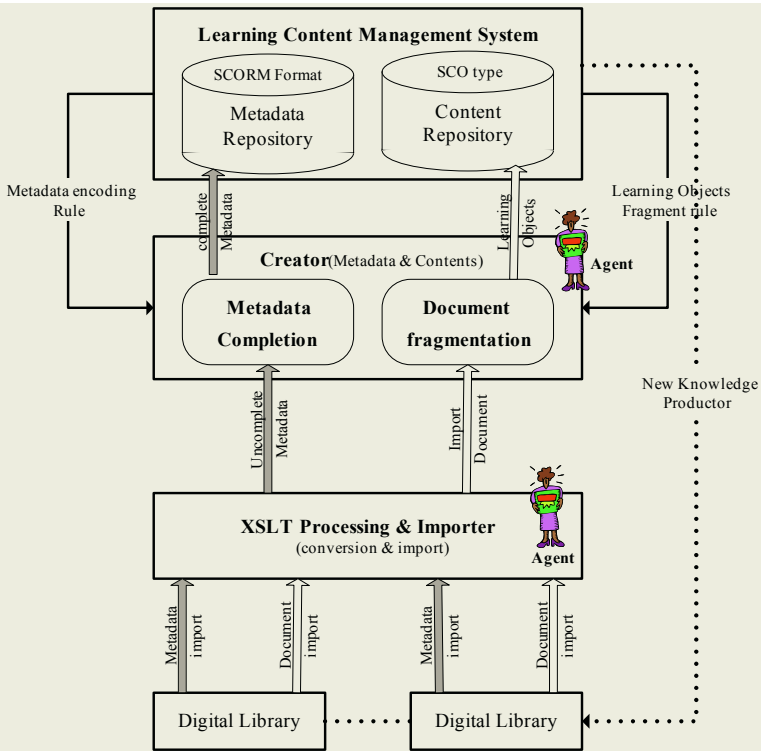
Fig. 3. Abstract architecture for metadata and document import

various services as adaptive learning, content reusing and remaking in digital reposi-
tory, content searching and retrieval, and intelligent testing service.

5   Conclusions

To the integration of the digital library into the LCMS of e-Learning system we pre-
sented technically method and architecture. We briefly described how digital libraries
and e-Learning systems have to be modified to enable the integration. And we de-
scribed the technically and methodological solutions.

We strongly believe that future digital libraries will be knowledge network. And
also such digital libraries will offer useful contents as the core knowledge to e-
Learning systems. As well as the digital libraries can get the benefits from the e-
Learning systems. The learners of e-Learning need contents of digital libraries. And
then the integrating system can easily connect the learners with the contents of digital
libraries as their learning ability in e-Learning environments.

We think that it needs to be done to refine and validate these initial formulations
through clearly much empirical works. And we believe that digital library and e-
Learning system will be able to provide a lot of useful information to each other
through system integrating in the future.

References

1. Aaron Krowne, Martin Halbert, Combined Searching of Web and OAI Digital library Resources, ACM/IEEE Conference, pages 343-344, ACM Press (2004)
2. Andreas Schmidt, Claudia Winterhalter, "User Context Aware Delivery of E-Learning Material", Journal of Universal Computer Science (JUCS), Vol. 10 No. 1, pp. 28-36(2004)
3. Birmingham, W. P., E. H. Durfee, T. Mullen, et al, The distributed agent architecture of the University of Michigan Digital library. AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments, Stanford, CA, AAAI Press(1995).
4. Catherine C. Marshall, Sara Bly, Sharing Encountered Information: Digital Libraries Get a Social Life, ACM/IEEE Conference, pages 218-227, ACM Press (2004)
5. Christine L.B, Gergory H.L, Anne Gilliland-Swetland, Kelli.M, Leslie. C, and Jason. F, How Geography Professors Select Materials for Classroom Lectures: Implications for the Design of Digital Libraries, ACM/IEEE Conference, pages 179-185, ACM Press (2004)
6. Frank Oldenettel, Michael Malachinski, and Dennis Reil, Integrating Digital Libraries into Learning Environments: The *LEBONED* Approach, ACM/IEEE Conference, pages 280-289, ACM Press (2004)
7. IEEE Information Technology - Learning Technology - Learning Objects Metadata LOM: Working Draft 6.1 (2001-04-18). As referenced by the IMS Learning Resource Meta-data Specification Version 1.2.http://ltsc.ieee.org/. (2001)
8. Institute of Electrical and Electronics Engineers (IEEE) Learning Technology Standards Committee (LTSC), (2003)
9. Jan Brase, Wolfgang Nejdl " Ontologies and Metadata for e-Learning" in "Handbook on Ontologies", Springer-Verlag, (2003)
10. S. Griffin, J. Merriman, E-Learning and the Digital library-A Report on Collaboration between IMS and OKI, CNI Fall Task Force Meeting, (2002)
11. Tamara Sumner, Mary Marlino, Digital Libraries and Educational Practice: A Case for New Models, ACM/IEEE Conference, pages 170-178, ACM Press (2004)

A Pedagogical Overview on e-Learning

Javier Andrade, Juan Ares, Rafael García, Santiago Rodríguez,
María Seoane, and Sonia Suárez

Facultad de Informática, Universidade da Coruña
Campus de Elviña s/n 15071, A Coruña, Spain
{jag,juanar,rafael,santi,maria,ssuarez}@udc.es

Abstract. Today, there are multiple approaches that serve as guide for the elaboration of an *e-Learning* system by means of theoretical viewpoints as constructivism. However, these approaches are not fit for stating, in a clear and solid way, their basic pedagogical principles. The analyses of the approaches though this paper aims to link the concepts that they present with the specific pedagogical concepts that, in the authors' opinion, they include. In this way, it is intended to show that this type of proposals should be explicitly sustained by the educative theories that they contain, therefore they would act as a theoretical grounding for developers of *e-Learning* systems, who would then perform their choices with a pedagogical basis.

1 Introduction

In the last decades, the economical organization in the world has advanced towards a service-based economy, whose epicentre is the knowledge [1]. This fact has induced a reduction regarding the useful life of both, knowledge and learned skills, so, the continuous acquisition of knowledge and flairs is fundamental in order to actively participate in society [2]. This need has been the basis for the emergence of the *e-Learning* discipline, which is understood as the use of Internet technologies in order to provide a wide range of solutions that might improve knowledge and performance [3]. However, technology should not be considered as a goal in itself, but as a mean that facilitate learning. The later has been corroborated by, among others, Mioduser et al. [4], who stated, "one step ahead for the technology, two steps back for the pedagogy". Therefore, it is necessary to bear in mind that during a good *e-Learning* course, not only the technical aspects but also—and mainly—the pedagogical ones should work correctly.

From this point of view, several authors as [5, 6, 7 and 3] have devoted their efforts to set approaches out from a theoretical viewpoint such as constructivism in order to produce the elements that should be at students' disposal. Despite this, and as it can be extracted from [8], their descriptions fail when trying to explain the pedagogical foundations of the approaches in a clear and solid way. Consequently, the resulting *e-Learning* systems reflect "common sense" rather than theoretically informed design.

Bearing in mind all that has been previously said, and in order to link concepts with their correspondent pedagogical principles, this paper presents the analysis of the most significant approaches for the elaboration of *e-Learning* systems from their pedagogical point of view. Obviously, the inferred relationships should not be con-

sidered as dogmatic because they derive from the inner perceptions of the authors. Precisely, the aim of this paper is to show the need that the development of e-*Learning* systems might be supported by an explicit pedagogical basis in order to both, avoid the present task of "discovery" and also allow systems developers to make their choices from a pedagogical grounding.

This paper has been structured as follow. Section 2 presents the pedagogical approaches that have been considered for the work. Section 3 shows the main proposals that might guide the elaboration of *e-Learning* systems. Section 4 depicts the relations than, from the authors' viewpoint, exist between pedagogical approaches and proposals. Finally, section 5 presents the obtained conclusions.

2 Pedagogical Approaches

The multiple existing pedagogical theories might be framed within two main trends: behaviourism and constructivism [9]. The behaviourism [10] proposes the modelling of people behaviour by means of concepts as *stimulus* and *response*, so the measurement of learning effectiveness would be made in terms of results, that is, final behaviour. In this way, the cognitive processes used for learning, which are considered as the essential part of psychology, are kept apart. On the contrary, the constructivist trend considers this last aspect that besides represents the core from where this trend organizes its proposals [11].

Regarding the behaviourist trend, the Skinner's Operant Conditioning [10] states that a new conduct might be acquired by means of the *punishments* or *reinforcements* that the student perceives after his/her own *responses*. A *reinforcement* is the presentation (*positive reinforcement*) or withdrawal (*negative reinforcement*) of a *stimulus* after a *response*, so that the frequency of the *response* would be consequently higher. A *punishment* is the presentation (*positive punishment*) or withdrawal (*negative punishment*) of a *stimulus* after a *response*, so that the frequency of the *response* would be consequently lower. The third key concept in this theory is the *shaping*, that allows the gradual consecution of a *response* by means of its *reinforcement*, consisting on successive approximations to that *response*.

Apart from anything else, some of the theories that might be highlighted regarding constructivist trend are, among others, Piaget's Genetic Epistemology, Vygotsky's Social-Cultural Theory, Ausubel's Subsumption Theory and Bruner's Constructivist Theory.

Piaget [12] maintains that the *cognitive development* takes place during different *stages* that individually determine the possibilities of learning. This author also emphasizes the individual aspects of development, so that the core concept of the theory is *equilibrium* notion. In this sense, the teacher should not only induce the student to have *cognitive conflicts* that generate *disequilibrium* among his/her *cognitive structures* or fixed patterns of behaviour, but also lead the learning asks in such way that the pupil might re-establish the *equilibrium*. Piaget also considers that the individual learns mainly by means of the manipulation of the learning subject.

Vygotsky [13] does not coincide with Piaget, who says that an individual cannot learn more than the capacity limit of his/her *stage* of *cognitive development*. Vygotsky differentiates between *Actual Development Level*—the pupil is able to do and to learn on his/her own—and *Potential Development Level*—the pupil is able to do and

to learn if helped by other people. The zone between these two levels is named *Zone of Proximal Development*, and it is the area that should be approached for the pupil to advance from *Actual* to *Potential Development Level* with the cooperation of other people. In such way, learning is based on the progressive assimilation of shared meanings, highlighting, unlike Piaget, its social aspect.

Ausubel [14] starts from Piaget's *cognitive structure* and *development stage* concepts in order to later focus on *significative learning,* which is the establishment of non–arbitrary and substantive links between the new contents and those that the pupil already knew. Therefore, it is fundamental to find out which is the *previous knowledge* that the student has before the learning. With this grounding, the teacher not only should preferably follow a strategy of transmission of already elaborated knowledge (*reception learning*) but also should provide the appropriate *previous organisers*, i.e. those materials that act as cognitive bridges between what the pupil already knows and what he/she needs to know.

Bruner [15] agrees with Ausubel and emphasizes the importance of *significative learning*. However, whereas Ausubel advocated the importance of *reception learning*, Bruner defends the *guided discovery learning.* The later means that, instead providing an already elaborated content to the pupil, the teacher not only should offer appropriate tools and support—named by the author as *scaffolding*—but also stimulate the pupil to discover the knowledge, for instance, by *induction* or *deduction*. In any case, the student acquires the knowledge when it is discovered by him/herself.

3   e-Learning Approaches

There are nowadays multiple approaches to the process of elaboration of an online formative action, that is, the preparation of those elements that should be at pupils' disposal within this context. The main proposals for the elaboration of these systems were selected for the present study, such as the works of Hannafin [5], Horton [6], Jonassen [7] and Rosenberg [3].

Hannafin [5] focuses on the study of Open Learning Environments (OLEs), that this author defines as learning environments focused on the pupil, where individual investigation is predominantly encouraged. These environments are based on the activity and they are supported by four basic elements: *context, resources, tools* and "*scaffolds*". The *context* guides the pupil towards a specific problem or a certain need, whereas the *resources* are reference materials that reinforce learning. The different types of *tools* allow operating with concepts: *information processing (search, compilation, organisation, integration* and *generation), manipulation* and finally, *communication (synchronous* and *asynchronous).* The "*scaffolds*" are those processes that facilitate the learning tasks, so it can be distinguished among *conceptual* "*scaffold*"— what should be considered for task resolution—, *meta-cognitive*—how to reason during learning—, *procedural*—how to uses tools and resources—and *strategic*— alternative proposals that might help.

Horton [6] considers an online course as a combination, mainly, of *lessons, activities* and *collaboration mechanisms*. If *lesson* is understood as a collection of experiences that accomplish one of the sub-goals of the course, the author describes the most common *lesson* structures and also offers tips and guidelines for their use. Horton also presents the most usual learning *activities*, referred to coordinated actions

that exercise basic intellectual skills; that can be used to teach, to exercise and to test skills and knowledge. On the other hand, collaboration can energize learners, promote deeper learning and make learner more self-reliant. However, the author considers that it is very important to select *collaboration mechanisms* attending to each course specific particularities.

Jonassen [7] establishes that knowledge is individually and socially elaborated by the pupils, who base on the interpretations of their experiences. With this aim, the author defines the Constructivist Learning Environments (CLEs), which conceive a *problem*, a *question* or an *example* as the centre of the environment. Around this centre they are arranged *related examples* and *information sources*—which help to understand the problem and also show possible solutions—, *cognitive tools*—that help to interpret and manipulate the different aspects of the problem—, *conversation* and *collaboration tools*—that allow the communities of pupils to negotiate and collaborate on the elaboration of meanings—, and *social/contextual support systems*—that contribute to the implementation of CLE by the users; all of them, supported by *modelling*, *coaching* and *scaffolding*. The *modelling* is focused on showing how expert actors operate when faced with a certain situation. The role of *coaching* not only involves controlling and regulating pupil performance, but also disturbs pupil designs and provides motivating guidelines in order to back the learning beyond pupils' possibilities.

Rosenberg [3] defines the *e-Learning* as the combination of both, *training* and *knowledge management* (KM). The *training* is the way in which instruction is transmitted in order to shape the learning; whereas, KM regards the use of information and tools which might help the students to learn and improve their performance. More specifically, this author highlights certain characteristics that a good *e-Learning* course should have, such as: significant and motivating goals, content authenticity, opportunities for practice (learning by doing), coaching and feed back, possibility of learn from failure (protected failure) and finally models and stories from experts.

4   Possible Connections

4.1   Hannafin's Open Learning Environments (OLEs)

The OLEs seem to be supported by Bruner's proposals. More specifically, the *context* determines the type of problem the pupil has to face up, which shows more keeping with *discovery learning* than with, for instance, Ausubel's *reception learning*. At the same time, for the search of solutions, the student has *tools* and "*scaffolds*" at his/her disposal, which can be related to the *scaffolding* concept. However, Hannafin's work also hides some original concepts, regarding other pedagogical theories, which are mainly reflected in the several tools than the author proposes.

In this sense, the *organisation tools* facilitate the representation of ideas, that is, they help the student to organise maps of concepts that represent complexes relationships such as those named by Piaget as *cognitive structures.* Along the same line, the *integration tools* allow the pupil to link new knowledge with the one he/she already has, therefore to integrate new relationships that modify or re-formulate those *cognitive structures*.

By means of the *manipulation tools* the student can change the content, the value and the parameters of his/her ideas with the aim of evaluate their validity. This task allows the pupil to find out his/her own knowledge; that is the Ausubel's concept of *previous knowledge*.

Lastly, *communication tools* act as a support during the initiation and maintenance of information exchange. Nevertheless, in the authors' opinion, they are not designed to promote the social learning, as Vygotsky proposes, specially because OLEs are defined as environments where individual investigation is the key.

4.2   Horton's Approach

As it is shown at section 3, Horton compiles the most common *lesson* structures and *activities* at *e-Learning* courses, as well as the *collaboration mechanisms* that match more with the peculiarities of each course; therefore, all this work well might represent a significant sample from most of the existing pedagogical proposals. Consequently, it was chosen to perform a categorisation of those structures according to the educative theory that might match better in each case, because a detailed analysis of each concept implies an exclusive extension.

With regards to the *lesson* structures that were considered by Horton, they mainly reflect Ausubel's proposals (e.g. classic tutorials, learner-centred lessons, learner-customized tutorials, knowledge-paced tutorials and generated lessons) and Bruner's ones (e.g. activity-centred lessons and exploratory tutorials).

The learning *activities* might be focused on either Ausubel's *reception learning* (e.g. web casts y presentation sequences), Bruner's *discovery learning* (e.g. scavenger hunt, guided research and case study), Vygotsky's *social learning* (e.g. team design, brainstorming, role-playing scenario, group critique and learning game), Piaget's manipulation of the learning subject (e.g. virtual laboratory) or Skinner's *shaping* concept (e.g. hands-on activity).

4.3   Jonassen's Constructivist Learning Environments (CLEs)

The denomination of Jonassen's proposal shows itself its constructivist character. Likewise OLEs, the CLEs seems to be based on Bruner's proposals. This owns to the fact that they place the *problem*, the *question* or the *example*, for whose solution the student has multiple support resources, in the centre of the environment. Nevertheless, a more detailed analysis of this proposal reveals the presence of concepts that belong to other constructivist as well as behaviourist authors.

In this sense, the *related examples*, the *modelling* and the *information sources* can be related to Ausubel's *previous organisers*, because they help the student to understand the problem. Jonassen does not qualify the need of explicitly determining the *previous knowledge* of the pupil, although he sporadically mentions the *significative learning* concept, therefore linking with Ausubel's proposals.

On the other hand, the use of *cognitive tools* is related to Piaget's viewpoint which says that the pupil learns from the manipulation of the learning subject. *Coaching* disturb the designs of the student and also provide motivating guidelines for reaching success, all of which appear to be quite similar to the *disequilibrium* and *equilibrium* conditions that were described by Piaget.

Regarding Vygotsky's ideas, they are reflected in both, *conversation* and *collaboration tools* as well as in *scaffolding*. The first ones make learning possible, whereas the later intends to back the pupils beyond their possibilities, that is, to help them to advance from the *Actual* to the *Potential Development Level*.

Finally, Jonassen thinks that it is advisable that pupils receive a *feed back* that might indicate the effects of their actions, which, although disguised as constructivism, appears to be related with Skinner's *positive reinforcement* and *positive punishment.* It is then intended to make the student aware about if it is advisable or not to repeat a certain action when circumstances might lead him/her to the same problem or to a similar one, which is in itself a kind of behaviour modelling.

4.4 Rosenberg's Approach

Differently from the general tendency and in accordance with Ausubel's, this approach emphasizes the need of preserving the authenticity of the contents offered to the students by means of constant revision and actualisation. This viewpoint matches with *reception learning* where the contents are elaborated before providing the pupils with them. In fact, this proposal shows clearly that this type of learning is not synonymous with contents memorisation because it prefers to put them in practice (learning by doing) so that at the same time, learning can be extracted from failures (protected failure). By means of the protected failure, the student can manipulate the learning subject and, at the same time, face up a wide range of challenges, all of which is in accordance with the learning method that Piaget defends.

Rosenberg also highlights the importance of the models and stories from experts as support for learning, so that the pupil might analyse the similarities and differences between them and the task that has to be solved, which is in accordance with the essence of Bruner's reasoning by *induction* or *deduction*.

Lastly, the author proposes the use of *coaching* and *feed back* for the student to be conscious about the consequences of his/her acts, which, as in Jonassen model, is related with Skinner's *positive reinforcement* and *positive punishment* concepts.

5 Conclusions

It is widely accepted that, within an *e-Learning* system, is fundamental that, not only the technological aspects but also the pedagogical ones must work fine. Consequently, the approaches for the elaboration of this type of systems, as a general rule, try to take a theoretical stand from a constructivist viewpoint. However, they do not explicitly show the pedagogical principles on which they are based, therefore final systems usually reflect "common sense" rather than theorically informed design. Due to what has been said, this paper presents an analysis of the main existing approaches, with the aim of linking the proposed concepts with their correspondent specific pedagogical principles. In this way, a reference theoretical basis is offered to the developers of *e-Learning* systems for their choices to be theoretically referenced. Logically, the detected relationships should not be dogmatically born in mind because they are the result of the authors' own perceptions, which reinforce these proposals' need to be supported by explicit educative theories.

References

1. Tissen, R., Andriessen, D., Deprez, F.L.: The Knowledge Dividend. Financial Times Prentice – Hall, London (2000)
2. European Commission: e-Learning – Designing tomorrow's education (2000). URL: http://europa.eu.int/comm/education/doc/official/keydoc/com2000/com2000-318en.pdf
3. Rosenberg, M.J.: E-Learning: Strategies for Delivering Knowledge in the Digital Age. McGraw-Hill, New York (2000)
4. Mioduser, D., Nachmias, R., Oren, A., Lahav, O.: Web-based learning environments: Current states and emerging trends. Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications, Seattle (1999)
5. Hannafin, M., Land, S., Oliver, K.: Open Learning Environments: Foundations, methods, and models. In: C. Reigeluth (ed.) Instructional-design theories and models (Volume II). Erlbaum, New Jersey (1999)
6. Horton, W.: Designing Web-Based Training. Wiley, New York (2000)
7. Jonassen, D.: Designing constructivist learning environments. In C. Reigeluth (ed.) Instructional-design theories and models (Volume II). Erlbaum, New Jersey (1999)
8. Oliver, R.: Winning the toss and electing to bat: maximising the opportunities of online learning. In C. Rust (ed.) Proceedings of the 9th Improving Student Learning Conference, Oxford (2002) 35-44
9. Bou, G., Trinidad, C., Huguet, L.: e-learning. Anaya Multimedia, Madrid (2003)
10. Skinner, B.F.: About Behaviorism. Vintage/Ebury, London (1975)
11. Twomey, C.: Constructivism: Theory, Perspectives, and Practice. Teachers College Press, New York (1996)
12. Piaget, J.: Psychology and epistemology: towards a theory of knowledge. Viking Press, New York (1978)
13. Wertsch, J.V.: Vygotsky and the social formation of mind. Harvard University Press, Massachusetts (1985)
14. Ausubel, D.P.: Educational psychology: A cognitive view. Holt Rinehart and Winston, New York (1978)
15. Bruner, J.S.: The Process of Education. Harvard University Press, Massachusetts (1977)

Modeling Understanding Level
of Learner in Collaborative Learning
Using Bayesian Network

Akira Komedani, Tomoko Kojiri, and Toyohide Watanabe

Department of Systems and Social Informatics
Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan
Phone: +81-52-789-2735, FAX: +81-52-789-3808
{k-akira,kojiri,watanabe}@watanabe.ss.is.nagoya-u.ac.jp

Abstract. In the collaborative learning, the learner often focuses on the particular person based on the person's understanding level. If the information about particular person whom the learner focuses on is acquired automatically, the learner is able to understand the target person easily. So, it is necessary to estimate the understanding levels of others. However, in the collaborative learning environment, since the learner does not utter all the knowledge that he knows, to externalize explicitly the understanding levels of the learners is difficult. The understanding level about the knowledge which is not uttered by the learner is also estimated from the uttered information. So, in this paper, we define solution network which represents the relations of knowledge in the exercise with their strengths. When the utterance is generated, the understanding level of the uttered knowledge and its related knowledge is estimated by means of the solution network.

1 Introduction

With a rapid spread of computer and network technology, many systems which support collaborative learning have been constructed[1]. The collaborative learning is one of learning styles in which various learners exchange their opinions and solve common problems through a network. In such a learning environment, learners sometimes cannot communicate with each other effectively because communication means are restricted.

A learner often focuses on particular person in order to know the person deeply and acquire knowledge from the person efficiently. If the information about particular person whom the learner focuses on is acquired and is shown to the learner automatically, the learner is able to understand the target person easily. Ito, et al.[2] constructed a system which determines a target person whom the learner focuses on, based on patterns of utterances, such as a type of utterance, target person and utterer. The system also acquires the face of the target person through network cameras automatically. However, in the collaborative learning, a learner often focuses on the particular person based on

the person's understanding level. Thus, In our approach, the other learners are identified based on their understanding levels. The following steps are needed in order to determine the target person for the learner;

- estimating understanding levels of others,
- modeling understanding level of the learner itself, and
- determining the target person by comparing the understanding levels of others based on that of the learner.

In this paper, we discuss a method for estimating understanding levels of others.

In order to determine the target person whom the learner is interested in, it is necessary to estimate the understanding levels of others. In the research of ITS(Intelligent Tutoring System), understanding levels of learners are estimated based on answers written by them. Student modeler in SINT[3] estimates understanding levels of learners based on a dialogue. Andrew, et al.[4] models learners based on thirty questions multiple-choice test. In such a learning environment, a learner tends to write all the knowledge that he knows stepwisely so that it is easier for the system to estimate his understanding level correctly. However, in the collaborative learning environment, the learner does not utter all the knowledge that he knows. The system cannot assume that the learner dose not understand the knowledge which is not uttered by the learner[5]. Therefore, in this paper, the understanding level of knowledge which is not uttered by the learner is estimated by the uttered knowledge. The knowledge which is closely related to the uttered knowledge is also regarded to be understood by the utterer. So, in our approach, the knowledge used in the exercise and their relations are modeled as the Bayesian Network structure called *solution network*. The solution network represents the relations of knowledge in the exercise with their strengths. Based on the solution network, the learner's understanding levels of knowledge which are related to the uttered knowledge are able to be estimated. However, the understanding levels of the knowledge which is not uttered are not estimated correctly. Therefore, the probability of correctness is attached to the understanding levels of knowledge which are estimated by the system. In this paper, the collaborative learning of mathematical exercises in high school is addressed.

2   Approach

Answers of mathematical exercises consist of several solution steps in each of which corresponds to a derivation method or a formula. In the collaborative learning, the learner wants to know which steps are derived by other learners respectively. So, the understanding level in each step needs to be managed. The followings are features of estimating understanding level of a learner.

1. The understanding level for a exercise can be estimated by his utterances. The learner indicates a particular solution step in generating utterance. The understanding levels in all solution steps are determined by all utterances

which he generated. Therefore, the understanding level for a exercise needs to be estimated by his utterances. From now on, the understanding level for a exercise estimated according to each utterance is defined as *understanding rate* and understanding level estimated by all utterances is defined as *understanding level*.

2. Degree of understanding rate is estimated according to types of utterances. The learner does not always understand the solution step that he mentions. For example, when a learner inquires about a solution step, he hardly understand the target solution step.

3. The understanding rates in solution steps which are not uttered are estimated by the utterance which indicates a particular solution step.
   For example, if a solution step is understood by a learner, solution steps that should be derived before the derived solution step are also regarded to have been understood already. Of course, the understanding rates in solution steps which are not uttered explicitly is ambiguous.

According to the features above, we propose the system which estimates the understanding levels of others based on their utterances, in order to model the others from a viewpoint of a learner.
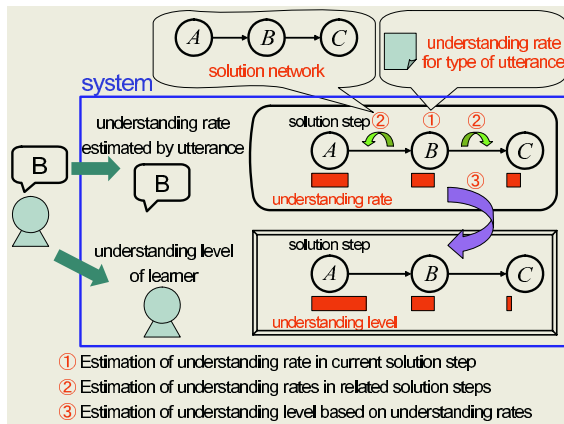
Fig. 1. Framework of estimating understanding level

Figure 1 shows the framework of estimating an understanding level of another learner in our system. The system contains the degree of understanding rate for each type of utterance. It also has the solution network, which holds solution steps of the exercise with strengths of their relations. When the utterance which indicates the particular solution step has been generated, the understanding rate in the target solution step is estimated according to its type. Then, understanding rates in related solution steps are estimated based on understanding rate in the target solution step and strengths of their relations represented as the solution network. Moreover, a probability of correctness of estimated understanding

rates are also calculated according to the solution network. The understanding rate in solution step whose probability of correctness is bigger should reflect to the understanding level.

3   Understanding Level Based on Utterance

3.1   Understanding Rate in Accordance with Type of Utterance

The understanding rate in a solution step in which a learner indicates is estimated from his intention to utter his opinion. For example, when the learner inquires about the solution step, he hardly understands the target solution step. In order to discriminate the understanding rate, six types of utterances are introduced initially: "proposition", "opinion", "inquiry", "reply", "agreement" and "objection". Then, the understanding rate for each type is defined. The understanding rate takes a value of 0 to 1. The higher the value of the understanding rate is, the more the learner understands the solution step. Table 1 shows the example of understanding rates for types of utterances. When the utterance contains a keyword of a solution step, the understanding rate in the solution step is estimated according to its type.

Table 1. Example of understanding rate for types of utterances

types of utterances	reply	opinion	agreement	objection	proposition	inquiry
understanding rate	0.9	0.8	0.6	0.6	0.4	0.1

3.2   Solution Network

The solution network represents the knowledge, which is used in the exercise of the collaborative learning, with the strengths of their relations. Since the solution network is structured as Bayesian Network, the solution network consists of nodes and links. The nodes represent solution steps and contain keywords which characterize the solution steps. The nodes also hold the degree of dependency on the former solution step, which show the probability that the current solution step is also understood if the former solution step is understood, and the probability that the solution step is understood even if the former solution step is not understood. The links represent an order among the solution steps.

Figure 2 shows an example of the solution network. In the mathematical exercises, a solution step is assigned to the statement which corresponds to one formula or derivation method. The solution step is characterized by the equations derived from the formula. In this example, the solution step 2 is characterized by the equation $f(4) = 4 + a$ and the statement "maximum value, 4+a". So, this node contains keywords such as, "$f(4) = 4 + a$" and "maximum value, 4+a". Moreover, this node holds the probabilities which represent dependency on the solution step 1.

The understanding rate in the related solution steps is calculated by the following equations.
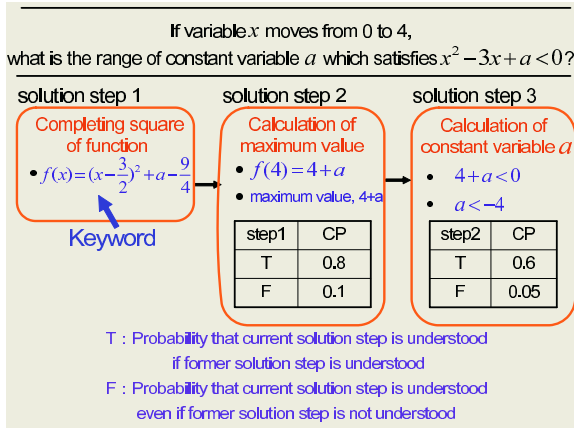
Fig. 2. Example of solution network

– Understanding rate in the next solution step for current solution step:

$$u(N) = CP(N|C)u(C) + CP(N|\bar{C})(1 - u(C)) \tag{1}$$

– Understanding rate in the previous solution step for current solution step:

$$u(B) = \frac{u(C) - CP(C|\bar{B})}{CP(C|B) - CP(C|\bar{B})} \tag{2}$$

In these equations, C indicates a current solution step, N the next solution step, and B the previous solution step. $u(C)$ is the understanding rate in the current solution step. $CP(C|B)$ corresponds to the probability that solution step C is understood if the solution step B is derived and $CP(C|\bar{B})$ the probability that the solution step C is understood even if the solution step B is not derived. For example, if the "opinions" in the solution step 2 whose understanding rate is 0.8 is generated in the solution network of Figure 2, the understanding rate in the solution step 3 is calculated as 0.49 by the equation (1).

3.3   Estimation of Understanding Level

The understanding rates which are estimated by using the solution network are not always correct. That is, the understanding rates in solution steps which are closely to the uttered solution step are more correct than those that are apart from the uttered solution step. The understanding rates with large accuracy should have great influences on the estimation of the understanding level in all the solution steps. Therefore, the probability of correctness is attached to each understanding rate. The following is an equation for determining the probability of correctness for the understanding rate:

$$t(I) = 1.00 - \alpha \times NL(I - C), \tag{3}$$

where I is an estimated solution step and C an uttered solution step. $t(I)$ corresponds to the probability of correctness of the understanding rate in the solution step I and $NL(I-C)$ indicates the number of links between the solution steps I and C. α, which is a constant, represents a decreasing degree of correctness per link. If the understanding level is calculated by the understanding rate with low accuracy, the understanding level is not trustworthy as well. The probability of correctness of the understanding level is defined as follows:

$$T_n(X) = \frac{T_{n-1}(X)}{T_{n-1}(X) + t(X)} \times T_{n-1}(X) + \frac{t(X)}{T_{n-1}(X) + t(X)} \times t(X), \qquad (4)$$

where n is the number of utterances which are input in the discussion and X is a target solution step. $t(X)$ is the probability of correctness of the understanding rate, while $T_n(X)$ is that of the understanding level when n utterances have been generated. In this equation, the value whose probability of correctness is larger is selected in precedence. Based on these probabilities, the current understanding level is calculated. The following is an equation for determining the understanding level in the solution step X after n utterances have been generated.

$$U_n(X) = \frac{T_{n-1}(X)}{T_{n-1}(X) + t(X)} \times U_{n-1}(X) + \frac{t(X)}{T_{n-1}(X) + t(X)} \times u(X) \qquad (5)$$

$u(X)$ is the understanding rate in the solution step X, while $U_n(X)$ is the understanding level when the n utterances have been generated. According to this equation, the value which is more correct is reflected to the current understanding level as well.

4   Experiment

We implemented the proposed mechanism as a part of our collaborative learning environment called HARMONY[6]. In this learning environment, a learner needs to indicate the types of utterances on inputting his utterances.

In this experiment, five students in our laboratory were participated to solve a mathematical exercise of high school collaboratively. The solution steps and their keywords of the mathematical exercise are shown in Figure 2. The objective of our experiment is to evaluate whether the calculated understanding rate in the solution step is adequate. So, the examinees were asked to point out the degrees of the understanding rates in individual solution steps of other examinees when the utterances with keywords had been generated. The average of the understanding rates pointed out by the examinees are compared with the understanding rate calculated by the prototype system.

During the learning, 3 out of 101 utterances included keywords. Tables 2, 3 and 4 show the average of the understanding rates of utterer that are pointed out by examinees, the understanding rate calculated by the system, and differences between them. The understanding rates calculated by the system to be validate, if the difference is less than 1.5. In Table 2, the understanding rate in the solution

Table 2. Understanding rates for utterance 1

utterance 1. $f(x) = (x - \frac{3}{2})^2 + a - \frac{9}{4}$ (proposition)			
solution steps	solution step 1	solution step 2	solution step 3
average value of examinees	0.825	0.3	—
estimated value of system	0.4	0.42	0.365
difference	0.425	0.12	—

Table 3. Understanding rates for utterance 2

utterance 2. So, the answer is $4 + a < 0$, isn't it? (opinion)			
solution steps	solution step 1	solution step 2	solution step 3
average value of examinees	0.5	0.566	0.866
estimated value of system	1.0	1.0	0.8
difference	0.5	0.434	0.066

Table 4. Understanding rates for utterance 3

utterance 3. I was mistaken. That is $a < -4$. (opinion)			
solution steps	solution step 1	solution step 2	solution step 3
average value of examinees	0.55	0.7	0.933
estimated value of system	1.0	1.0	0.8
difference	0.45	0.3	0.133

step whose keyword was input was not estimated correctly. In this experiment, the understanding rate for the proposition is defined as 0.4, because proposition is generated by the learner who comes up the new idea but is not sure to its correctness. However, examinees thought that the utterer who input proposition understood the target solution step. So, the understanding rates assigned to each type of utterance should be reviewed so as to satisfy the examinees' senses.

On the other hand, Tables 3 and 4 are the results of utterances whose types are "opinion". These utterances indicate the solution step 3. Based on the result, the system can estimate the understanding rate in the solution step 3 almost correctly. However, the understanding rates in other solution steps are not estimated correctly. This is because the examinees solved the exercise by the answering path which was different from what we assumed. Since these solution steps 1 and 2 are not the solution steps that were derived by the examinees, these value were not sure about the utterer's understanding rate in those solution steps. If the system held nodes in these solution steps, it could have estimated understanding rates in the derived and underived solution steps more correctly.

5   Conclusion

In this paper, we proposed a method of estimating understanding levels of others from utterances so as to identify the target person of the learner. Based on the experimental result, the understanding rate in the solution step whose keyword

is included is estimated correctly if the understanding rate for the types of utterances is assigned validly. In order to set appropriate understanding rate for the types of utterances, we need to investigate impressions that each type of utterance gives to learners. Moreover, the system cannot estimate the understanding rates in the solution steps which are not prepared as the nodes in the solution network. To represent the several paths for the answer as the solution network should be needed.

For our future work, we need to construct the mechanisms for modeling the understanding level of the learner itself and determine the target person of the learner by comparing others' understanding levels based on that of the learner. Then, we should define the information of the target person to be proposed to the learner so as to make their discussion active.

References

1. T. Kojiri, and T. Watanabe: HARMONY: Web-based Adaptive Collaborative Learning Environment, Proc. of ICCE/SchoolNet 2001, pp.559-566 (2001).
2. T. Kojiri, Y. Ito and T. Watanabe: User-oriented Interface for Collaborative Learning Environment, Proc. of ICCE 2002, Vol.1, pp.213-214 (2002).
3. A. Mitrovic: SINT -a Symbolic Integration Tutor, Proc. of ITS 1996, pp.587-595 (1996).
4. A. Mabbott and S. Bull: Alternative Views on Knowledge: Presentation of Open Learner Models, Proc. of ITS 2004, pp.689-698 (2004).
5. A. S. Gertner and K. Vanlehn: Andes: A Coached Problem Solving Environment for Physics, Proc. of ITS 2000, pp.133-142 (2000).
6. T. Kojiri and T. Watanabe: Agent-oriented Support Environment in Web-based Collaborative Learning, Proc. of ICCE/ICCAI 2000, Vol.1, pp.419-426 (2000).

Dynamic Generation of Diagrams for Supporting Solving Process of Mathematical Exercises

Yosuke Murase[1], Tomoko Kojiri[2], and Toyohide Watanabe[1]

[1] Department of Systems and Social Informatics
Graduate School of Information Science, Nagoya University
[2] Information Technology Center, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan
{murase,kojiri,watanabe}@watanabe.ss.is.nagoya-u.ac.jp

Abstract. We have constructed the agent that takes a role of a teacher and assists the group for deriving answer of mathematical exercises in the collaborative learning support system. In mathematical exercises, the complementary diagram is drawn for the purpose of grasping the current learning situation and deducing the derivation method for the next answering step. So, the objective of our research is to construct the advising mechanism for the agent that proposes the diagram in the next answering step. In the environment that students can freely draw a diagram, it is unsuitable to prepare the appropriate diagram for each answering step in advance. So, our system analyzes the students' diagram and generates a new diagram dynamically based on the analyzed information. In our approach, stepwise generation of the complementary diagram corresponding to each step is regarded as a forward reasoning in the production system. The situation is regarded as the learner's diagram and the prepared rules are regarded as rules for drawing diagram. So, the diagram for each answering step is generated dynamically by applying particular rules to the current diagram. Consequently, by the proposed diagram, students can arrange the current learning situation and understand the derivation method for the next answering step easily.

1 Introduction

In recent years, the collaborative learning through network is one of hottest subjects. In the collaborative learning, plural students solve common exercises collaboratively by exchanging their opinions with each other. However, if the understanding levels of students are different and the communication among students is ineffective, students sometimes cannot attain to their learning goals. In order to deal with such situation, to grasp the learning situation of the group and generate appropriate advices is very important.

There are some researches which aim at assisting collaborative learning. Nakamura, et al. constructed a pseudo learner that grasps the understanding level of the individual student and makes statements instead of the student, when the learning group had a passive discussion and students fell into impasse situation[1]. Supnithi, et al. proposed the integrated model "Opportunistic

Group Formation(OGF)", which is a function to form a collaborative learning group dynamically[2]. When OGF detects the situation for a learner to shift from individual learning mode to collaborative learning mode, it forms a learning group whose members are assigned a reasonable learning goals and social roles. These researches assume that each student learns the learning materials by him/herself. And, by learning these materials in the learning group, these researches attempt to step up students' problem solving ability. On the other hand, in our research, the learning materials are assumed to be prepared for the group learning, and to be learned through the interaction with other learners.

We focus on collaborative learning of high school-level mathematics. We have developed the agent which grasps global learning situation of the group by extracting keywords uttered through the chat-based interaction[3][4]. Then, if the agent detects an impasse situation, it gives the advice as one of the utterances. In general, in solving mathematical exercises, students often draw diagrams so as to arrange the current learning situation and deduce the derivation method for the next answering step. Therefore, the objective of our research is to construct the advising mechanism of proposing appropriate diagrams which enable the group to derive the next answering step according to the current learning situation.

Ito, et al. analyzed the student activities for drawing diagrams for geometry exercises[5][6]. Based on this analysis, students tend to add particular figures to current diagram supplementarily in deducing the next answering step. By adding supplementary figures, students change viewpoints for deducing the next answering step and also make the learning goal and the answering steps to the learning goal clear. To show the diagram to the group is not to give answers directly but to promote more active discussion. Diagrams drawn by students are all different, especially for the mathematical exercises of function. So, to show the diagram which is prepared for each answering step in advance is not suitable for the current situation. In addition, the supplementary figure is able to be defined according to the derivation method. Therefore, in our approach, the rules for drawing supplementary figures are prepared. Then, the diagram in each answering step is generated dynamically by applying particular rules to the current diagram.

In Section 2, the approach to generate the diagram and the outline of our system are introduced. In Section 3, the mechanism for generating diagram is shown in detail. The prototype system and its evaluation are discussed in Section 4. Finally, our conclusion is shown in Section 5.

2   Approach

In the collaborative learning environment, the students share a diagram supplementarily and derive the correct answer through discussing with each other. The diagram which each student draws is not always identical even if the learning situations are the same. Therefore, in the environment in which students can freely draw a diagram, it is unsuitable to prepare the diagram statically.

So, our system analyzes the students' diagrams and generates a new diagram dynamically based on students' diagram.

A process of solving mathematical exercise consists of many answering steps and the derivation method is applied to proceed the learning from one answering step to another. An example of answering steps in mathematical exercises is shown in Figure 1. The derivation method is corresponded to theorem, rule, or formula. In each answering step, the diagrams correspond to derived equations are drawn in order to arrange the current learning situation. On the other hand, the figures that indicate the derivation method for the next answering step are added to the current diagram. These figures are defined according to the derivation method. In the mathematical exercises of function, supplementary figures that indicate sub-goals or preconditions of the derivation method are used mostly.
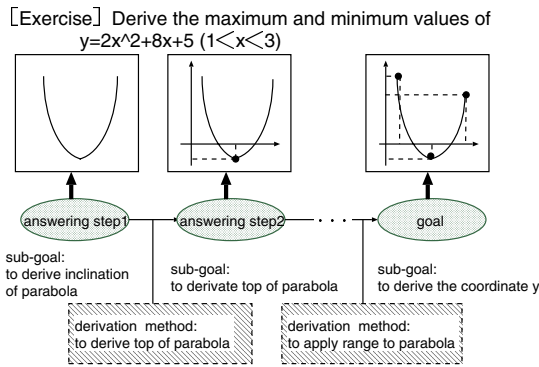
Fig. 1. Answering steps in mathematical exercise

In the collaborative learning, it is effective for the system to give the advice so that students grasp the next derivation method. One of the advantages in the collaborative learning is that various answering viewpoints are derived from plural students. Therefore, to prepare the diagram for each learning step in advance and provide the same diagram for every learning situation limits the derivation method to be discussed in the group. In our approach, stepwise generation of the complementary diagram corresponding to each step is regarded as a forward reasoning in the production system. That is, the appropriate drawing rule that corresponds to a method for drawing supplementary figure of particular derivation method is selected and applied to the diagram of current answering step. By applying the rules whose preconditions are satisfied with the current answering step, different diagrams are generated according to the learning situation and even in the same situation. In addition, whether they are appropriate or not, various diagrams may become a trigger of an active discussion and of discovering various answering viewpoints.

Figure 2 shows a conceptual configuration of our system for generating supplementary figures. Our system consists of two layers; model conversion layer
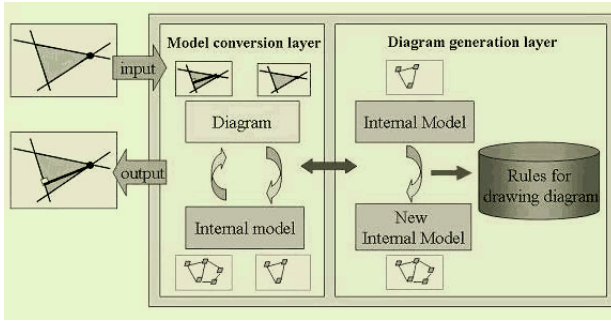
Fig. 2. System architecture

and diagram generation layer. In the model conversion layer, diagrams drawn by the group are transformed into/from internal model that is the representation of a diagram inside the system. If the impasse situation is detected, the internal model for the diagram of the group is sent to the diagram generation layer. In the diagram generation layer, a new internal model is generated by applying rules for drawing diagram to current internal model. Rules for drawing diagram indicate the supplementary figure to be added for each derivation method. Then internal model is sent back to the model conversion layer, is transformed into the diagram, and then is provided to the group.

3   Mechanism for Generating Diagram

3.1   Internal Model

The internal model is the representation of a diagram inside the system. In order to add supplementary figures to the diagram of the group, the system needs to grasp the figures that compose the diagram. However, the scale or range of figures drawn by students is different from each other. If figures of two diagrams are about the same, the same rules for drawing diagram can be applied. Therefore, the internal model is sufficiently definedd as the figures and the relations between figures.

Figures and relations between figures are expressed as predicates. Table 1 shows the internal model of mathematical exercise in two-dimensional function. Variables x and y indicate the figures. Currently, 6 kinds of figures and 12 kinds of relations are defined, which are enough to represent all preconditions of the derivation method for answering two-dimensional function.

3.2   Rules for Drawing Diagrams

Rule for drawing diagrams indicate the rule to add supplementary figures for each derivation method. The rule for drawing diagrams is expressed in the form

Table 1. Internal model of mathematical exercise in two-dimensional function

	point(x)	line(x)
figure	segment(x)	parabola(x)
	x-axis(x)	y-axis(x)
	cross(x, y)	on(x, y)
relation	contact(x,y)	apart(x, y)
between	parallel(x, y)	vertical(x, y)
two	top(x, y)	pivot(x, y)
figures	edge(x, y)	same-long(x, y)
	same-shape(x, y)	opposite-shape(x, y)

of "($\langle$rule number$\rangle$ IF $\langle$conditional part$\rangle$ THEN $\langle$operation part$\rangle$)". In the conditional part, the internal model that corresponds to the precondition of the derivation method is described. When applying the derivation method, particular figures and relations are added to the current internal model. So, the operation part is defined as "(add $\langle$supplementary figures and relations between the figures$\rangle$)".

Table 2. Example of rule for drawing diagrams

$\langle\langle$The rule to add for deriving the top of parabola$\rangle\rangle$
(P1   IF parabola(obj1)   THEN add (point(obj2)), add (top(obj2,obj1)))

An example of the rule for drawing diagram is shown in Table 2. This example is a rule for deriving the top of parabola. If the internal model includes parabola, the system adds the points to the top of parabola.

4   Prototype System

Currently, the prototype system is implemented in the stand-alone learning environment. This system is mainly implemented using Java and, figures and relations are managed as independent objects. Prolog is used to select and apply the rule for drawing diagrams.

The interface of the prototype system is shown in Figure 3. The student draws a diagram by pushing buttons for drawing figures and buttons for drawing relations, and by clicking the particular coordinates on the canvas of drawing diagrams. Moreover, the student can delete the object easily by choosing the target object in the list. When the figure is added, the input figure is drawn in a different color in order to highlight it. After "next button" is pushed, the diagram is transformed into the internal model. Then, appropriate rules are selected and a new diagram with the supplementary figures is displayed as the advice. If "cansel button" is pushed, another rule which fits to the current diagram is applied and another diagram is generated.
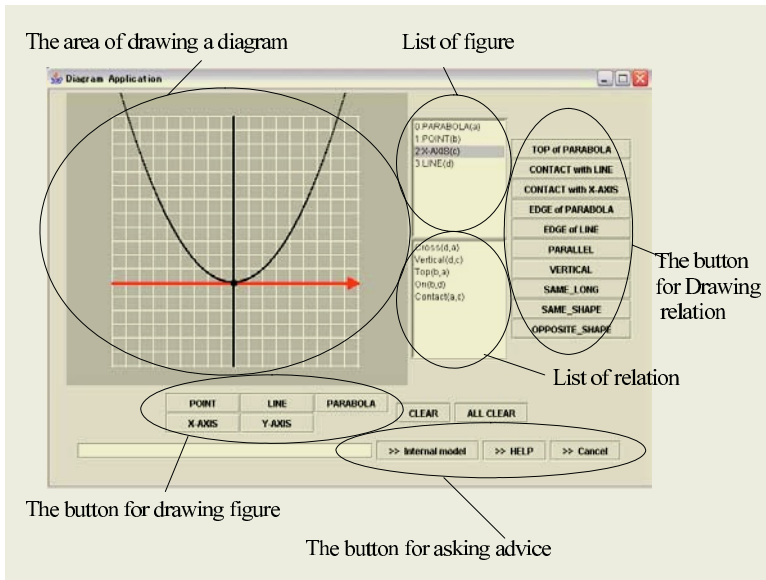
Fig. 3. Interface of prototype system
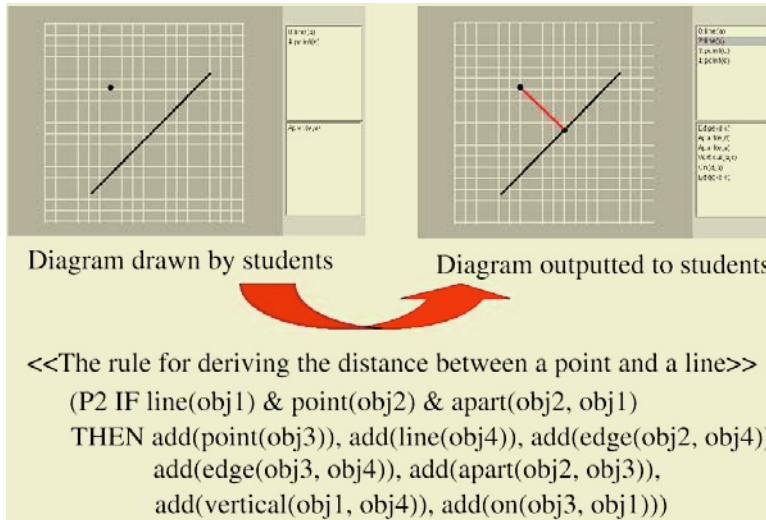
Fig. 4. Example of applying rule for drawing diagrams

If plural rules match with the current internal model, the rule which has the biggest number of predicates in the precondition part is selected. Figure 4 shows the example of generating diagrams. The left side in Figure 4 is a diagram which the student draws, and the right shows the diagram generated after the rule for drawing diagrams has been applied. Based on this internal model, the line which connects to a point and a line and crosses over a line perpendicularly is added to the diagram as a supplementary figure.

We evaluate our system by analyzing the prepared rule for drawing diagrams. Currently, we define 9 kinds of rules for drawing diagrams for high school-level mathematics exercise in two-dimensional function. Figure 5 shows the relation of rules for drawing diagrams according to their preconditions. Each node corresponds to the rule and the link indicates the inclusive relation. That is, the precondition of the rule in an upper level contains the precondition of the rule in a lower level. Preconditions of type 5 and 6 are different from others. On the other hand, types 1, 2, 3, 4, 7, 8, and 9 are highly related to other rules. So, the lower rules are difficult to be selected even if it was appropriate to the current learning situation. To avoid applying the inappropriate rules for drawing diagram, it is necessary to select the applicable rules for drawing diagram according to the types of exercises and control the order to apply.
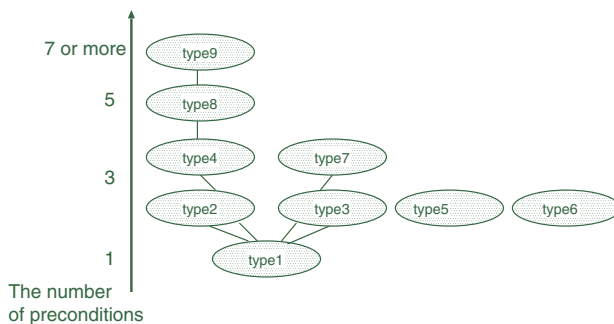
Fig. 5. Inclusive relation of rules according to their precondition

5   Conclusion and Future Work

In this paper, we proposed the mechanism of generating diagrams as an advice for the mathematical exercise in two-dimensional functions. In our approach, the diagram is generated by applying the rule corresponding to each derivation method to the internal model of the current diagram. By generating the diagram based on the forward reasoning, various kinds of diagrams from various viewpoints would be expected. Currently, we have constructed the prototype system that is executed in a stand-alone learning environment.

For our future work, we must introduce the prototype system into the collaborative learning environment and evaluate whether it can promotes the active discussions of various viewpoints. In the collaborative learning, diagrams are not drawn stepwisely. Therefore, the history of applying the rules for drawing diagrams should be managed and rules are needed to be selected in considering the history. Moreover, currently we focus on only mathematical exercises in two-dimensional function. In order to generate the advice that makes learners induce various viewpoints, we need to extend the target area, such as three-dimensional function, trigonometric function and so on. Of course, the internal model and rules for drawing diagrams should be extended according to the target area.

References

1. M. Nakamura and S. Otsuki: Group Learning Environment Based on Hypothesis Generation and Inference Externalization, Proc. of ICCE'98, Vol.2, (1998) pp.535-538.
2. T. Supnithi, A. Inaba, M. Ikeda, J. Toyoda and R. Mizoguchi: Learning Goal Ontology Supported by LearningTheories for Opportunistic Group Formation, Proc. of AIED'99, (1999) pp.67-74.
3. T. Kojiri and T. Watanabe: Diagrammatic Advices in Mathematical Exercises, Proc. of ICCE'03, (2003) pp.293-296.
4. T. Watanabe and T. Kojiri: Architecture of Collaborative Learning Environment as Infrastructure for Personal Activity, Proc. of CATE'03, (2003) pp.420-425.
5. T. Ito, N. Ohnishi and N. Sugie: A Cognitive Model of Diagram Based Geometric Problem Solving, Transactions on IEICE, Vol.24, No.11, (1993) pp.84-97.
6. T. Ito, N. Ohnishi and N. Sugie: Problem-Solving SCRIPT and Classification of Drawings for Explaining Human Drawing Processes, Transactions on IEICE, Vol.26, No.5, (1995) pp.29-42.

MOLEAS: Information Technology-Based Educational Software Framework*

Su-Jin Cho[1] and Seongsoo Lee[2]

[1] Department of Korean Language Education, Seoul National University, 151-742, Korea
chosoojin@chollian.net
[2] School of Electronics Engineering, Soongsil University, 156-743, Korea
sslee@ssu.ac.kr

Abstract. Recently, the advent of Internet introduced E-learning, although there are still many practical problems to be solved. This paper presents an educational software framework based on the information technologies called MOLEAS (MOdular Learning Engines with Active Search). The software framework is modularized with standardized interface based on IEEE 1484 LTSA (Learning Technology Standard Architecture) so that the user can reuse other users' modules to save time and effort. Using peer-to-peer technology, a user can search other users with same educational interests over Internet.

1 Introduction

Recently, Internet and multimedia gave a tremendous impact all over the world, and changed our daily life in every aspect. One of such changes is E-learning. It is regarded as an effective method to solve various problems of conventional education. Along with the progress of information technologies, Internet infrastructure and its users grow rapidly. However, E-learning does not play an important role in the Internet applications. For example, high-speed (>1Mbps) Internet users in Korea are 21.7 million, covering 50% of whole population [1]. Nevertheless, only 7% of them use Internet for educational purpose, while 67% for computer games. This comes from the fact that conventional E-learning has many practical problems to be solved. In this paper, a novel software framework called MOdular Learning Engines with Active Search (MOLEAS) is proposed to avoid these problems.

2 Information Technologies for e-Learning

In this section, emerging information technologies are briefly described from the viewpoint of E-learning.

Intelligent Software Agent [2]: One serious problem in E-learning is information flooding, i.e. how to find proper educational materials from the Internet full of almost infinite number of various materials. For example, suppose a learner is interested in the history of nuclear weapon and wants to search appropriate educational materials from Internet. However, Yahoo, one of the most famous search engines in the world, reports about 100 web sites and about 17,000 web pages with keywords of "history", "nuclear weapon", and "education".

Intelligent software agent is one solution for such a problem. By definition, software agent is autonomous software, interacting with its environment and with other agents reactively or proactively in order to achieve its own goals [3]. In the multi-agent technologies, many agents cooperate with each other using the messages communicable among themselves. This technology is very effective in practical intelligent tutoring system (ITS) [4]. Based on user's interests and preferences, it navigates Internet space, contacts search engines, retrieves/analyzes/refines search results, and selects proper contents. In addition, it automatically recommends best materials for each learner by monitoring log files and analyzing learning inclination, which enables personalized instruction.

Peer-to-Peer [5]: In the Internet environment, most information exchange is performed by client-server architecture, where information is transferred from client to server, and then server to another client. Most information exchange services such as email hosts, web sites, file transfer protocol (FTP) consist of a server with powerful host machine and many clients with small personal computers. However, this approach suffers from following problems:

- Expensive server cost: The server cost increases exponentially with the amount of information to be stored or processed.
- Access bottleneck: Too many concurrent accesses of same information results in access failure or server down.
- Global service failure: When the server goes out of order, whole services stop for all users.

As high-speed Internet connection becomes popular, communication without server becomes possible between individual users. Peer-to-peer (P2P) is emerging as an alternative to client-server architecture, which is summarized as follows:

- Users register their materials to their network neighbors.
- A user search across the network to find right materials he wants.
- After he gets the location, he directly gets the materials without any servers.

This technology is widely used in many fields such as cycle sharing (SETI@home [6]) and file sharing (Napster [7]). In E-learning, P2P is effective in many ways. For example, suppose a learner wants to find another learner who is interested in the history of the Holy Roma Empire. With P2P, he can easily find many learners with common interest and perform collaborative learning with them. He can also find many experts on the Holy Roma Empire and get proper advices from them.

MPEG-7 [8]: Currently, most Internet search engines are text-based, i.e. they provide search results primarily by matching keywords. This lead to serious problems in multimedia contents, although it works pretty well for documents, books, and databases. Many multimedia contents are just audiovisual objects without any text descriptions. For example, assume that somebody uploaded home video of his family in his personal homepage without any text descriptions. In such a case, how can we find it over Internet? Furthermore, it is difficult to label proper text keywords fully describing given multimedia contents. In the above example, what are appropriate keywords if that video is just "several people enjoying dinner around large table"?

Although multimedia contents are described in detailed text keywords, it sometimes happens that somebody wants to find out appropriate contents based on impre-

cise and abstract descriptions. For example, suppose a teacher develops educational materials on antiwar movement. He wants to find out some audiovisual contents such as "documentary films of soldiers killing unarmed civilians, the more victims the better". How can he find it with conventional search engines?

One solution is MPEG-7, formally named as "Multimedia Content Description Interface". It is an international standard for describing the features of multimedia contents so that users can search, browse, and retrieve those contents more efficiently than they could using today's text-based search engines. Because the descriptive features must be meaningful in the context of the application, they will be different for different user domains and different applications. This implies that the same material can be described using different types of features, tuned to the area of application. For example, MPEG-7 provides many description levels as follows:

- Video object description: shape, size, color, movement, spatial/temporal position…
- Audio object description: key, mood, tempo, change of tempo, temporal position…
- Structural description: scene cuts, segmentation in regions, motion tracking…
- Conceptual description: objects and events, interactions among objects…
- Semantic description: "this is a scene where some spaceships are attacking a big black artificial planet, with the sound of terrible blast…"

One important feature of MPEG-7 is that it enables human-like queries in multimedia database search. Some query examples are:

- Typing some paragraphs of "King Lear" on a keyboard, retrieve some literary works matching in terms of emotions.
- With an Olympic logo, obtain a list of cities where Olympic Games were held.

MPEG-7 is a power tool for E-learning in many ways. For example, suppose a teacher develops multimedia materials for communicative conversation skills. He wants to show various sets of conversations to his learners. Using MPEG-7 search engines, he can find many film clips of polite, friendly, or enthusiastic conversation between businessmen, lovers, and family members out of many famous movies.

3   MOLEAS: An Educational Software Framework

This paper proposes MOdular Learning Engines with Active Search (MOLEAS), a novel software framework for E-learning over Internet environment. It has the following features:

- Standardized architecture based on IEEE 1484.1 Learning Technology Standard Architecture (LTSA) [9]
- Flexible architecture with module-based learning engine
- Distributed architecture over Internet with P2P and intelligent software agent
- Reconfigurable architecture covering various E-learning models
- Learning inclination analyzer, enabling personalized instruction
- Various communication tools between teachers and learners such as email, BBS, chatting, and instant messenger
- Powerful authoring tools with MPEG-7 multimedia search engine
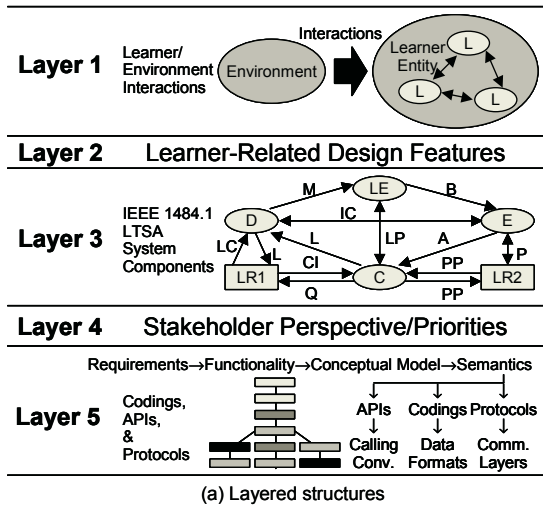
(a) Layered structures
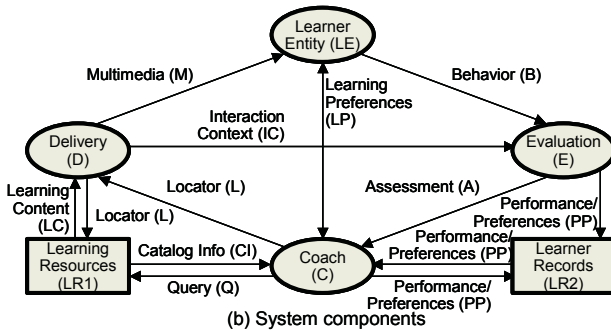
(b) System components

Fig. 1. IEEE 1484.1 learning technology system architecture

LTSA is an international standard for information technology-supported learning, education, and training systems. It has five refinement layers as shown in Fig. 1 (a), but only layer 3 is specified in the standard, including 16 system components as shown in Fig. 1 (b) - four processes (learner entity, evaluation, coach, delivery), two storages (learning resources, learner records), and ten flows (learning preferences, assessment, performance/preferences, query, catalog info, locator, learning content, multimedia, interaction context). The hardware platforms and the software programs are implemented as layer 5, which corresponds to the proposed software framework.

As shown in Fig. 2 (a), MOLEAS consists of five basic modules: learning module, teaching module, content module, search module, and control module. Each module corresponds to several LTSA system components, as shown in Fig. 2 (b). Each module has standardized LTSA interface scheme so that any two modules can communicate and combine with each other. Some modules locate in the users' personal computers, some in the web sites, and some in the Internet.

By reconfiguring connection status, this module-based approach enables flexible system architecture for various education models. For example, in Fig. 3 (a), all modules locate in a CD-ROM, implementing stand-alone educational software. Fig. 3 (b)

(a) MOLEAS modules
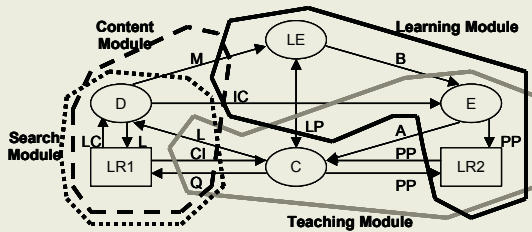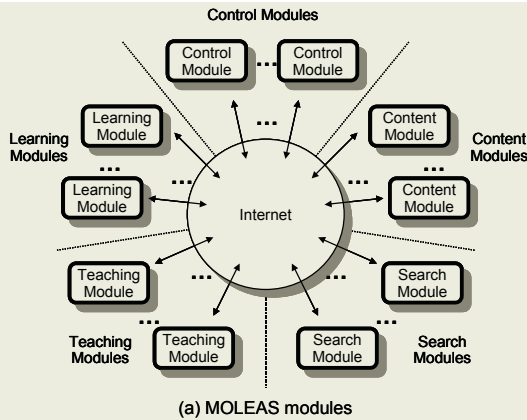
	Processes				Stores		Flows									
	LE	C	E	D	LR1	LR2	M	LC	L	CI	Q	IC	LP	B	A	PP
Learning Module	√		√		√							√	√	√		√
Teaching Module		√	√		√			√	√	√		√		√	√	√
Content Module				√	√		√	√	√	√	√					
Search Module				√	√			√	√	√	√					
Control Module	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√

(b) MOLEAS modules and LTSA system components

Fig. 2. Basic modules in MOLEAS

is an example of an educational web site, where the web site has many modules acting as a server and each learner/teacher has only learning/teaching module acting as a client. Search module in the web site finds proper contents or teachers. Note that it serves as conventional distance learning through "teacher → web site → learner".

Furthermore, when learners register their own interests and preferences, intelligent software agents automatically finds other learners with common interest by exploiting peer-to-peer. Once they are found, he can access and utilize their modules to compose an effective learning engine. In this case, the learning engine is not stand-alone software on his personal computer, but distributed software over Internet. Fig. 3 (c) shows an example for collaborative learning among learners.

MOLEAS has many powerful tools to support personalized instruction. Learning inclination analyzer is an expert system combined with intelligent software agent. It analyzes log files and behaviors of learner's learning modules, and recommends best
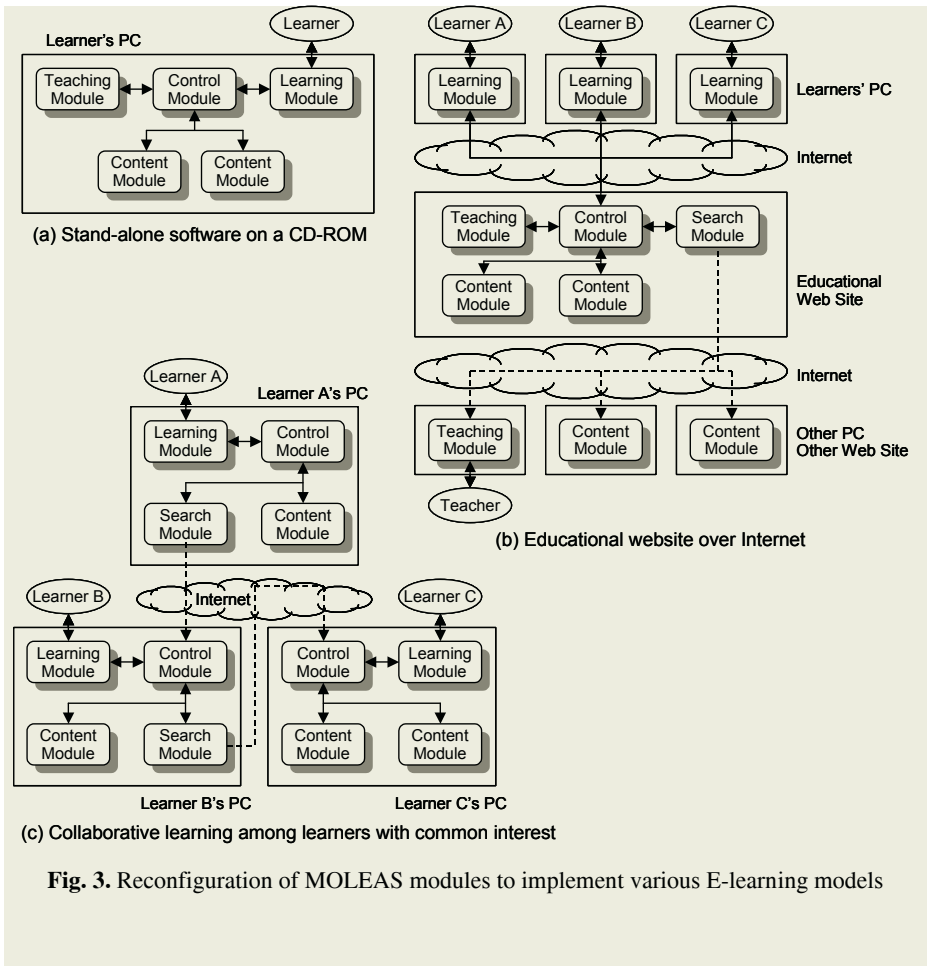
Fig. 3. Reconfiguration of MOLEAS modules to implement various E-learning models

materials to both teachers and learners. For example, it recommends detailed and profound educational materials if a learner studies for a long time on one subject, while it recommends short but various materials if he connects frequently. MOLEAS also provides various communication tools between users so that teachers can easily teach, guide, and encourage learners. It has powerful authoring tools with MPEG-7 multimedia search engine so that teachers can easily develop high-quality educational contents. The proposed learning engine has following advantages:

- A user can easily find other learners and teachers with common interest or proper educational materials he really needs.
- It can be easily applied to various education models with flexibility and adaptability, since it only needs to change connection status between modules.
- A user can easily drag and add various modules of other users to his learning engine, meaning that it saves a lot of time and money by reusing them.
- With powerful built-in tools, it can be applied to various fields of E-learning including distance learning, personalized instruction, and collaborative learning.

4   Conclusion

In this paper, MOLEAS, a novel educational software framework is proposed based on the emerging information technologies. It has five modules: learning module, teaching module, content module, search module, and control module. It has standardized interface based on IEEE 1484 LTSA so that it can be easily reconfigured for various educational applications. By exploiting peer-to-peer technology, a user can easily find other learners and teachers with common interest or proper educational materials he really needs. Once they are found, he can easily drag and add various modules of other users to his learning engine to save time and effort.

References

1. Korea National Computerization Agency: Korea Internet White Paper 2002, http://www. nca.or.kr/main/notice/2002_11/2002whitepaper.pdf
2. Shoham, Y.: Agent-Oriented Programming, Artificial Intelligence **1** (1993) 51-92
3. Sleeman, D., Brown, J.S.: Intelligent Tutoring System, Academic Press (1982)
4. Wooldridge, M., Jennings, N.R.: Intelligent Agent: Theory and Practice, The Knowledge Engineering Review **2** (1995) 115-152
5. Awduche, D.O., Agogbua, J., McManus, J.: An Approach to Optimal Peering between Autonomous Systems in the Internet, International Conference on Computer Communications and Networks (1998) 346 –351
6. http://setiathome.ssl.berkeley.edu/
7. http://www.napster.com/
8. ISO/IEC JTC1/SC29/WG1 15938: Multimedia Content Description Interface, http://www. cselt.it/mpeg/standards/mpeg-7/mpeg-7.zip
9. IEEE 1484.1 Standard for Learning Technology - Learning Technology System Architecture, http://ltsc.ieee.org/wg1/files/IEEE_1484_01_D09_LTSA.pdf

A Web-Based Information Communication Ethics Education System for the Gifted Elementary School Students in Computer

Woochun Jun[1] and Sung-Keun Cho[2]

[1] Dept. of Computer Education, Seoul National University of Education, Seoul, Korea
wocjun@snue.ac.kr
[2] Seoul Ujang Elementary School, Seoul, Korea
vision1970@hanmail.net

Abstract. The development of information communication technologies has brought great changes to our society. These advancements enable people to work more efficiently and make daily life more convenient. However, the pervasive application of information communication technologies also has some side effects such as hacking and computer game addiction. While current school education on computers and communication for the gifted students has mainly focused on knowledge and skills on computer, ethical issues on computer knowledge and technologies have not been dealt with sufficiently. In this paper, we introduce a Web-based information communication ethics education system for the gifted elementary school students in computer. The proposed system has the following characteristics: First, extensive contents for the gifted elementary school students are organized. Second, the system adopts various types of avatars for delivering hard contents in an interesting manner. Third, the system supports various types of interactions by providing interactive environments. Finally, the proposed system provides online survey to check students' progress in information communication ethics.

1 Introduction

Advances in computer and communication technologies have rapidly changed our life in various ways so that people are able to communicate on a global scale and assess vast amounts of information. However, along with positive effects, we have to deal with side effects such as hacking and game addiction. In order to reduce the harmful effects raised in current information-oriented society, people used to set up strict rules or develop programs as a part of technical steps. However, those actions taken have turned out to be effective within a boundary. Experts are now moving into developing educational countermeasures against such harmful effects. They suggest that people can eradicate harmful effects and casualties by instructing students the essential information communication ethics. Therefore, the information communication ethics education is considered to be the fundamental way of rooting up serious side effects [3].

For elementary school students, information communication ethics education is more important since they usually do not have firm morality. For those students,

teachers must not to force them to learn the information communication ethics by cramming education. Instead, teachers need to help the elementary school students realize the proper code of conduct on the Web, and then internalize it naturally.

For the gifted students, teaching information communication ethics is more important since they usually have higher sensibility and bigger knowledge and skills in computer than ordinary students [6]. Thus, it is necessary to select appropriate contents and teach those contents at their early stage. In this paper, a Web-based information communication ethics education system for the gifted elementary school students in computer is introduced.

This paper is organized as follows. In Section 2, the basic concepts and principles of information communication ethics are presented. Also, definition and characteristics of gifted students in computer are discussed. In Section 3, the design and implementation of the proposed system is explained. Finally conclusions and further research issues are provided in Section 4.

2 Backgrounds

2.1 Definition and Principles of Information Communication Ethics

The information communication ethics education is developed for establishing healthy information-oriented society utilizing information communication technology with justifiable ethic code by bringing forward opportunities for people to learn the norm they should respect and observe. In brief, it's the essential subordinate field belonging to ethology. Also, it has the significant educational value and importance.

In cyberspace, people often make mistakes like hurting another person's feelings due to anonymity. It is, therefore, necessary to learn the right "netiquette," or manners required for cyber space, as it is to learn etiquette in real life. Although a definite standard for netiquette has not yet been established, some basic rules have been introduced as follows [1,2]:

  1) Remember the human
  2) Adhere to the same standards of behavior online that you follow in real life
  3) Know where you are in cyberspace
  4) Respect other people's time and bandwidth
  5) Make yourself look good online
  6) Share expert knowledge
  7) Help keep flame wars under control
  8) Respect other people's privacy
  9) Don't abuse your power
  10) Be forgiving of other people's mistakes

2.2 Definition and Characteristics of the Gifted Students in Computer

There have been some works for defining gifted student in computer [4,5,7,8,9,10]. Based on those definitions, we define the gifted student as follows. First, gifted students have above-average ability in general intelligence, strong curiosity in computer, high creativeness, high mathematical and linguistic talent and tenacity. Second, gifted

students are interested in application software, programming, game and multimedia and excellent in intuitiveness, generalization, inference, and adaptability for computer. Third, gifted students have talent in expressing ideas using computer and applying computers to other real-world problems.

In [5], conditions for gifted students in computer are defined as in Table 1.

Table 1. Conditions for the gifted student in computer

Area	Contents
Comparative predominance	-Excellence in vocabulary use and linguistic expression over the same age -Possession of above-average ability for mathematics and linguistics -Spending time and efforts on reading
Investigation	-Strong curiosity and high scholastic accomplishment for computer-related subjects -Strong will to accomplish for a specific subject -Keen observance and good memory
Analysis and planning	-Ability to grasp and generalize facts and relationships among elements -Ability to solve problems with efficient methods
Applicability	Excellence in applying computer knowledge to new situation
Mental state	Possession of infinite imaginative power, applicability and initiative power
Expression	Excellence in expressing new ideas and creative contents with computer

On the other hand, characteristics of gifted students in computer are presented in [7].

Table 2. Characteristics of the gifted Students in computer

Area	Characteristics
General characteristics	-Excellence in understanding and manipulating things -Quick acquisition of basic functions -Right and quick decisive power -High curiosity -Enthusiasm about new thinking and challenge
Application software	-High imaginative power and applicability -Ability to grasp relationships -Ability to set up hypothesis and conjecture
Programming	-Excellence in grasping and understanding main principles -Insight to cause and effect -Enjoying new way of thinking and method
Multimedia	-Infinite imaginative power -Excellent artic sense -Composed and delicate -Excellent creative activity -Ability to observe things sharply
Digital Contents	-Tenacity -Infinite imaginative power and applicability -Desire to win a game -Desire to have control -Desire to show off -Resolute decisive power

3   Design and Implementation of a Web-Based Information Communication Ethics Education System for the Gifted Elementary School Students in Computer

3.1   Design of the Proposed System

Various Curricula, materials and contents have been investigated to provide contents for gifted elementary school students in computer. Since the gifted students are likely to learn advanced techniques quicker than ordinary students, we emphasize on contents leading the gifted students to understand dysfunction of computer and harmfulness of various types of addictions, resulted from misuse of computer. The contents can be classified into 4 subjects, Information communication ethics Pros and cons of cyber world, Netiquette and Cyber symptoms, respectively. Table 3 shows those arrears and their corresponding subjects.

Table 3. Areas and subjects of contents

Area	Subject
Information communication ethics	-Meaning of information communication ethics -Ethics code of netizen -Terminologies of information communication -Chatting language
Pros and cons of cyber world	-Pros and cons of Internet -Illegal copy -Computer Virus -Hacking -Cyber crime
Netiquette	-Meaning of netiquette -Rules of netiquette -Observance of netiquette -Spam mail
Cyber symptoms	-Game addiction -Internet addiction -Unsound site -Chatting addiction

The proposed system has the following characteristics. First, the proposed system provides comprehensive contents of information communication ethics. Second, various types of avatars are used to explain contents and guide students with voice message. Third, for each subject, there are three steps, Introduction, Development and Discussion to follow. Students identify ethical issues introduced by Flash video (Introduction stage), grasp and try to solve problems based video data (Development stage) and discuss and exchange their ideas for themselves (Discussion stage). The purpose of these steps is to let students learn by heart rather than cramming. Fourth, the system supports various types of interactions by providing interactive environments such as Q & A, FAQ on BBS.

3.2   Implementation of the Proposed System

3.2.1   Hardware and Software Environment

Our system has been implemented using ordinary hardware and software. Table 4 shows our hardware and software environment.

Table 4. Hardware and software environment of the proposed system

	Element	Option	
	CPU	P-III 733 MHz	
Hardware	RAM	256 MB	
	HDD	27GB	
	Operating system	Linux	
	Web server	Apache	
	Database	Mysql	
Software	Web browser	MS Explorer 6.0	
	Programming	Contents production	Flash Mx
	language	BBS production	PHP 4.0

Web site address of our system is http://comedu.snue.ac.kr/~m20011513.

3.2.2 Application of the Proposed System
Figure 1 shows the overall menu structure of the proposed system.
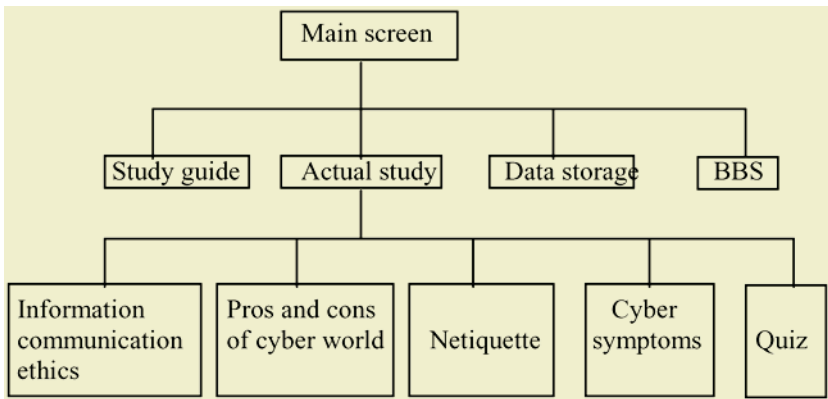
Fig. 1. Overall menu structure of the proposed system

Figure 2 shows overall flow diagram of study procedures for the gifted students.

4 Conclusions and Further Work

In this paper we introduce a Web-based information communication ethics education system for the gifted elementary school students in computer. First, the proposed system provides comprehensive contents of information communication ethics. Second, various types of avatars are used to explain contents and guide students with voice message. Third, for each subject, there are three steps, Introduction, Development and Discussion to follow. Students identify ethical issues introduced by Flash video (Introduction stage), grasp and try to solve problems based video data (Development stage) and discuss and exchange their ideas for themselves (Discussion stage). The purpose of these steps is to let students learn by heart rather than cramming. Fourth, the system supports various types of interactions by providing interactive environments such as Q & A, FAQ on BBS.

Fig. 2. Overall flow diagram of study procedure for students

We plan to evaluate the effectiveness of the proposed system in diverse manners such as survey and tests. We also plan to extend the proposed system to the gifted elementary school students in other areas.

References

1. http://www.albion.com/netiquette/corereuls.html
2. How will instructor teach the cyber ethics?, Seoul Metropolitan Office of Education, 2001.
3. http://www.kedi.re.kr/Exec/Pds/Cnt/128-16.htm
4. Korea Education and Research Information Service, Ministry of Education & Human Resources Development, "Study on Preventing the Reverse Effects of Information-oriented Educational Institution", 2000.
5. K. Lee and W. Jun, "A Creativity Development Study for the Gifted Elementary School Students in Computer", Proceedings of the 8th Korean Association of Information Education Winter Conference, pp. 404-412, 2003.
6. D. Na, "Development of a Curriculum of Education for the Elementary Gifted Children of Information Science, Elementary Computer Education major, Graduate School, Inchon National University of Education, 2003.
7. S. Oh, "Definition and Recognition System for Gifted Students in Computer", Master Thesis, Computer Education Major, Graduate School of Education, Sungkyunkwan University, Korea, 2002.
8. K. Ryu and J. Lee, "A Study on Information & Communication Ethics Education for the Gifted Elementary School Students in Information Science", Proceedings of the 8th Korean Association of Information Education Summer Conference, pp. 199-206, 2003.
9. R. J. Severson, The principles of information ethics. ME Sharpe, 1997.
10. S. Shin, "A Study on Development of a Program for the Identification of Information Gifted Children based on the Ability of a Discrete Thinking", Master Thesis, Elementary Computer Education major, Graduate School, Korea National University of Education, 2004.

Design of a Web Based
Lab Administration System

Sujan Pradhan and Hongen Lu

Department of Computer Science and Computer Engineering, La Trobe University
Bundoora, Melbourne, VIC 3086, Australia
{sujan,helu}@cs.latrobe.edu.au

Abstract. Web-based learning, be it in higher education, corporate, government or health care sector, has been rapidly developing over the past several years. The current manual system of allocating labs to students is not only time consuming for lecturers and head tutors but is also inefficient and inconvenient. In this paper, we present an evaluation on the design of a web based lab administration system; the system provides a flexible and comprehensive tool for e-learning. The software development life cycle is described, as well as functions which are innovative and unique in this system.

1 Introduction

Web-based learning, be it in higher education, corporate, government or health care sector, has been rapidly developing over the past several years. One of the prominent advantages of e-learning is the flexibility of delivery. Asynchronous learning, such as this, allows the learners not only to work at their own set pace but from virtually any location they choose. Of these, higher education has been the most notable in setting the trend. One of the challenges has been in providing more practical approaches to e-learning. Such approaches need proper refinement and delivery during design, especially considering the convergence of these different teaching methods, which can add further complexity. The focus of this paper is in evaluating a design of a web based lab administrative system, a type of web-based learning, with a knack for software agents.

Many research works have been conducted in the field of web-based learning, such as [3] and [4], which apply agent technology in online collaborative learning. On the other hand, there is not any notable commercial software which solely target laboratory administration; however, there are a few commercially available software which offer similar capability of our Automated Lab Information Administrative System (ALIAS). WebCT [6], for example, offers many integrated features in web-based learning. It targets the higher education sector in delivering e-learning by letting lecturers design a subject's structure online. Yet another relevant example in e-learning is MIT (Massachusetts Institute of Technology) OpenCourseWare [5] project introduced in 1999. Its initiative to make e-learning a thing of the past has given a new meaning to education-"knowledge and ideas are shared openly and freely for the benefit of all" [5].

2   System Analysis and Design

2.1   Domain Analysis

The current manual system of allocating labs to students is not only time con-
suming for lecturers and head tutors but is also inefficient, inconvenient, causes
lapse in communication, and increase in human resources. The whole process
begins with a design of a lab signup sheet made up of preference features (i.e.
lab day, lab time) by the staff. They are distributed during lecture where the
students indicate which labs they would like to attend in order of preference,
usually from 1-7. Once all of the sheets have been turned in these preferences
are then manually sorted according to the student's preference(s) by the staff.
The staff may opt to use a spreadsheet tool or sort them manually on paper.
The final allocation is posted in a bulletin board visible to students.
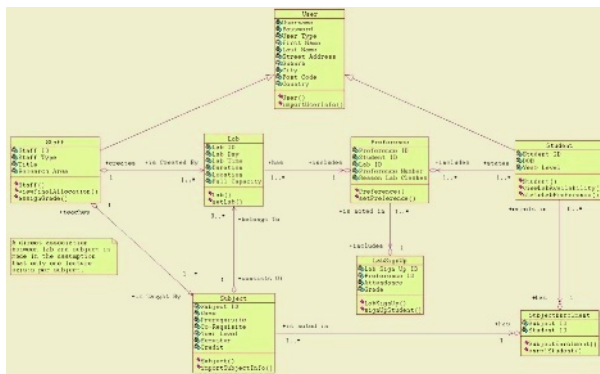
Fig. 1. Class Diagram

Using a manual system such as the one above has drawbacks. The first notable
point that comes to mind is the increase in error. For example, lab clashes may
not be avoided if a chosen lab times are not checked for each of the student's
subject during lab signups; this checking could save student time and hassle
later. The other point is a waste of valuable time; the tasks become repetitive
and administrative for the staff. Moreover, the students are obliged to fill out
the preference sheets if they wish to be allocated; this causes unnecessary hassles
if the student was absent on the day the signups. Depending on the number
of students enrolled in a subject the role of assigning labs can become quite
challenging if there are conflicts such as those who are not able to attend certain
labs due to unknown restrictions. These problematic reasons have a given a
rise to the creation of ALIAS, where a reduction or a complete elimination in
administrative tasks is the key. The core of the high level domain class diagram,
Figure 1, includes the following classes: Student, Lab, Preference and LabSignUp.
These four classes represent the main thrust of the lab allocation process. The
identified classes and their relationships are shown in a domain class diagram at
an analysis level.

2.2   Data Analysis

ALIAS will use a relational database for data storage. However, not all the data will be stored in the data storage as some data are referenced to/from another system, such as the university Lightweight Directory Access Protocol (LDAP) system. Most database communication will take place through dynamic SQL and stored procedures. The use of stored procedures will make some functions reusable (compared to using queries directly from the server side program modules) and will also simplify maintenance.

2.3   UI Analysis

The screenshot, Figure 2, is an example of a look and feel of the menus in the system. Each subject has its own menus which are more relevant to a given situation. For example the menus may be different for a student than for a staff given the same subject. The figure captures a screen for a staff once labs have been set up.

Fig. 2. Menu- screen after setting up lab

2.4   Systems Design Description

ALIAS will be accessible to students and staff alike from anywhere with an internet connection. Students will be prompted to log in to ALIAS with the same username and password they use for other study hall accounts at La Trobe University. Once logged in, they will be shown their current registered classes; they will have an option to click on the classes to view lab availability for each class. They will also be given an option to view lab availability, which corresponds to classes they have not yet registered. Students will be allowed to sign up for

labs even though they have not yet officially enrolled in a subject, as the staff is obliged to sign them up per Department policy. The lab administrative system will have four consecutive modes each with time deadlines: preference, allocation adjustment, final, and post-mortem mode.

During preference mode, students are allowed to state their preferences for each subject they are enrolled in or will be enrolling in the near future. They will also be able to view, adjust, or delete their preferences during this mode. The system has intelligence built-in; a software agent will coherently check for any lab clashes and display only the labs which are available to the student. If students are not able to attend any of the given lab choices options will be provided to state the reasons for the time conflict. These students will be flagged as unallocated students. Since there is a great deal of autonomy within staff, an option will be provided during the post-mortem mode for them to allocate these students manually to any given lab which may have spots remaining.

During allocation adjustment mode, students are allowed to view their allocated labs once the allocation has occurred.Currently, time will be used as a trigger to allocate students, The time, in this case, refers to the expiry date for signups. Students may also be allowed to adjust their allocation provided that there are still spots remaining. If students who are not placed in labs for any reason (i.e. the labs become full, unavailability of labs) they will be automatically flagged as unallocated students.

During final mode, students are able to view the final student allocation. During this mode, lab adjustment by the students is not allowed. An algorithm heavily based on recursive formulation [1] is used for allocation (i.e. students with first preference are allocated until labs are full; remaining first preference students are allocated to second preference labs and so on until all the students have been allocated). The staff will be able to manually assign those students, who were unallocated during the allocation adjustment mode, to an appropriate lab if they choose to do so.

During post-mortem mode, any lab related issues, such as assigning of grades or recording lab attendance by head tutors can be configured per each need. Staff will have access to view their corresponding labs either before or after the final allocation is complete. They should be able to view all the students per each lab allocation for their own subject. They can also adjust the final list as per their own requirement, as well as allocating those students who were flagged as unallocated. A facility will be provided to export the lab allocation list to a spreadsheet. Staff will also be able to send individual or group e-mails relating to any lab work to corresponding lab lists if desired.

It should be noted that before students are able to state their preferences during the preference mode above, either a lecturer or a head tutor will be able to set up labs. During the set up, they will be given options to define any parameters and constraints (i.e. define the minimum, maximum and optimum size for labs) for each lab. Please note that tutorials are treated in the same manner as labs in the context of this paper.

This web based learning tool is designed to automate the current manual gruelling task of assigning labs to students. It will streamline the whole process, from setting up of labs to final allocation, and conducting post-mortem if required. Furthermore, students will also have a great degree of convenience in selecting and viewing lab preferences as long as they have a connection to the Internet.

3 System Architecture and Implementation

The J2EE design patterns used in ALIAS assists in the way of rapid development as they are highly extensible due to reusability. This is critical when time to market is important. The portability aspect of J2EE is helpful when changing vendors as this does not greatly affect the system already being built. J2EE's many components have built-in support which greatly reduces the problems associated with integration and connections. For example, XML, JDBC, RMI/IIOP and countless other technologies are seamlessly integrated into the J2EE framework.
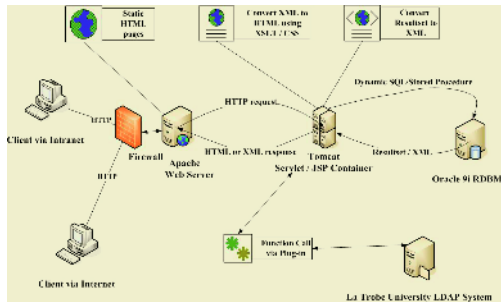
Fig. 3. System Architecture

The language level system architecture, Figure 3, reflects the types of programming languages/techniques used to accomplish a certain architectural goal. The middleware software components consist of the following- web (Apache Web Server) and application server software (Jakarta Tomcat), firewall software, relational database management system software (Oracle 9i), and fault tolerant web server software all deployed under Linux Red Hat operating system. Tomcat will provide support for JSP and servlet, the main language used for this architecture.

A HTTP request is received by servlet/JSP container (Tomcat) via Apache web server, which resides within a firewall. If the request is for a static web page however, Apache will not forward the request. If the request is for authenticating a user (password will be encrypted via HTTPS), a function call from a business tier Java class will be made via a plug-in to the university LDAP system. If the request involves accessing a data store, a dynamic SQL or stored procedure

will be performed to extract the data. The data will be sent in the form of a result set or in XML format by the Oracle 9i RDBMS. The result set data will be converted to XML and wrapped inside an object. Before forwarding the response on to the Apache web server, XML data is then converted to XHTML using XSLT and CSS. It is assumed that all connections are made through HTTP, thus the reason for XHTML conversion. The following will be used for client side GUI: JavaScript, XHTML, CSS, XML and XSLT.

4   Testing

Structural and functionality testing examines the extent to which ALIAS meets the expected functional requirements. They verify that the system performs all the desired functions and conforms to the requirement specifications. Before completion of the project a wide range of normal and erroneous input data would be used to validate each feature of all the behaviours of the system, to ensure that the applications perform efficiently, correctly, and most importantly, bug-free in a real world environment. Different types of testing, white box, black box and usability, were conducted to ensure that the software conforms to the highest of standards.

The survey in Figure 4 is used to identify the level of acceptance from the users' point of view. In analysing the survey, 4th year students seems to have the highest level of satisfaction whereas 1st year students seems to have the least.

Fig. 4. Useability Testing Sample Survey

This result is consistent with their computer skills and programming capacities. Furthermore, it can be concluded from the survey that an overall satisfaction has been reached within the department.

5   Conclusion

This paper describes a tool for automating a lab administrative system in a web-based learning environment. The current manual system consists of too much overhead. It not only takes up extra human resources, but it also comprises of several inconveniences which make it difficult to focus on the primary goal of teaching. The redundancies and discrepancies caused by the system can spill over to a domino effect within the staff in the department. This amounts to avoidable human errors or loss of time due to menial administrative tasks. In meeting the demands of a viable solution to the problem at hand, a tool in web-based learning has been proposed. The tool, ALIAS, solves most, if not all, problems related to the administration of lab allocation. This innovative tool not only enriches student experience through interactive learning similar to [2] but also assists both the students and staff in making the process less time consuming; the end result is higher efficiency and more time for academic learning.

References

1. Bender, E.: Mathematical Methods in Artificial Intelligence, Wiley-IEEE Computer Society Press, 1st Edition, 1996.
2. García, A., Rodríguez, S., Rosales, F., Pedraza, J.L.: Automatic Management of Laboratory Work in Mass Computer Engineering Courses, IEEE Transaction on Education, Vol. 48, No.1, February 2005.
3. Lu, H.: Mediator based Open Multi-Agent Architecture for Web based Learning, In Proceedings of the 2nd International Conference on Web-based Learning (ICWL 2003), Lecture Notes of Computer Science, Springer-Verlag, Melbourne, Australia, August 2003.
4. Lu, H.: Open Multi-Agent Systems for Collaborative Web based Learning, International Journal of Distance Education Technologies, A Publication of Idea Group Publishing, Information Science Publishing, USA, 2004.
5. http://ocw.mit.edu/index.html
6. http://www.webct.com

Experiences with Pair and Tri Programming in a Second Level Course

Maryam Purvis, Martin Purvis, Bastin Tony Roy Savarimuthu,
Mark George, and Stephen Cranefield

Department of Information Science, University of Otago
P O Box 56, Dunedin, New Zealand
{tehrany,mpurvis,tonyr,mgeorge,scranefield}
@infoscience.otago.ac.nz

Abstract. In this paper we describe our experiences with multi-programming (pair and tri programming) in a second level course. The course, "Application Software Development" is a second year course, which has a heavy emphasis on java programming in the labs as well as the development of a full-fledged project. The objective of the course is to build an entire project that comprises of various software engineering activities that span across the semester. In general, we observe that multi-programming improves the students' ability in analytical thinking and communicating the conceptual ideas. It also raises certain issues when this approach is adopted in the educational context. In this paper we discuss some of these issues. Overall, multi-programming experience has been a rewarding experience for the students in spite of certain problems that were encountered.

1 Introduction

The concept of collaborative/cooperative learning [3] has been widely researched and advocated across the literature. The term "collaborative learning" refers to an instruction method in which students at various performance levels work together in small groups towards a common goal. The students are responsible for one another's learning as well as their own. Collaborative learning has proven effective in improving critical thinking of the students [5]. The concept of tri programming is inspired by concepts of collaborative learning and the pair programming [1, 8, 9, 10] activity associated with the Extreme Programming [14]. In tri programming the groups will have three people assigned to work together at a single computer to create software. Similar to pair programming, one person types and takes the initiative for the design of the particular code being developed at that time, while the other two catch minor errors and maintain a more strategic view of how the code fits into the rest of the system. The programmers swap roles as often as they like. All the three programmers should be involved simultaneously in all code development.

The group members will also interact with each other in discussing the design and implementation issues related to the programs that they are developing. In this process, they will reinforce each other's knowledge and work towards a common understanding. They also develop the communication skills that are required for their future careers.

It is believed that both pair and tri programming help both the experienced as well as novice programmers. The expert benefits by practicing his or her ability in explaining a particular idea. Often during this process the expert may identify certain errors in his or her own logic. At the same time the novice benefits by getting exposed to various problem-solving techniques and ideas and also by questioning the reasons behind certain design and implementation decisions.

2   Background

2.1   Motivation for Multi-programming

The motivations for tri programming are a) To provide more opportunities for interactions, b) To provide anonymous evaluation of team members and c) To lessen the likelihood of a dominant person taking over the whole work

2.2   Previous Work

Many researchers have been looking into collaborative programming. The term collaborative programming has been mostly used in the context of pair programming. Our work on multi-programming incorporates teams of three as well as teams of two. Some work has been done on teams comprising of more than two members such as tri programming [7]. Most of the work on group programming concentrates on an advanced programming course in which students tend to be more cooperative and mature, and also it focuses on division of tasks among team members as opposed to simultaneous development of a block of code.

The pair programming practice is associated with the Extreme Programming (XP) methodology. XP promotes a set of practices such as planning game, up-front testing, pair programming, code re-factoring and having on-site customer and so on [14]. Due to the success of XP process in certain applications, some researches have attempted to integrate these basic principles into the educational courses. Some have adopted the whole process [6] while others have tried a subset such as applying just pair programming [5, 11, 12].

The research done by Noble et. al [6], in applying some of the XP principles in their software engineering courses were reported to be successful. They were able to shift the focus from a document-centric approach to a more iterative project-planning and communication-centric approach. Others who have tried primarily the pair programming approach in the educational context [11, 12] have reported an improvement in terms of the quality of the code produced, increases in personal satisfaction and decreases in student drop-outs, in particular, in introductory courses. In addition, it has been reported that the students who participated in pair programming performed as well as students who did solo programming in the final exam [11, 12]. In our experience, we believe that, when participating in pair programming, the process of communicating and rationalizing the decisions made during the project development helps the students in mentally formulating the concepts in a better way and generally these students do better in their final exam. There are many factors involved in the success of the pair programming. Thomas et al. [13], report on the selection criteria. According to Thomas, the pairs with similar abilities and similar personalities tend to

work together the best. Other factors are how often the roles of the pair (driver, observer) should be swapped, how often the pairs should be rotated – (new pairs should be formed), how the work done by the pairs should be evaluated, and so on. In this paper we attempt to address these issues.

2.3   Class Statistics in 2003 and 2004

This paper concerns with the student data and feedback provided by the students for the years 2003 and 2004. The sample sizes were 144 and 110 respectively. In 2003 multi-programming (both pair and tri programming) was done and in 2004 pair programming was undertaken. The students had the option to go to solo programming if a problem such as scheduling was negatively impacting their work. In 2003 there were 29 groups of three, 16 groups of two and 25 solos. In 2004 there were 35 pairs and 40 solo programmers.

2.4   Evaluation of the Group Work

Each group submitted one copy of the project work. The project work contributes to 10% of the total mark for the paper. 80% of the project mark comes from marking the quality of the project submitted (this part of the mark is the same for all the group members); the remaining 20% of the project mark was determined from peer assessment.

2.5   Group Assessment

Assume that the three members in a team are A, B and C. While evaluating a student (say A) if two team members (B & C) are happy with the performance of the other team member (A), then A gets 2 points. If B is happy and C is not happy, then A gets 1 point and if both B and C are not happy about A's performance, then A gets 0 points. The pair programmer's work was assessed only based on what they handed in. This is because the lack of anonymity would put the team members on a difficult position in order to be objective.

3   Issues Related to Multi-programming

In this section, we point out some of the issues that came up while conducting pair and tri programming practices. Most of the results that have been reported are based on our observations and casual interactions with the students during the laboratory work and feedback that we have received from the students throughout our experiment.

3.1   Performance of Multi vs. Solo Programmers

Overall, multi programmers performed at least as well in the projects as the solo programmers. This has been reported in the literature widely [5, 11, 12]. Table 1 shows the average marks scored by multi programmers and solo programmers in two different phases of the project in 2003. It is observed that multi programmers perform better than the solo programmers.

Table 1. Average marks scored by multi and solo programmers in two different phases of the project

	Phase I	Phase II
Multi programmers	88.16	85.94
Solo programmers	74.82	57.5

3.2   Scheduling

Most students suggested that they had scheduling problems when the group size was 3. It is natural that when the size of the team increases it creates problems in reaching a common agreement on problem solving methodologies and also in agreeing to a common schedule. This has also been reported in the XP community [7].

3.3   Formation of Groups Based on Skills

In 2003, the groups were based on the previous performance of the students. The team was chosen in such a way that it comprised of a good, average and below average students. In 2004 the groups were based on voluntary pairing. It has been observed that, though the grouping was arranged based on the skill levels of the students, the students are not enthusiastic to work with team members that they do not know and expressed their reservations in the feedback. However, most of them realised the need for teamwork in real life projects with unknown team members. Still, most students preferred voluntary team formation over the assignment of team members according to their skill levels or previous performances.

3.4   Impact of Peer Pressure

Researchers have reported that students develop a positive peer pressure while programming with a partner [2, 5]. We also believe that peer pressure plays a constructive role in adhering to the deadlines and fulfilling commitments. Some students have mentioned that peer pressure forced them to learn better and prepare well ahead of labs and project meetings. Multi-programming imparts the sense of joint responsibility among the team members. This results in timely presence in project/lab meetings and during collaborative work sessions.

3.5   Performance of International Students on Multi-programming

It has been observed [4] that it is always ambiguous to measure the performance of international students as multi-cultural issues played a major role in their performance and also their written and communication skills. We have observed that, international students, especially Asian students had problems while communicating with their partners in the multi-programming. In a group consisting of only international students, the students often discussed their issues in English and the accent played an important role in the level of understanding between the students. If the group was composed of students belonging to the same nationality, they tended to discuss their ideas in their native language. In a group consisting of mixed nationalities (interna-

tional as well as native English speakers) the international students as well as the native English speakers had difficulty in explaining their views clearly to each other. Some international students felt that this approach resulted in better understanding of the concepts when their group mates explained it in English and also they considered this as a chance to improve their communication skills. On the other hand, most of the native English speakers found it hard to work with the international students, as they had to explain the students over and again till they understood. They considered this to be a waste of time.

3.6   Policies for Swapping the Roles in Multi-programming

Most of the students complained about one person being the driver all the time. It is normal to have totally different personalities in a group. Some students had quoted that some of their teammates had dominating personalities and their own voices were never heard within the team. So policies concerning swapping of roles must be developed.

4   Recommendations to Improve the Process

4.1   Forming Well Balanced Groups/Pairs

To encourage better interaction and cooperation between the team members, the students may select their own partners if they choose to, otherwise an effort should be made to form groups of people with similar skills as well as similar interests (this can be done by asking students to fill a brief set of questionnaire that identifies the students like or dislikes). Similar thoughts have also been expressed by McDowell et.al [11, 12] and Thomas et.al [13].

4.2   Student Assessment

Firstly, in order to improve the team sprit, the evaluation mechanism should be based on the quality of the project as opposed to the performance of an individual. This shifts the focus of students in creating a good quality project and may resolve some of the personal conflicts for the sake of creating good quality software. Secondly, the assessment criteria should be as closely related to the training provided in the course as possible in order to avoid confusion with regard to what is learnt and what is expected in the course. For example, when pair programming is adopted for the practical aspect of the course, the practical test should be also based on the pair programming. There would be other opportunities to evaluate the student understanding of the concepts covered in other course assessment forms, such as final exam and individual lab work that might be designed for learning the basic techniques and principles associated with the course. Thirdly, to apply pair programming in the classrooms, some modification might be needed to the normal process. For example, in the educational setting, we need to ensure that the students have had adequate opportunities to learn the required concepts and techniques - in particular, in a course that is taught at the first or the second year level. These students are still struggling to learn the basic ideas and need to practice on all aspects of the programming including both coding

and code reviewing. In the industrial setting, one assumes that the participants know the basic principles. So, to address this issue and ensure that the students get adequate practice in learning the basic programming technique, we propose the following process.

For a group of 3 people (assume A, B, C denote students with ascending order of capability), choose three problem sets (assume X, Y, Z denotes the problems). The problems have ascending difficulty levels. For the simple problem X, A is the driver and the rest of the team are observers. For the problem Y with medium difficulty, B will be the driver and the other team members act as observers. To solve the difficult problem Z, C plays the role of the driver while the other team members are observers. In the scenario described above, all the students are engaged, and each get a chance to play the roles associated with pair programming (drivers and observers). The problem sets are similar in the basic concepts; they can include some extensions to keep the more capable students interested while providing enough repetitions for the less able student. By this process we have introduced enough redundancy and repetition to ensure that the students have had adequate chance to learn the material. Also, there is a clear structure for the students to swap roles in order to experience the benefits of different roles associated with the pair programming.

4.3 Improving Communication Skills Through Swapping Team Members

In order to improve the communications skills and provide better chances for interacting with different types of students, the project can be partitioned into several sub-tasks. At the end of each task, the students may have the opportunity to swap teams. This way not only the students get to work with new partners, but they get to experience how to understand a piece of code written by another group and how to interface it and maintain it in order to add additional functionalities.

4.4 Motivating Students on Pair Programming

McDowell et. al [11, 12] have discussed the need for motivating students for pair programming. We agree with their viewpoint that students must be motivated and enough background information must be provided so that pair programming will be effective. To prepare and motivate the students for the process associated with the pair programming and the benefits gained, the students should be initially exposed to the appropriate literature in this topic and perhaps be given a chance to write a short document on their understanding of the process. As a part of this process, the students can get engaged in a simple task and outline some of the ground rules to be followed as well as reach an agreement on when and how the roles should be swapped, so that all participants benefit from different aspects of pair and tri programming.

5 Conclusions

As long as there is active participation by all the members of the group, multi-programming will be fruitful activity. Though it has been observed that most three-member groups faced scheduling problems, three people can generate a better discussion on rationalizing some of the programming decisions that are made. Most students

also felt that voluntary pairing should be preferred over non-voluntary pairing. We have also presented certain recommendations to improve the multi-programming method of programming. We believe most students benefit from pair and tri programming. Similar to the benefits of code review in the traditional software development process, pair programming subjects the project to constant code and design review. For this reason it has been reported that the quality of the code is generally improved. In addition, we believe that multi-programming tends to make students reflect more on the concepts and improves the analytical thinking ability of the students. In a way, it helps students to move toward a deeper learning style. It also facilitates knowledge sharing among students and improves the communication and interpersonal skills that are required to work effectively in a team environment. We are planning to formalize and adopt other XP practices in our courses. We have already incorporated some aspects of XP testing practice and in the future we will introduce other practices such as game planning, project iteration, and re-factoring.

References

1. Williams, L., R. R. Kessler, W. Cunningham, and R. Jeffries. "Strengthening the Case for Pair Programming." *IEEE Software* 17.4 (July/Aug. 2000): 19-25.
2. Williams, L., Wiebe, E., Yang, K., Ferzli, M., Miller, C., *In Support of Pair Programming in the Introductory Computer Science Course* Computer Science Education, 2002.
3. Antil, L., J. Jenkins, S. Wayne, and P. Vadasy. "Cooperative Learning: Prevalence, Conceptualizations, and the Relationship between Research and Practice." American educational research journal 35, no.3 (1997): 419-454.
4. George, P. G. (1994). The Effectiveness of Cooperative Learning Strategies in Multicultural University Classrooms. *J. of Excellence in College Teaching*, 5(1), 21-30.
5. Gokhale, A. (1995). Collaborative learning enhances critical thinking. Journal of Technology Education, Vol 7, no 1, Fall 1995
6. Noble, J, Marshall, Stuart, Marshall, Stephen, Biddle, R. Less Extreme Programming. ACE 2004: 217-226
7. TriProgramming: www.c2.com/cgi/wiki?TriProgramming, Accessed on 20th Feb, 2005.
8. Williams, Laurie, Robert R. Kessler, Ward Cunningham, and Ron Jeffries, Strengthening the Case for Pair-Programming, IEEE Software, July/Aug 2000.
9. Williams, Laurie and Kessler, Robert R., *All I Really Need to Know about Pair Programming I Learned In Kindergarten*, Communications of the ACM, May 2000.
10. Williams, Laurie and Kessler, Robert R. The Effects of Pair-Pressure and Pair-Learning on Software Engineering Education. Conference of Software Engg. Edu. and Training, 2000.
11. McDowell, Charlie, Linda Werner, Heather Bullock, and Julian Fernald, The Impact of Pair Programming on Student Performance, Perception, and Persistance, In Proc. of the 25th International Conference on Software Engineering (ICSE 2003), pp. 602 - 607, 2003.
12. Charlie McDowell, Brian Hanks, and Linda Werner, Experimenting with Pair Programming in the Classroom, Proceedings of the 8th Annual Conference on Innovation and Technology in Comp. Sci. Education (ITiCSE 2003), 2003, Thessaloniki, Greece.
13. Thomas, L., M. Ratcliffe, and A. Robertson, Code Warriors and Code-a-Phobes: A Study in Attitude and Pair Programming", Proceedings of SIGCSE 2003, pages 363-367, 2003.
14. Beck, K. & Fowler, M. (2000), Planning Extreme Programming, Addison-Wesley. Beck, K. *Extreme Programming Explained: Embracing Change*. Addison-Wesley, 2000.

An Intelligent Safety Verification
Based on a Paraconsistent Logic Program

Kazumi Nakamatsu[1], Seiki Akama[2], and Jair Minoro Abe[3]

[1] University of Hyogo, Himeji, Japan
`nakamatu@shse.u-hyogo.ac.jp`
[2] Teikyo Heisei University, Chiba, Japan
`akama@thu.ac.jp`
[3] University of Sao Paulo, Paulista University, Sao Paulo, Brazil
`jairabe@uol.com.br`

Abstract. We introduce an intelligent safety verification method based on a paraconsistent logic program EVALPSN with a simple example for brewery pipeline valve control. The safety verification is carried out by paraconsistent logic programming called EVALPSN.

Keywords: EVALPSN (Extended Vector Annotated Logic Program with Strong Negation), safety verification, paraconsistent logic program, intelligent control.

1 Introduction

We have developed an annotated logic program called EVALPSN (Extended Vector Annotated Logic Program with Strong Negation) in order to deal with defeasible deontic reasoning and contradictions [2], and shown that EVALPSN can be applied to some safety verification based control [3, 4]. The basic ideas of the EVALPSN safety verification and control are that the safety properties for control in deontic expression can be easily translated into EVALPSN and the control can be performed by EVALPSN programming.

The safety verification for pipeline control is a crucial issue to avoid unexpected mixture of different kinds of liquid. In fact, for example, nitric acid and caustic soda are used in cleaning process for brewery plant pipelines, then such pipeline valve control is strongly required to avoid the mixture of them, however, formal safety verification methods for such control have not been applied yet. In this paper, we introduce how to apply the EVALPSN safety verification method to pipeline valve control with a simple example.

The outline of the EVALPSN safety verification is as follows : we suppose the safety properties P in deontic expression for the pipelines, which are translated into EVALPSN ; if we want to start a process in the pipelines, the current environment(sensor) information for the process such as valve open-close states is also translated into EVALPSN ; then the safety for the process is verified by EVALPSN programming.
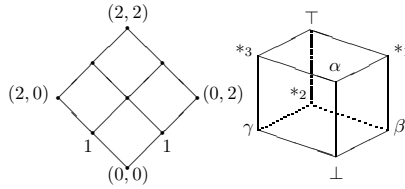
Fig. 1. Lattice $\mathcal{T}_v(n = 2)$ and Lattice $\mathcal{T}_d$

2   EVALPSN

We review the outline of EVALPSN briefly and the details of EVALPSN can be found in [2]. Generally, a truth value called an *annotation* is explicitly attached to each literal in annotated logic programs. For example, let p be a literal, μ an annotation, then $p:\mu$ is called an *annotated literal*. The set of annotations constitutes a complete lattice. An annotation in EVALPSN has a form of $[(i, j), \mu]$, which is called an *extended vector annotation* such that the set of the first components (i, j) constitutes the lattice, $\mathcal{T}_v = \{ (x, y) | 0 \leq x \leq n, 0 \leq y \leq n, x, y$ and n are integers $\}$. The ordering$(\preceq_v)$ of the lattice $\mathcal{T}_v$ is defined as : let $v_1 = (x_1, y_1) \in \mathcal{T}_v$ and $v_2 = (x_2, y_2) \in \mathcal{T}_v$, $v_1 \preceq_v v_2$ iff $x_1 \leq x_2$ and $y_1 \leq y_2$. For each vector annotated literal $p:(i, j)$, the first component i of the vector annotation denotes the amount of positive information to support the literal p and the second one j denotes that of negative one as well, and the set of the second ones constitutes the lattice, $\mathcal{T}_d = \{\perp, \alpha, \beta, \gamma, *_1, *_2, *_3, \top\}$. The ordering$(\preceq_d)$ of the lattice $\mathcal{T}_d$ is described by the Hasse's diagram in **Fig. 1**. Therefore, the complete lattice $\mathcal{T}_e$ of extended vector annotations is defined as the product $\mathcal{T}_v \times \mathcal{T}_d$. The ordering$(\preceq)$ of the lattice $\mathcal{T}_e$ is defined as : let $[(i_1, j_1), \mu_1]$ and $[(i_2, j_2), \mu_2]$ be extended vector annotations, $[(i_1, j_1), \mu_1] \preceq [(i_2, j_2), \mu_2]$ iff $(i_1, j_1) \preceq_v (i_2, j_2)$ and $\mu_1 \preceq_d \mu_2$. The intuitive meaning of each member in the lattice $\mathcal{T}_d$ is $\perp$ (unknown), α (fact), β (obligation), γ (non-obligation), $*_1$ (fact and obligation), $*_2$ (obligation and non-obligation), $*_3$ (fact and non-obligation), and $\top$ (inconsistency). Therefore, EVALPSN can deal with not only inconsistency between usual knowledge but also between permission and forbiddance, obligation and forbiddance, and fact and forbiddance.

There are two kinds of epistemic negation $\neg_1$ and $\neg_2$ in EVALPSN, which are defined as mappings over $\mathcal{T}_v$ and $\mathcal{T}_d$, respectively.

DEFINITION 1 (Epistemic Negations, $\neg_1$ and $\neg_2$)

$$\neg_1([(i, j), \mu]) = [(j, i), \mu], \forall \mu \in \mathcal{T}_d,$$
$$\neg_2([(i, j), \perp]) = [(i, j), \perp], \quad \neg_2([(i, j), \alpha]) = [(i, j), \alpha],$$
$$\neg_2([(i, j), \beta]) = [(i, j), \gamma], \quad \neg_2([(i, j), \gamma]) = [(i, j), \beta],$$
$$\neg_2([(i, j), *_1]) = [(i, j), *_3], \quad \neg_2([(i, j), *_2]) = [(i, j), *_2],$$
$$\neg_2([(i, j), *_3]) = [(i, j), *_1], \quad \neg_2([(i, j), \top]) = [(i, j), \top].$$

These epistemic negations, $\neg_1$ and $\neg_2$, can be eliminated by the above syntactic operation. On the other hand, the ontological negation(strong negation, $\sim$) in EVALPSN can be defined by the epistemic negations, $\neg_1$ or $\neg_2$, and interpreted as classical negation.

DEFINITION 2 (Strong Negation) $\sim F =_{def} F \rightarrow ((F \rightarrow F) \wedge \neg(F \rightarrow F))$, where F be a formula and $\neg$ be $\neg_1$ or $\neg_2$, and $\rightarrow$ is a classical implication. Here we can define EVALPSN formally.

DEFINITION 3 (well extended vector annotated literal)   Let p be a literal. p: $[(i,0),\mu]$ and p:$[(0,j),\mu]$ are called *well extended vector annotated literals*(weva-literals for short), where $i,j \in \{1,2,\cdots\}$, and $\mu \in \{\ \alpha,\ \beta,\ \gamma\ \}$.

DEFINITION 4 (EVALPSN)   If $L_0,\cdots,L_n$ are weva-literals,

$$L_1 \wedge \cdots \wedge L_i \wedge \sim L_{i+1} \wedge \cdots \wedge \sim L_n \rightarrow L_0$$

is called an *Extended Vector Annotated Logic Program clause with Strong Negation*   abbr. EVALPSN clause. An *Extended Vector Annotated Logic Program with Strong Negation* is a finite set of EVALPSN clauses.

Deontic notions and fact are represented by extended vector annotations as follows : "fact" is represented by an extended vector annotation $[(m,0),\alpha]$; "obligation" by an extended vector annotation $[(m,0),\beta]$; "forbiddance" by an extended vector annotation $[(0,m),\beta]$; "permission" by an extended vector annotation $[(0,m),\gamma]$; where m is a positive integer.

3   Safety Verification

Pipeline Network

We take a pipeline network described in Fig. 2 as an example. In Fig. 2, arrows indicate liquid flows, home-plate figures do tanks, and cross figures do valves. We consider logical safety verification for the network like railway network in-
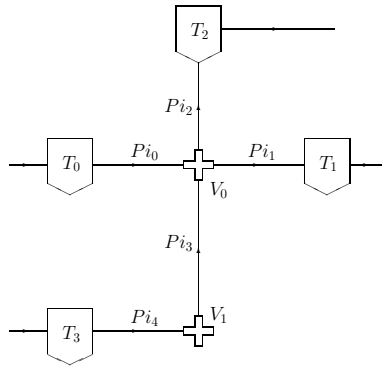
Fig. 2. Pipeline Network
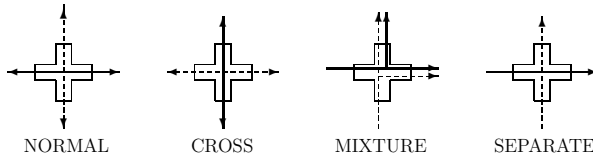
Fig. 3. Valve Control Directions
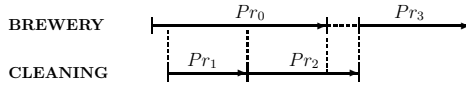
Fig. 4. Process Schedule Chart

terlocking. We have physical entities : tanks $\mathcal{TA} = \{T_0, T_1, T_2, T_3\}$; pipes $\mathcal{PI} = \{Pi_0, Pi_1, Pi_2, Pi_3, Pi_4\}$; (a pipe is a pipeline including neither valves nor tanks) ; valves $\mathcal{VA} = \{V_0, V_1\}$; and logical entities : processes $\mathcal{PR} = \{Pr_0, Pr_1, Pr_2, Pr_3\}$; (processes are logically defined as a set of sub-processes and valves) ; sub-processes $\mathcal{SPR} = \{SPr_0, SPr_1, SPr_2, SPr_3, SPr_4\}$ in the network. Each entity has logical or physical states.

Sub-processes have two states *locked*($\mathtt{l}$) and *free*($\mathtt{f}$), "the sub-process is locked" means that the sub-process is logically locked(reserved) for one process (liquid flow) and "free" means unlocked.

Processes have two states *set*($\mathtt{s}$) and *unset*($\mathtt{xs}$), "the process is set" means that all the sub-processes in the process are locked and "unset" means not set. Valves in the network are supposed to control two liquid flows in the normal and cross directions as shown in **Fig.**3.

Valves have two controlled states ; *controlled mixture*($\mathtt{cm}$) which means that the valve is controlled to mix the two liquid flows in the normal and cross directions, *controlled separate*($\mathtt{cs}$) which means that the valve is controlled to separate the two liquid flows in the normal and cross directions as shown in **Fig.**3.   We suppose five kinds of cleaning liquid, cold water(cw), warm water(ww), hot water(hw), nitric acid(na) and caustic soda(cs). Then we consider the following four processes in the pipeline network with the process schedule (**Fig.** 4) : Pr_0, a brewery process ; the tank T_0 to the valve $V_0(\mathtt{cs})$ to the tank T_1 : Pr_1 and Pr_2, cleaning processes by nitric acid and cold water ; the tank T_2 to the valve $V_1(\mathtt{cm})$ to the valve $V_1(\mathtt{cs})$ to the tank T_3 : Pr_3, a brewery process with mixing ; the tank T_0 to the valve $V_0(\mathtt{cs})$ to the tank T_1, and the tank T_2 to the valve $V_1(\mathtt{cm})$ to the valve $V_1(\mathtt{cs})$ to the tank T_3 :

In order to verify the safety for valve control, we introduce three safety properties, **SPr** (for sub-processes), **Val** (for valves), and **Pr** (for processes).

Pipeline Safety Properties
SPr : It is a forbidden case that the sub-process over a given pipe is simultaneously locked by different kinds of liquid.
Val : It is a forbidden case that valves are controlled to mix different kinds of liquid unexpectedly.

Pr : Whenever a process is set, all its component sub-processes are locked and all its component valves are controlled consistently.

Safety Verification in EVALPSN

First of all, we translate the safety properties for the pipeline network into EVALPSN. The following predicates are used in the EVALPSN safety verification.

- $Pr(i, l) : [\mu_1, \mu_2]$ represents that the process i for the liquid l is set($\mathtt{s}$) or unset($\mathtt{xs}$), where $i \in \{p0, p1, p2, p3\}$ is a process id and $l \in \{b, cw, ww, hw, na, cs\}$, where $\mu_1 \in \mathcal{T}_{v_1} = \{\bot_1, \mathtt{s}, \mathtt{xs}, \top_1\}$ and $\mu_2 \in \mathcal{T}_d$.
- $SPr(i, j, l) : [\mu_1, \mu_2]$ represents that the sub-process from the valve i (or the tank i) to the valve j (or the tank j) occupied by the liquid l is locked($\mathtt{l}$) or free($\mathtt{f}$). Moreover, if a sub-process is free then the kind of liquid in the pipe is not cared, and the liquid is represented by the symbol "0"(zero). Therefore, we have $l \in \{b, cw, ww, hw, na, cs, 0\}$. $i \in \{v0, v1\}$ and $j \in \{t0, t1, t2, t3\}$ are valve id and tank id corresponding to the valves $V_{0,1}$ and the tanks $T_{0,1,2,3}$. where $\mu_1 \in \mathcal{T}_{v_2} = \{\bot_2, \mathtt{l}, \mathtt{f}, \top_2\}$, and $\mu_2 \in \mathcal{T}_d$.
- $Val(i, l_n, l_c) : [\mu_1, \mu_2]$ represents that the valve i occupied by the two kinds of liquid $l_n, l_c \in \{b, cw, ww, hw, na, cs, 0\}$ is controlled separate($\mathtt{cs}$) or mixture($\mathtt{cm}$), where $i \in \{v0, v1\}$ is a valve id, and $\mu_1 \in \mathcal{T}_{v_3} = \{\bot_3, \mathtt{cm}, \mathtt{cs}, \top_3\}$ and $\mu_2 \in \mathcal{T}_d$. The arguments l_n and l_c of the predicate Val represent the liquid flows in the normal and cross directions, respectively. Generally, if a valve is released from the controlled state, the liquid flow in the valve is represented by the symbol 0 that means "free".
- $Eql(l_1, l_2) : [\mu_1, \mu_2]$ represents the liquids l_1 and l_2 have the same($\mathtt{sa}$) kind or different($\mathtt{di}$) ones, where $l_1, l_2 \in \{b, cw, ww, hw, na, cs, 0\}$, and $\mu_1 \in \mathcal{T}_{v_4} = \{\bot_2, \mathtt{sa}, \mathtt{di}, \top_4\}$ and $\mu_2 \in \mathcal{T}_d$.
- $Tan(ti, l) : [\mu_1, \mu_2]$ represents that the tank T_i has been filled fully($\mathtt{fu}$) with the liquid l or empty($\mathtt{em}$), where $i \in \{0, 1, 2, 3\}$, $l \in \{b, cw, ww, hw, na, cs, 0\}$, and $\mu_1 \in \mathcal{T}_{v_5} = \{\bot_5, \mathtt{fu}, \mathtt{em}, \top_5\}$ and $\mu_2 \in \mathcal{T}_d$.

Now we formalize the three safety properties **SPr**, **Val** and **Pr** in EVALPSN. Then we interpret the safety properties deontically.

SPr. This condition can be intuitively interpreted as derivation rules of forbiddance. If a sub-process from a valve (a tank) i to a valve (a tank) j is locked by one kind of liquid, it is forbidden for the sub-process to be locked by different kinds of liquid simultaneously. This condition is translated into the following EVALPSN clauses :

$$SPr(i, j, l_1) : [\mathtt{l}, \alpha] \wedge \sim Eql(l_1, l_2) : [\mathtt{sa}, \alpha] \rightarrow SPr(i, j, l_2) : [\mathtt{f}, \beta], \qquad (1)$$

where $l_1, l_2 \in \{b, cw, ww, hw, na, cs\}$. Moreover, in order to derive permission for locking sub-processes we need the following EVALPSN clauses :

$$\sim SPr(i, j, l) : [\mathtt{f}, \beta] \rightarrow SPr(i, j, l) : [\mathtt{f}, \gamma], \qquad (2)$$

where $l \in \{b, cw, ww, hw, na, cs\}$.

Val. This condition also can be intuitively interpreted as derivation rules of forbiddance. We have to consider two cases : one is for deriving the forbiddance from changing the control state of the valve, and another one is for deriving the forbiddance from mixing different kinds of liquid without changing the control state of the valve.

Case 1. If a valve is controlled separate, it is forbidden for the valve to be controlled mixture, conversely, if a valve is controlled mixture, it is forbidden for the valve to be controlled separate. Thus, generally we have the following EVALPSN clauses :

$$Val(i, l_n, l_c) : [\mathsf{cs}, \alpha] \wedge \sim Eql(l_n, 0) : [\mathsf{sa}, \alpha] \wedge \sim Eql(l_c, 0) : [\mathsf{sa}, \alpha]$$
$$\rightarrow Val(i, l_n, l_c) : [\mathsf{cs}, \beta], \tag{3}$$
$$Val(i, l_n, l_c) : [\mathsf{cm}, \alpha] \wedge \sim Eql(l_n, 0) : [\mathsf{sa}, \alpha] \wedge \sim Eql(l_c, 0) : [\mathsf{sa}, \alpha]$$
$$\rightarrow Val(i, l_n, l_c) : [\mathsf{cm}, \beta], \tag{4}$$

where $l_n, l_c \in \{b, cw, ww, hw, na, cs, 0\}$.

Case 2. Next, we consider the other forbiddance derivation case in which different kinds of liquid are mixed even if the valve control state is not changed. We have the following EVALPSN clauses :

$$Val(i, l_{n_1}, l_{c_1}) : [\mathsf{cs}, \alpha] \wedge \sim Eql(l_{n_1}, l_{n_2}) : [\mathsf{sa}, \alpha] \wedge \sim Eql(l_{n_1}, 0) : [\mathsf{sa}, \alpha]$$
$$\rightarrow Val(i, l_{n_2}, l_{c_2}) : [\mathsf{cm}, \beta], \tag{5}$$
$$Val(i, l_{n_1}, l_{c_1}) : [\mathsf{cs}, \alpha] \wedge \sim Eql(l_{c_1}, l_{c_2}) : [\mathsf{sa}, \alpha] \wedge \sim Eql(l_{c_1}, 0) : [\mathsf{sa}, \alpha]$$
$$\rightarrow Val(i, l_{n_2}, l_{c_2}) : [\mathsf{cm}, \beta], \tag{6}$$
$$Val(i, l_{n_1}, l_{c_1}) : [\mathsf{cm}, \alpha] \wedge \sim Eql(l_{n_1}, l_{n_2}) : [\mathsf{sa}, \alpha] \rightarrow Val(i, l_{n_2}, l_{c_2}) : [\mathsf{cs}, \beta], \tag{7}$$
$$Val(i, l_{n_1}, l_{c_1}) : [\mathsf{cm}, \alpha] \wedge \sim Eql(l_{c_1}, l_{c_2}) : [\mathsf{sa}, \alpha] \rightarrow Val(i, l_{n_2}, l_{c_2}) : [\mathsf{cs}, \beta], \tag{8}$$

where $l_{n_1}, l_{c_1} \in \{b, cw, ww, hw, na, cs, 0\}$ and $l_{n_2}, l_{c_2} \in \{b, cw, ww, hw, na, cs\}$. Note that the EVALPSN clause $\sim Eql(l_n, 0) : [\mathsf{sa}, \alpha]$ represents there does not exist information such that the normal direction with the liquid l_n in the valve is free (not controlled). As well as the case of sub-processes, in order to derive permission for controlling valves, we need the following EVALPSN clauses :

$$\sim Val(i, l_n, l_c) : [\mathsf{cm}, \beta] \rightarrow Val(i, l_n, l_c) : [\mathsf{cm}, \gamma], \tag{9}$$
$$\sim Val(i, l_n, l_c) : [\mathsf{cs}, \beta] \rightarrow Val(i, l_n, l_c) : [\mathsf{cs}, \gamma], \tag{10}$$

where $l_n, l_r \in \{b, cw, ww, hw, na, cs, 0\}$.

Pr. This condition can be intuitively interpreted as derivation rules of permission and directly translated into EVALPSN clauses as a rule "if all the components of the process can be locked or controlled consistently, then the process can be set". For example, if the brewery process Pr_0 consists of the sub-process the tank T_0 to the valve V_0, the valve V_0 with controlled separate by beer in the

normal direction, and the sub-process the valve V_0 to the tank T_1, then we have the following EVALP clause to obtain the permission for setting the process Pr_0,

$$SPr(t0, v0, b)\!:\![\mathtt{f}, \gamma] \wedge SPr(v0, t1, b)\!:\![\mathtt{f}, \gamma] \wedge Val(v0, b, l)\!:\![\mathtt{cm}, \gamma] \wedge$$
$$Tan(t0, b)\!:\![\mathtt{fu}, \alpha] \wedge Tan(t1, 0)\!:\![\mathtt{em}, \alpha] \rightarrow Pr(p0, b)\!:\![\mathtt{xs}, \gamma], \tag{11}$$

where $l \in \{b, cw, ww, hw, na, cs, 0\}$.

Control Examples. We suppose that all the sub-processes and valves in the pipeline network are unlocked (free) and no process has already started at the initial stage. Then, if we have all the requests for the processes $Pr_{0,1,2,3}$, in order to verify the safety for all the processes, the following fact EVALP clauses (as the current states) are input to the EVALPSN pipeline control :

$$SPr(t0, v0, 0)\!:\![\mathtt{f}, \alpha], \quad Val(v0, 0, 0)\!:\![\mathtt{cs}, \alpha], \quad SPr(v0, t1, 0)\!:\![\mathtt{f}, \alpha],$$
$$Val(v1, 0, 0)\!:\![\mathtt{cs}, \alpha], \quad SPr(v0, t2, 0)\!:\![\mathtt{f}, \alpha], \quad SPr(v1, v0, 0)\!:\![\mathtt{f}, \alpha],$$
$$SPr(t3, v1, 0)\!:\![\mathtt{f}, \alpha], \quad Tan(t0, b)\!:\![\mathtt{fu}, \alpha], \quad Tan(t1, 0)\!:\![\mathtt{em}, \alpha],$$
$$Tan(t2, 0)\!:\![\mathtt{em}, \alpha], \quad Tan(t3, na)\!:\![\mathtt{fu}, \alpha].$$

Then all the sub-processes and valves in the network are permitted to be locked or controlled. However the tank conditions do not permit the processes Pr_2 and Pr_3 to be set. We show that the beer process Pr_0 can be verified to be set as follows : we can have neither the forbiddance from locking the sub-processes SPr_0 and SPr_1, nor the forbiddance from controlling the valve V_0 separate with beer in the normal direction, by the EVALPSN clauses (1), (4), (5), (6) and the above fact EVALP clauses ; therefore we have the permission for locking the sub-processes SPr_0 and SPr_1, and controlling the valve V_0 separate with beer in the normal direction and any liquid in the cross direction, $SPr(t0, v0, b)\!:\![\mathtt{f}, \gamma]$, $Val(v0, b, l) : [\mathtt{cm}, \gamma]$, $SPr(v0, t1, b) : [\mathtt{f}, \gamma]$, where $l \in \{b, cw, ww, hw, na, cs, 0\}$, by the EVALPSN clauses (2) and (9) ; moreover, we have the tank conditions, $Tan(t0, b)\!:\![\mathtt{fu}, \alpha]$ and $Tan(t1, 0)\!:\![\mathtt{em}, \alpha]$, thus we have the permission for setting the beer process Pr_0, $Pr(p0, b)\!:\![\mathtt{xs}, \gamma]$, by the EVALPSN clause (11).

4   Conclusion

In this paper, we have introduced EVALPSN based safety verification for pipeline process control. Furthermore, If we consider the safety verification for process order we need to verify the safety for the temporal order relations such as "the process 1 must be before the process 2".

References

1. Nakamatsu,K., Abe,J.M., and Suzuki,A., "A Defeasible Deontic Reasoning System Based on Annotated Logic Programming", Proc. the 4th Int'l Conf. Computing Anticipatory Systems, AIP Conf. Proc. Vol.573 (2001) 609–620

2. Nakamatsu,K., Abe,J.M., and Suzuki,A., "Annotated Semantics for Defeasible Deontic Reasoning", Proc. the 2nd Int'l Conf. Rough Sets and Current Trends in Computing, LNAI Vol.2005, Springer-Verlag (2001) 432–440
3. Nakamatsu,K., Abe,J.M., and Suzuki,A., "A Railway Interlocking Safety Verification System Based on Abductive Paraconsistent Logic Programming", Soft Computing Systems, Frontiers in AI Applications, Vol.87 (2002) 775–784
4. Nakamatsu,K., Seno,T., Abe,J.M., and Suzuki,A., "Intelligent Real-time Traffic Signal Control Based on a Paraconsistent Logic Program EVALP", Proc. the 9th Int'l Conf. Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, LNCS Vol.2639, Springer-Verlag (2003) 719–723

Paraconsistent Artificial Neural Network: An Application in Cephalometric Analysis

Jair Minoro Abe[1,2], Neli R.S. Ortega[2],
Maurício C. Mário[2], and Marinho Del Santo Jr.[3]

[1] Institute For Advanced Studies – University of São Paulo
jairabe@uol.com.br
[2] School of Medicine - HCFMUSP/LIM01, University of São Paulo
neli@dim.fm.usp.br, cmario@unisanta.br
[3] Orthodontist and Researcher
marinho@delsanto.com.br

Abstract. Paraconsistent artificial neural network (PANN) is a mathematical structure based on paraconsistent logic, which allows dealing with uncertainties and contradictions. In this paper we propose an application of PANN to analyze cephalometric measurements in order to support orthodontics diagnostics. Orthodontic and cephalometrical analysis taking into account several uncertainties and contradictions, ideal scenario to be treated by paraconsistent approach.

1 Introduction

Craniofacial discrepancies, either skeletal or dental, are assessed in lateral cephalograms by cephalometric analyses. In Orthodontics, such quantitative analysis compares an individual with a sample of a population, matched by gender and age. However, cephalometric measurements hold significant degrees of insufficiency and inconsistency, making its clinical application less effective than ideal.

A conventional cephalometric analysis compares individual measurements to a pattern, i.e., a norm for that variable, assessed in a sample of patients which has the same age and gender. Such piece of information is, in the best scenario, a suggestion of the degree of deviation from the norm, for such particular variable. A better scenario would be to knowing how much the value of a variable of a certain patient is deviated from its norm. It would be better if we could quantify the "noise" carried by a cephalometric value and filter its potential damage to the contextual result. In this sense, developing a mathematical structure able to provide quantitative information and modeling the inconsistencies, contradictions and evidences of abnormality of these variables is relevant and useful.

In order to analyze skeletal and dental changes we selected a set of cephalometric variables based on an expert knowledge. Figures 1 and 2 show these variables and analysis. In this paper we propose an expert system able to assess the degree of evidence of abnormality of each variable, suggesting a diagnosis for the case and, consequently, an adequate treatment plan. The expectance is that system increases the potential clinical application of the cephalometric analysis, potentially better addressing a more efficient therapy.

1.	Basion
2.	Sella
3.	Nasion
4.	Posterior Nasal Spine
5.	Anterior Nasal Spine
6.	Inter-Molars
7.	Inter-Incisors
8.	Gonion
9.	Menton
10.	Gnathion
11.	A Point
12.	B Point
13.	Pogonion
14.	Incisal Edge - Upper Incisor
15.	Apex - Upper Incisor
16.	Incisal Edge - Lower Incisor
17.	Apex - Lower Incisor

Fig. 1. Cephalometric Variables

1.	Anterior Cranial Base
2.	Palatal Plane (PP)
3.	Oclusal Plane (OP)
4.	Mandibular Plane (MP)
5.	Cranial Base
6.	Y Axis
7.	Posterior Facial Height
8.	Anterior Facial Height - Median Third
9.	Anterior Facial Height - Lower Third
10.	Anterior Facial Height
11.	SNA
12.	SNB
13.	Long Axis - Upper Incisor
14.	Long Axis - Lower Incisor
15.	A Point - Pogonion Line

Wits: distance between the projections of the A and B Points on the occlusal plane.

Fig. 2. Proposed Cephalometric Analysis

2   Paraconsistent Artificial Neural Network

The Artificial Neural Network has been extensively studied in AI, particularly for pattern recognition and dynamical process. In this paper we use a new theory of artificial neural network based on a paraconsistent annotated logic $E\tau$. Roughly speaking,

paraconsistent logics allow formulas of the form A & $\neg A$ (A and the negation of A) to be applied in a non-trivial manner in deductions. Also, the paraconsistent annotated logic $E\tau$ allows to manipulate the concepts off uncertainty and paracompleteness in its structure. The atomic formulas of $E\tau$ is of the type $p_{(\mu, \lambda)}$, where $(\mu, \lambda) \in [0, 1]^2$ and $[0, 1]$ is the real unitary interval (p denotes a propositional variable). An order relation is defined on $[0, 1]^2$: $(\mu_1, \lambda_1) \leq (\mu_2, \lambda_2) \Leftrightarrow \mu_1 \leq \mu_2$ and $\lambda_1 \leq \lambda_2$, constituting a lattice that will be symbolized by τ. A detailed account of annotated logics is to be found in [1]. $p_{(\mu, \lambda)}$ can be intuitively read: "It is assumed that p's belief degree (or favorable evidence) is μ and disbelief degree (or contrary evidence) is λ." Thus, (1.0, 0.0) intuitively indicates total belief, (0.0, 1.0) indicates total disbelief, (1.0, 1.0) indicates total inconsistency, and (0.0, 0.0) indicates total paracompleteness. In the Fig. 3 and Table 1 we consider the output states.
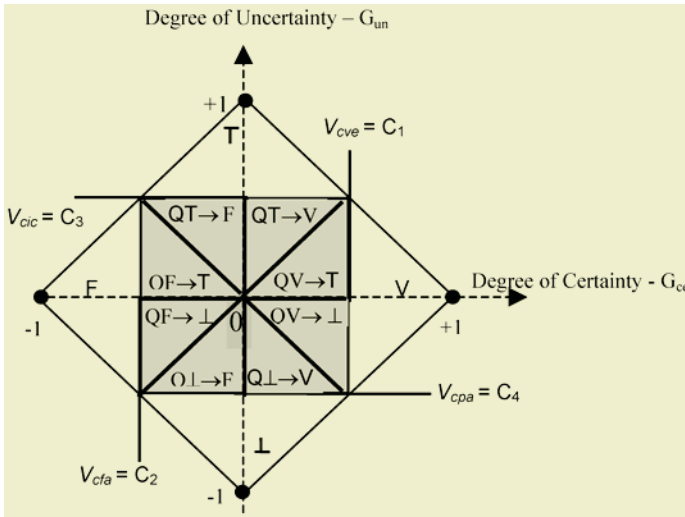
Fig. 3. Representation of the certainty degrees and of contradiction degrees

These states can be described with the values of the certainty degree and uncertainty degree by means of suitable equations. In this work we have chosen the resolution 12 (number of the regions considered according to the Fig. 3, but the resolution is totally dependent on the precision of the analysis required in the output and it can be externally adapted according to the applications. So, such limit values called Control Values are:

V_{cic} = maximum value of uncertainty control = C_3
V_{cve} = maximum value of certainty control = C_1
V_{cpa} = minimum value of uncertainty control = C_4
V_{cfa} = minimum value of certainty control = C_2

For the discussion in the present paper we have used: $C_1 = C_3 = ½$ and $C_2 = C_4 = -½$.

In the paraconsistent analysis the main aim is to know how to measure or to determine the certainty degree concerning a proposition, if it is False or True. Therefore, for this, we take into account only the certainty degree G_{ce}.

Table 1. Logical states

Extreme States	Symbol	Non-extreme states	Symbol
True	V	Quasi-true tending to Inconsistent	$QV{\rightarrow}T$
False	F	Quasi-true tending to Paracomplete	$QV{\rightarrow}\perp$
Inconsistent	T	Quasi-false tending to Inconsistent	$QF{\rightarrow}T$
Paracomplete	$\perp$	Quasi-false tending to Paracomplete	$QF{\rightarrow}\perp$
		Quasi-inconsistent tending to True	$QT{\rightarrow}V$
		Quasi-inconsistent tending to False	$QT{\rightarrow}F$
		Quasi-paracomplete tending to True	$Q\perp{\rightarrow}V$
		Quasi-paracomplete tending to False	$Q\perp{\rightarrow}F$


```
*/Adjustable values */
  Vcve =C1 */maximum value of
certainty control */
  Vcfa =C2 */minimum value of
certainty control */
  Vcic =C3 */maximum value of
uncertainty control */
  Vcpa =C4 */minimum value of
uncertainty control*/
*/Input /Variables */
  μ, λ
*/Output /Variables */
  Digital output = S1
  Analog output = S2a
  Analog output = S2b
*/Mathematical expressions */
  0≤ μ ≤ 1 e 0≤ λ≤ 1
  Gun = μ + λ - 1
  Gce = μ - λ
*/ determination of the extreme
states */
  if Gce ≥ C1 then S1 = V
  if Gce ≥ C2 then S1 = F
  if Gun ≥ C3 then S1 = T
  if Gun ≤ C4 then S1 = ⊥
  If not: S1 = I -
Indetermination
  Gun = S2a
  Gce = S2b
```

Fig. 4. The Basic Paraconsistent Artificial Neural Cell

In the paraconsistent analysis the main aim is to know how to measure or to determine the certainty degree concerning a proposition, if it is False or True. Therefore, for this, we take into account only the certainty degree G_{ce}. The uncertainty degree G_{un} indicates the measure of the inconsistency or paracompleteness. If the certainty degree is low or the uncertainty degree is high, it generates an indefinition.

The resulting certainty degree G_{ce} is obtained as follows:

If: $V_{cfa} \le G_{un} \le V_{cve}$ or $V_{cic} \le G_{un} \le V_{cpa} \Rightarrow G_{ce} =$ Indefinition

For: $V_{cpa} \le G_{un} \le V_{cic}$

If: $G_{un} \le V_{cfa} \Rightarrow G_{ce} =$ False with degree G_{un}

$V_{cic} \le G_{un} \Rightarrow G_{ce} =$ True with degree G_{un}

The algorithm that expresses a basic Paraconsistent artificial neural cell (PANC) is presented and represented in Fig. 4 above.

3   Architecture of the Paraconsistent Artificial Neural Network

The selected cephalometric variables are inserted in the paraconsistent network in the following three units: Unit I, considering the antero-posterior discrepancy, Unit II, considering vertical discrepancy, and Unit III, taking into account dental discrepancy (see Fig. 5).
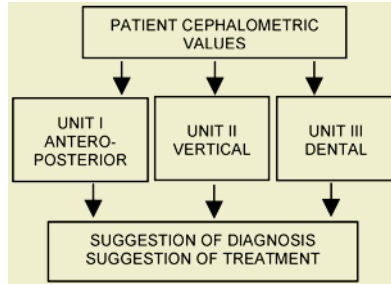
Fig. 5. Functional macro view of the neural architecture

Unit I is made by 2 levels. The first one involves the ANB and Wits variables. At the second level, there is a combination of the result of the level 1 and the variables SNA and SNB. The output of the second level is regard to the position of the maxilla and the mandible, classifying it as: well positioned, protruded, retruded, tending to protruded and tending to retruded. The classes protruded and retruded coming with their respective degrees of evidence of abnormality. Moreover, the classes "tending to" suggest the assessment of the outputs of the Units II and III. The variables pertaining to the Unit II are divided in three different groups: Group I: Se-Go/Na-Me proportion. The value of this proportion may result in normal, vertical or horizontal face. Group II: Y Axis. The value of this angle may also result in normal, vertical or horizontal face. Group III: angles SeNa/PP, SeNa/OP and SeNa/MP. Each one of the three angles may also result in normal, vertical or horizontal face. The combination of the output form Groups I, II and III will also result in normal, vertical or horizontal face. The variables pertaining to the Unit III are divided in three different groups: Group I: U1.PP and U1.SN angles and linear measurement U1-NA, taking in account the SNA angle (Unit I). The upper incisors may be in a normal position, proclined or retroclined. Group II: L1.APg, L1. NB and L1.GoMe angles and linear measurements L1-APg, L1-NB, taking in account the SNB angle. The lower incisors may be in a normal position, proclined or retroclined. Group III: angle U1.L1. This value results in three possible positions: normal, proclined and retroclined. The combination of the outputs of the Groups I, II and III will result in normal, proclined, retroclined, tending to proclined and tending to retroclined. Each unit has the following components, represented in the Fig. 6:

 a) Standardization: the difference between the data from the patient radiographs and its relative norm, by age and gender;
 b) Data modeling: matrices with possible degrees of evidence of abnormality.

Modeling Tool: matrices with all possible values of evidence of abnormality related to the variables of the patients, using the standard deviation values and providing the degree of evidence of abnormality.

Learning Cell: "learn" the functions of the matrices which contain the degrees of evidence of abnormality of the variables.

Neural Network of Evidence: file the learned information and return with the degree of evidence according to the standard deviation of each variable in a contextualized form, i.e., considering the degrees of evidence of abnormality of the other variables.

Expert System: provides the diagnosis, inferred from the neural degrees of evidence of abnormality of each unit.

Treatment Plan: based upon the specialist system provides a suggestion of diagnosis for that specific orthodontic case.

Fig. 6. Functional micro view of the structure of each unit of the Fig. 5

The system works with four craniofacial measurements in the Unit I, resulting in 46 inferences, giving 33 outputs of types of malocclusion (diagnosis). In the Unit II, 5 craniofacial measurements allow 90 inferences and 4 outputs of diagnosis. In the Unit III, 9 cephalometric measurements were assessed, giving 87 inferences and 12 outputs of diagnosis. Suggestions of treatment are proposed for all diagnosis pointed. At total, the expert system is based upon 18 craniofacial cephalometric measurements, 223 inferences and 49 outputs of diagnosis and suggestions of treatment.

4   Discussion

Cephalometrics is the most useful tool for orthodontic diagnosis, since assess craniofacial skeletal and dental discrepancies. However, conventional cephalometrics holds important limitations, mostly due to the fact that the cephalometric variables are not assessed under a contextualized scope and carry on important variation when compared to samples norms. Because of that, its clinical application is relative, subjective, and routinely less effective than the expected. In addiction, discordance between orthodontists about diagnosis and treatments it is not uncommon, due to the inevitable uncertainties involved in the cephalometrics variables. In our point of view, this is a perfect scenario to evaluate the paraconsistent neural network capacity to perform with uncertainties, inconsistencies, and paracompleteness in a practical problem.

In this work an expert system to support orthodontic diagnosis was developed based on the paraconsistent approach. In the structure proposed the inferences were

based upon the degrees of evidence (favorable and unfavorable) of abnormality for cephalometrics variables, which may have infinite values between "0" and "1". Therefore, the system may be refined with more or less outputs, depending upon the need. Such flexibility allows that the system can be modeled in different ways, allowing a more fine adjusting.

The system require measurements taken from the head lateral radiography of the patient that will be assessed. The precision of the system increase as much as data is added, enriching the learning cells for that specific situation. On the other hand, if the radiographic information provided was insufficient, the system give back an Undefined (*Un*) output, state that is also predictable in the paraconsistent logic. Therefore, possible "noise" as consequence of the lack of data does not prevent that the initial goals of neural network can be achieved.

In order to evaluate the practical aspects of this paraconsistent neural network we analyzed the concordance with the system and an expert opinion about 40 real cases. As preliminary results, the degrees of evidence of abnormality were tested for the three Units. Kappa values, comparing the software and the opinion of the expert were: Unit 1 = 0.485; Unit 2 = 0.463 and Unit 3 = 0.496 (upper incisors), = 0.420 (lower incisors) and = 0.681 (upper combined with lower incisors). The strength of agreement is at least moderate. It is important to highlight that the initial data, used for the classification of each group presented significant variation and the opinions of the specialist about particular problems hold important subjective weight.

Finally, although the system needs more accurate validation, the preliminary results are encouraged and show in a doubtless way that paraconsistent neural networks may contribute for the development of expert systems taking to account uncertainties and contradictions, presented in the most real problems, particularly heath areas, opening a new promising tool of research.

References

1. J. M. Abe, "*Fundamentos da Lógica Anotada*" (Foundations of Annotated Logics), (in Portuguese) Ph. D. Thesis, University of São Paulo, São Paulo, 1992.
2. J. M. Abe, Paraconsistent Artificial Neural Networks: an Introduction, *Lecture Notes In Artificial Intelligence* 3214, Springer, Eds. J.G. Carbonell & J. Siekmann, ISSN 0302-9743, ISBN 3-540-23206-0, pp. 942-948, 2004.
3. J. I. Da Silva Filho & J. M. Abe, *Fundamentos das Redes Neurais Paraconsistentes* – Destacando Aplicações em Neurocomputação, (in Portuguese) Editôra Arte & Ciência, ISBN 85-7473-045-9, 247 pp., 2001.
4. A.P. Dempster, Generalization of Bayesian inference, *Journal of the Royal Statistical Society*, Series B-30, 205-247, 1968.
5. R. Hecht-Nielsen, *Neurocomputing*. New York, Addison Wesley Pub. Co., 1990.
6. T. Kohonen, *Self-Organization and Associative Memory*. Springer-Verlag, 1984.
7. R. Sylvan & J. M. Abe, On general annotated logics, with an introduction to full accounting logics, *Bulletin of Symbolic Logic*, 2, 118-119, 1996.
8. C. Stephens & N. Mackin, The validation of an orthodontic expert system rule-base for fixed appliance treatment planning. *Eur J Orthod*, 20: 569-578, 1998
9. M. Del Santo Jr., Influence of the Occlusal Plane on ANB and Wits Assessments of Anteroposterior Relationship of the Jaws. *American Journal of Orthodontics and Dentofacial Orthopedics*, in press, 2005.

10. D. R. Martins, G. R. P. Janson, R. R. Almeida, A. Pinzan, J. F. C. Henriques, M.R. Freitas, *Atlas de Crescimento Craniofacial*. (in Portuguese) Ed. Santos, São Paulo, SP, 1998.

11. Y.Sorihashi Y, C. D. Stephens, K. Takada, An inference modeling of human visual judgment of sagittal jaw-base relationships based on cephalometry: Part II. *Am. J. Orthod. Dentofac. Orthop.* 117: 303-11, 2000.

12. C. Steiner, Cephalometrics for you and me, *Am. J. Orthod.* 39: 729-55, 1953.

13. A. Jacobson, "Wits" appraisal of jaw disharmony, *Am. J. Orthod.* 67: 125-38, 1975.

14. Jacobson, The application of the "Wits" appraisal. *Am. J. Orthod.* 70:179-89,1971.

15. R. M. Ricketts, Cephalometric analysis and synthesis, *Angle Orthod.* 31:141-56, 1961.

Non-alethic Reasoning in Distributed Systems

Jair Minoro Abe[1,2], Kazumi Nakamatsu[3], and Seiki Akama[4]

[1] Institute For Advanced Studies – University of São Paulo, Av. Prof. Luciano Gualberto
Trav. J, 374, Térreo, Cidade Universitária CEP 05508-900 – São Paulo – SP- Brazil
`jairabe@uol.com.br`
[2] Information Technology Dept., ICET – Paulista University, Rua Dr. Bacelar 1212
CEP 04026-002 – São Paulo – SP – Brazil
[3] School of H.E.P.T., Himeji Institute of Technology
1-1-12 Shunzaike, HIMEJI 670-0092, Japan
`nakamatu@shse.u-hyogo.ac.jp`
[4] Computational Logic Laboratory, Department of Information Systems
Teikyo Heisei University, 2289 Uruido, Ichihara-shi, Chiba 290-01 Japan
`akama@cn.thu.ac.jp`

Abstract. In this work we present some of our investigations concerning para-consistent, paracomplete and non-alethic systems for knowledge representation in AI. Inconsistency and paracompleteness are natural phenomena in describing parts of our reality. Thus, we need more generical tools, i.e., formal systems for modeling contradictory and paracomplete knowledge, and in this paper we focus non-alethic knowledge in distributed systems.

1 Introduction

When several agents are doing a task, the classical requirement that at least one of a proposition and its negation must be false does not always fit our intuitions. Contradictions arise for different reasons: for example, we may have conflicting goals or need to share limited resources. They may be avoided, solved, possessed, or even created deliberately. Also, if P is a "vague" predicate and a is a borderline individual we may feel that both $P(a)$ and $\neg P(a)$ are false. Similarly, if A is a contingent proposition and F is the future operator, we may think that $F(A)$ and $\neg F(A)$ are both false. In general a paracomplete theory can be conceived as the underlying logic of an incomplete theory in the strong sense, i.e. of a theory according to which a proposition and its negations are both false.

When it is taken questions of logical omniscience, one relevant concept that appears is that of contradiction. Roughly speaking, the problem of logical omniscience is an agent to know all logical consequences from a set of premises, in particular all tautologies. A good discussion is to be found in [7]. Also, it is well known that from a contradiction all formulas is provable, and so due previous observation, all agents know all formulas ! Definitely this is not also natural.

The use of modal systems for modeling knowledge and belief has been largely considered in AI. For instance, it seems that the first one to consider knowledge and belief to machines was [14]. Subsequently, [14], [15], [16], [8], [9], among others, have considered knowledge in multi-agent systems, besides other approaches.

The essential ideas underlying the systems proposed by [9] and collaborators can be summarized as follows: $\Box_i A$ can be read *agent i knows A, i = 1, ... , n. Common knowledge* and *distributed knowledge* are also defined in terms of additional modal

operators: $\square_G$ ("everyone in the group G knows"), $\square_G^C$ ("it is common knowledge among agents in G"), and $\square_G^D$ ("it is distributed knowledge among agents in G") for every nonempty subset G of $\{1, \ldots , n \}$.

Nevertheless, the most of these proposals use extensions of Classical Logic or at least part of it, keeping as much as fundamental characteristics of Classical Logic.

So, in this paper we propose a logical system for reasoning with inconsistencies and paracompleteness in distributed systems in a nontrivial way.

The proposed system is a logical framework that has the following characteristics:

1. The principle of contradiction, in the form $\neg(A \wedge \neg A)$, is not valid in general;
2. From two contradictory propositions, A and $\neg A$, we cannot deduce any formula B whatsoever;
3. The principle of middle excluded, in the form $A \vee \neg A$, is not valid in general; so, we can have formulas A and $\neg A$, both false.
4. They contain the most important schemes and rules of inference of the classical propositional calculus that are compatible with conditions 1, 2, and 3 above.

2   The Non-alethic Multimodal Logics Ω_n ($1 \leq n < \omega$)

The desired system is obtained by using heavily ideas of [6] and [7].

We present, in this section, a family of non-alethic multimodal propositional calculi Ω_n ($1 \leq n < \omega$). The language of the calculi Ω_n ($1 \leq n < \omega$) is the same for all of them; let us call it L. The primitive symbols of L are the following:

1. Propositional variables: a denumerable set of propositional variables
2. $\rightarrow$ (implication)
3. $\wedge$ (conjunction)
4. $\vee$ (disjunction)
5. $\neg$ (negation)
6. Modal operators: $\square_1, \square_2, \ldots , \square_n$, ($n \geq 1$), $\square_G$, $\square_G^C$, $\square_G^D$ (for every nonempty subset C, D of G of $\{1, \ldots , n\}$).
7. parentheses.

Formulas are defined in the usual manner. In L, we put:

Definition 1. Let A be any formula. Then A^o is shorthand for $\neg(A \wedge \neg A)$. A^i stands for $A^{o \cdots o}$, where the symbol ^{o} appears i times for $i \geq 1$. (So A^1 is just A^o). And we write $A^{(i)}$ for $(A \wedge A^1 \wedge A^2 \wedge \ldots \wedge A^i)$. A^* is an abbreviation for $A \vee \neg A$. A^{i*} stands for $A^* \wedge A^{**} \wedge \ldots \wedge A^{* \cdots *}$, where the symbol * appears i times for $i \geq 1$. (So A^{1*} is just A^*). And we write $A^{(i*)}$ for $(A \wedge A^{1*} \wedge A^{2*} \wedge \ldots \wedge A^{i*})$.

Definition 2. We write $(A \leftrightarrow B)$ for $(A \rightarrow B) \wedge (B \rightarrow A)$.

The postulates (axiom schemes and inference rules) of Ω_n ($1 \leq n < \omega$) are (A, B, C are formulas whatsoever):

(1) $A \rightarrow (B \rightarrow A)$
(2) $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$

(3) $((A \to B) \to A) \to A$

(4) $\dfrac{A,\ A \to B}{B}$

(5) $A \wedge B \to A$

(6) $A \wedge B \to B$

(7) $A \to (B \to (A \wedge B))$

(8) $A \to A \vee B$

(9) $B \to A \vee B$

(10) $(A \to C) \to ((B \to C) \to ((A \vee B) \to C))$

(11) $A^{(i*)} \wedge B^{(i)} \to ((A \to B) \to ((A \to \neg B) \to \neg A)$

(12) $(A^{(i)} \wedge B^{(i)}) \to ((A \wedge B)^{(i)} \wedge (A \vee B)^{(i)} \wedge (A \to B)^{(i)})$

(13) $(A^{(i*)} \wedge B^{(i*)}) \to ((A \wedge B)^{(i*)} \wedge (A \vee B)^{(i*)} \wedge (A \to B)^{(i*)})$

(14) $A^{(i)} \to (A \to \neg\neg A) \wedge (A \to (\neg A \to B))$

(15) $A^{(i*)} \to (\neg\neg A \to A)$

(16) $(A^{(i)} \vee A^{(i*)})$

(17) $\square_i(A \to B) \to (\square_i A \to \square_i B),\ i = 1, 2, \ldots, n$

(18) $\square_i A \to \square_i \square_i A,\ i = 1, 2, \ldots, n$

(19) $\text{⫪} \square_i A \to \square_i \text{⫪} \square_i A,\ i = 1, 2, \ldots, n$

(20) $\square_i A \to A,\ i = 1, 2, \ldots, n$

(21) $\dfrac{A}{\square_i A},\ i = 1, 2, \ldots, n$

(22) $\square_G A \leftrightarrow \wedge_{i \in G} \square_i A$

(23) $\square_{G}^{C} A \to \square_G(A \wedge \square_{G}^{C} A)$

(24) $\square_{\{i\}}^{D} A \leftrightarrow \square_i A,\ i = 1, 2, \ldots, n$

(25) $\square_{G}^{D} A \to \square_{G'}^{D},$ if $G' \subseteq G$.

(26) $\dfrac{A \to \square_G(B \wedge A)}{A \to \square_{G}^{C} B}$ (Induction Rule)

The postulates of Ω_ω are those of Ω_n with the exception of (3), (14) - (15).

Definition 3. We write $\neg^{(i)}A$ for $\neg A \wedge A^{(i)}$.

Theorem 1. In Ω_n ($1 \le n < \omega$), all valid schemes and rules of classical positive propositional logic are true. In particular, the deduction theorem is valid in Ω_n ($1 \le n < \omega$). Ω_n ($1 \le n < \omega$) contains intuitionistic positive logic.

Proof. Immediate consequence of postulates (1) – (10) for the case of Ω_n ($1 \le n < \omega$). Since scheme (3) is Peirce's Law, Ω_n ($1 \le n < \omega$) contains positive intuitionistic logic.

Let us denote by F the set of all formulas of Ω_1. Γ stands for subsets of F and let A be a formula. We introduce syntactical concepts like $\Gamma \vdash A$ and $\vdash A$ as usual.

Theorem 2. In Ω_n, ($1 \le n < \omega$), we have: $\vdash \neg A^{(i)} \to (\neg A)^{(i)}$

In Ω_n, $(1 \leq n < \omega)$, $A^{(n)}$ expresses intuitively that the formula A 'behaves' classically, so that the motivation of the postulates (11) - (15) are clear. Furthermore, in this calculus, the connectives $\rightarrow$, $\wedge$, $\vee$, and $\neg$ have all the properties of classical implication, conjunction, disjunction, and negation, respectively. Therefore the classical propositional calculus is contained in Ω_n, though it constitutes a strict sub-calculus of the former. It is not difficult to prove that conditions 1, 2, 3, and 4 of the previous paragraph.

Theorem 3. In Ω_n, $(1 \leq n < \omega)$, the following schemes are not valid:
$(A \wedge \neg A) \rightarrow B$; $(A \wedge \neg A) \rightarrow \neg B$; $(A \rightarrow (\neg A \rightarrow B))$; $(A \leftrightarrow \neg A) \rightarrow B$; $\neg\neg A \rightarrow A$; $(A \leftrightarrow \neg A) \rightarrow \neg B$; $A \rightarrow \neg\neg A$; $(\neg A \wedge (A \vee B)) \rightarrow B$; $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$; $\neg(A \wedge \neg A)$; $A \vee \neg A$; $\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$; $\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$; $(A \wedge \neg A) \leftrightarrow (B \wedge \neg B)$

If Ω_0 denotes the classical propositional multimodal calculus, then the hierarchy Ω_1, Ω_2, ... , Ω_n, ... Ω_ω. Ω_n is such that Ω_{i+1} is weaker than Ω_i for all $(1 \leq n < \omega)$. Ω_ω is the weakest calculus of the hierarchy. It is worthwhile to note that we can extend Ω_i to first and higher-order logics and very strong theories can be formalized on these logics. In addition, all these logics have a two-valued semantics relative to which they are sound and complete, as we will expose subsequently. Let us consider only the calculus Ω_1, to fix ideas. Ω_1 is a paraconsistent and paracomplete calculus, so we can use it to handle inconsistent and paracomplete sets of formulas without immediate trivialization (by this we mean that all formulas in our language may not get entailed by this inconsistent (or paracomplete) set of formulas). It is easy to show that in Ω_1, there are inconsistent (paracomplete) theories which have models, and as a consequence, they are not trivial. In other words, Ω_1 can be the underlying logic of paraconsistent and paracomplete theories. However, it should be noted that when we are working only with formulas that satisfy the principle of contradiction, then Ω_i reduces to Ω_0. In AI, we have to cope with inconsistent and paracomplete systems of information; therefore, the logics based on Ω_n $(1 \leq n < \omega)$ can in principle be employed in that task. We don't enter into details here since the main objective of this section is to give an idea how sophisticated formalisms there are in order to deal with inconsistencies and paracomplete systems of information in distributed systems.

3 Semantic Analysis of Ω_1

In this section we discuss a semantics of Ω_1. Let $\Gamma \subseteq F$. $\overline{\Gamma} = \{A \in F \mid \Gamma \vdash A\}$. Γ is called *trivial* if $\overline{\Gamma} = F$. Γ is called *inconsistent* if there is a formula A such that A and $\neg A \in \Gamma$; otherwise Γ is called *consistent*. We have in a obvious definition the concept of non-trivial maximal set of formulas.

Theorem 4. If Γ is a non-trivial maximal set of formulas and A and B are formulas, then:
1. $\Gamma \vdash A \Leftrightarrow A \in \Gamma$.
2. $A \in \Gamma \Leftrightarrow \daleth A \notin \Gamma$.
3. $\daleth A \in \Gamma \Leftrightarrow A \notin \Gamma$.
4. $A \in \Gamma$ or $\daleth A \in \Gamma$.

5. $\vdash A \Rightarrow A \in \Gamma$.
6. $A, A^o \in \Gamma \Rightarrow \neg A \notin \Gamma$.
7. $\neg A, A^o \in \Gamma \Rightarrow A \notin \Gamma$.
8. $A, A \rightarrow B \in \Gamma \Rightarrow B \in \Gamma$.
9. $A^o, B^o \in \Gamma \Rightarrow (A \wedge B)^o, (A \vee B)^o, (A \rightarrow B)^o \in \Gamma$.
10. $A^o \in \Gamma \Rightarrow (\neg A)^o \in \Gamma$.

We now present Kripke semantics for Ω_n ($1 \leq n < \omega$).

Definition 4. A Kripke structure for Ω_1 (or Ω_1-structure) is a set theoretical structure $K = [W, R_1, R_2, \ldots, R_n, \varphi]$ where W is a nonempty set of elements called 'worlds'. R_i ($i = 1, 2, \ldots, n$) is a binary relation on W such that it is an equivalence relation. φ is an interpretation function with the usual properties.

Definition 5. If A is a formula of Ω_1, and $w \in W$, we define the relation $K,w \vDash A$ (K,w *force A*) by recursion on A. If it is not the case that $K,w \vDash A$ we write $K,w \nvDash A$.

1. $K,w \nvDash A \Rightarrow K,w \vDash \neg A$;
2. $K,w \vDash \neg\neg A \Rightarrow K,w \vDash A$;
3. $K,w \vDash B^o$ and $K,w \vDash A \rightarrow B$ and $K,w \vDash A \rightarrow \neg B \Rightarrow K,w \nvDash A$;
4. $K,w \vDash A \rightarrow B \Rightarrow K,w \nvDash A$ or $K,w \vDash B$;
5. $K,w \vDash A \wedge B \Rightarrow K,w \vDash A$ and $K,w \vDash B$;
6. $K,w \vDash A \vee B \Rightarrow K,w \vDash A$ or $K,w \vDash B$;
7. $K,w \vDash A^o$ and $K,w \vDash B^o \Rightarrow K,w \vDash (A \rightarrow B)^o$ and $K,w \vDash (A \wedge B)^o$ and $K,w \vDash (A \vee B)^o$
8. $K,w \vDash A^*$ and $K,w \vDash B^* \Rightarrow K,w \vDash (A \rightarrow B)^*$ and $K,w \vDash (A \wedge B)^*$ and $K,w \vDash (A \vee B)^*$
9. $K,w \vDash \neg A \Rightarrow K,w \nvDash A$.
10. $K,w \vDash A$ or (exclusive) $K,w \vDash \neg A \Rightarrow K,w \vDash \neg A$ or $K,w \vDash \neg\neg A$
11. If $K,w \vDash A$ or $K,w \vDash \neg A$ and $K,w \vDash B$ or $K,w \vDash \neg B$, then $K,w \vDash A \rightarrow B$ or $K,w \vDash \neg(A \rightarrow B)$, $K,w \vDash (A \wedge B)$ or $K,w \vDash (\neg(A \wedge B))$, and $K,w \vDash (A \vee B)$ or $K,w \vDash (\neg(A \vee B))$
12. $K,w \vDash \Box_i A$ iff $K,w' \vDash A$ for each $w' \in W$ such that $w R_i w'$, $i = 1, 2, \ldots, n$

Definition 6. Let $K = [W, R_1, R_2, \ldots, R_n, I]$ be a Kripke structure for Ω_1. The Kripke structure K *forces* a formula A (in symbols, $K \vDash A$), if $K,w \vDash A$ for each $w \in W$. A formula A is called Ω_1-valid if for any Ω_1-structure K, $K \vDash A$. A formula A is called valid if it is Ω_1-valid for all Ω_1. We symbolize this fact by $\vDash A$. If Γ is a set of formulas, we can introduce the concept of $\Gamma \vDash A$.

Theorem 5. Let K be a Ω_1-structure. For all formulas A, B, then

1. If A is an instance of a propositional tautology then, $K \vDash A$
2. If $K \vDash A$ and $K \vDash A \rightarrow B$, then $K \vDash B$
3. $K \vDash \Box_i(A \rightarrow B) \rightarrow (\Box_i A \rightarrow \Box_i B)$, $i = 1, 2, \ldots, n$
4. $K \vDash \Box_i A \rightarrow \Box_i \Box_i A$, $i = 1, 2, \ldots, n$
5. $K \vDash \Box_i A \rightarrow A$, $i = 1, 2, \ldots, n$
6. If $K \vDash A$ then $K \vDash \Box_i A$, $i = 1, 2, \ldots, n$

Theorem 6. If K is an interpretation for Ω_1 and $w \in W$. Then:

1. $K,w \vDash A \Leftrightarrow K,w \nvDash {}_{\mathbb{T}}A$;
2. $K,w \nvDash A \Leftrightarrow K,w \vDash {}_{\mathbb{T}}A$;
3. $K,w \nvDash A^\circ \Leftrightarrow K,w \vDash A$ and $K,w \vDash \neg A$;
4. $K,w \nvDash A \Leftrightarrow K,w \vDash A) = 0$ and $K,w \vDash \neg A$;
5. $K,w \nvDash A^\circ \Leftrightarrow K,w \vDash (\neg A)^\circ$;
6. $K,w \vDash A \Leftrightarrow K,w \vDash A$ or $K,w \nvDash \neg A$;

Lemma 1. $\Gamma \vdash A \Rightarrow \Gamma \vDash A$.

Proof. By induction on the length of the deductions of A from Γ.

Lemma 2. Each non-trivial set of formulas is contained in a non-trivial maximal set. There are non-trivial inconsistent sets.

Proof. First part is like in the classical case. As in Ω_1 the scheme $(A \wedge \neg A) \to B$ is not valid, we can deduce immediately that the set $\{A, \neg A\}$ is inconsistent and non trivial, when A is a propositional variable.

Lemma 3. Every non-trivial maximal set Γ has a model.

Proof. We define a function $\varphi: F \to 2$ as follows: to each formula A, if A belongs to Γ, we put $\varphi(A) = 1$; if A does not belong to Γ, we set $\varphi(A) = 0$. We show then that conditions $1 - 7$ of definition of interpretation are satisfied.

Theorem 7. (Strong Completeness). $\Gamma \vDash A \Rightarrow \Gamma \vdash A$.

Proof. Similar to the classical one, by employing the strong negation ${}_{\mathbb{T}}$ (and not the operator $\neg$).

Corollary 1. (Weak Completeness). $\vDash A \Rightarrow \vdash A$. As a consequence, $\Gamma \vDash A \Leftrightarrow \Gamma \vdash A$.

Theorem 8. There are inconsistent (but non trivial) sets of formulas which has models. (Γ has model $\Leftrightarrow \Gamma$ is non trivial)

Observation. The first (or the second) condition of the Theorem 2 entails 1) e 3) of the definition of interpretation. In fact, suppose that $K,w \vDash A \Leftrightarrow K,w \nvDash {}_{\mathbb{T}}A$ or $K,w \nvDash A$ $\Leftrightarrow K,w \vDash {}_{\mathbb{T}}A$; so, if $K,w \vDash A$, then $K,w \nvDash {}_{\mathbb{T}}A$ and, consequently, $K,w \nvDash \neg A$. On the other hand, if $K,w \vDash B^\circ$ and $K,w \vDash A \to B$ and $K,w \vDash A \to \neg B$ and if $K,w \vDash A$, we have $K,w \vDash B^\circ$ and $K,w \vDash B$ and $K,w \vDash \neg B$; therefore, $K,w \vDash B$ and $K,w \vDash {}_{\mathbb{T}}B$, which it is a contradiction. We observe that in general, the value of an interpretation K for a given formula is not determined by the values of the values of K for the compound atomic formulas. Similar results on paracompleteness can be obtained but we will omit details.

Theorem 9. The propositional counterpart of Ω_n is decidable.

Corollary 2. $(A \to A)$ is not a thesis of Ω_n.

Definition 7. Let $\Delta = \{A^\circ \in F | \vdash A\}$; Γ is called strongly non trivial if $\Gamma \cup \Delta$ is not trivial. Let us suppose now that Δ be the set $\{A^\circ \in F | A$ is not a propositional variable$\}$; Γ is called strictly non-trivial if $\Gamma \cup \Delta$ is not trivial.

Theorem 10. There are strongly non-trivial sets and strictly non-trivial sets.

It is observed that the semantics discussed is such that the Tarski's criterion for truth (T) is valid. In fact, if A is a formula and $[A]$ its name, we have clearly:

$[A]$ is true (in an interpretation) iff A.

In a certain sense, the semantics proposed for Ω_n constitutes a generalization of the traditional semantics.

The preceding considerations is extended to the remaining calculi Ω_n ($1 \leq n < \omega$) as well as to the predicate calculi Ω_n^* and predicate calculi with equality $\Omega_n^=$ ($1 \leq n < \omega$). A similar semantics can be applied to Ω_ω and to Ω_ω^* and $\Omega_\omega^=$.

We also observe that the semantic of Ω_n extends the classic one. In general, the paraconsistent semantics generalizes the classic one. As a consequence, there are Tarskian "alternatives" for truth theory (by Tarski). In this sense, the paraconsistent logic becomes a start point for a dialectics of classic doctrine of truth of logicity.

4   Conclusions

The paper is somewhat highly technical, but it shows that it is possible a logic in which it is possible to deal with inconsistent and paracomplete knowledge. Our main point is that inconsistency and paracomplete is a natural phenomenon in the world. We are often faced with conflicting advice. In such situations, most human beings are able to make an appropriate choice. The design of very large knowledge bases poses a similar problem. Disagreements between agents in the domain of interest may lead to the construction of a knowledge base that is inconsistent. Worse still, the very fact that it is inconsistent may emerge only much later, after the expert system has been in use for a significant period of time. Thus it is clear that formal methods are needed to handle this problem. The system proposed has provided means to reasoning about inconsistent and paracompleteness.

References

1. Abe, J.M., 1992, *Fundamentos da lógica anotada*, (Foundations of annotated logics), in Portuguese, Ph.D. thesis, University of São Paulo, São Paulo, 135p.
2. Abe, J.M., Some Aspects of Paraconsistent Systems and Applications, Logique et Analyse, 15, 83-96, 1997.
3. Cresswell, M.J., 1972, Intensional logics and logical truth, *Journal of Philosophical Logic,* 1, 2-15.
4. Cresswell, M.J., 1973, *Logics and Languages,* London, Methuen and Co.
5. Da Costa, N.C.A., 1974, On the theory of inconsistent formal systems, Notre Dame J. of Formal Logic, 15, 497-510.
6. Da Costa, N.C.A., J.M. Abe, and V.S. Subrahmanian, 1991, Remarks on annotated logic, *Zeitschr. f. Math. Logik und Grundlagen d. Math.*, 37: 561-570.
7. Fagin, R., J.Y. Halpern, Y. Moses & M.Y. Vardi, 1995, *Reasoning about knowledge*, The MIT Press, London.
8. Fischer, M.J. and N. Immerman, 1986, Foundation of knowledge for distributed systems. In J.Y. Halpern (Ed.), *Theoretical Aspects of Reasoning about Knowledge: Proc. Fifth Conference,* pp. 171-186. San Francisco, Cal.: Morgan Kaufmann.

9. Halpern, J.Y. and R. Fagin, 1989, Modeling knowledge and action in distributed systems. *Distributed Computing* 3(4), 159-179.
10. Halpern, J.Y. and Y. Moses, 1990, Knowledge and common knowledge in a distributed environment, *Journal of the ACM 37(3),* 549-587.
11. Lipman, B.L., 1992*, Decision theory with impossible possible worlds.* Technical report Working paper, Queen's University.
12. Lipman, B.L., 1992, *Logics for nonomnisiscient agents: An axiomatic approach.* Technical report Working paper 874, Queen's University.
13. Lipman, B.L., 1994, An axiomatic approach to the logical omniscience problem. In R. Fagin (Ed.), *Theoretical Aspects of Reasoning about Knowledge: Proc. Fifth Conference,* pp. 182-196. San Francisco, Cal.: Morgan Kaufmann.
14. McCarthy, J., 1979, *Ascribing mental qualities to machines*. Technical report CRL 90/10, DEC-CRL.
15. Parikh, R. and R. Ramamujan, 1985, Distributed processing and the logic of knowledge. In R. Parikh (Ed.), *Proc. Workshop on Logics of Programs,* pp. 256-268.
16. Rosenschein, S.J., 1985, Formal theories of AI in knowledge and robotics. *New Generation Computing* 3, 345-357.
17. Rosenschein, S.J. and L.P. Kaelbling, 1986, The synthesis of digital machines with provable epistemic properties. In J.Y. Halpern (Ed.), *Theoretical Aspects of Reasoning about Knowledge: Proc. Fifth Conference,* pp. 83-97. San Francisco, Cal.: Morgan Kaufmann.

A Connectionist Model for Predicate Logic Reasoning Using Coarse-Coded Distributed Representations

Sriram G. Sanjeevi[1,*] and Pushpak Bhattacharya[2,**]

[1] Dept. of Comp. Science & Engg.,N.I.T. Warangal, Warangal 506004, 91-0870-2430440
sgsanjeevi@yahoo.com
[2] Dept. of Comp. Science & Engg.,I.I.T. Bombay, Mumbai 400076, 91-22-5767718
pb@cse.iitb.ac.in

Abstract. In this paper, we describe a model for reasoning using forward chaining for predicate logic rules and facts with coarse-coded distributed representations for instantiated predicates in a connectionist frame work. Distributed representations are known to give advantages of good generalization, error correction and graceful degradation of performance under noise conditions. The system supports usage of complex rules which involve multiple conjunctions and disjunctions. The system solves the variable binding problem in a new way using coarse-coded distributed representations of instantiated predicates without the need to decode them into localist representations. Its performance with regard to generalization on unseen inputs and its ability to exhibit fault tolerance under noise conditions is studied and has been found to give good results.

1 Introduction

Traditionally reasoning systems using predicate logic have been implemented using symbolic methods of artificial intelligence. Connectionist methods of implementation of reasoning systems describe an alternative paradigm. Among the connectionist systems they use two types of representational schemes. They are 1) localist and 2) distributed representational schemes.

Localist representational schemes represent each concept with an individual unit or neuron. In the distributed representational schemes [3] each unit or neuron is used in representation of multiple concepts and multiple units or neurons are used to represent a single concept. In the literature, some localist methods for reasoning using connectionist networks have been described. The connectionist inference system *SHRUTI* [1], [5] described a localist method where temporal synchrony was used to create bindings between variables and entities they represent. *CONSYDERR* [2] described another localist method for variable binding and forward reasoning. It used an assembly or a set of interconnected nodes to represent each predicate $p(x_1.....x_k)$. Since, these systems used localist representations, advantages of distributed representations are not obtainable by them and hence the motivation for a distributed representation based reasoning system.

---

* Asst. Professor

** Professor

2   Rule and Fact Base

Our system represents and reasons with predicate logic rules and facts. Following are rules and facts we use.

1. *give(x, y, z)* —> *own(y, z);*
2. *own(y, z)* —> *donate(y, z);*
3. *own(y,z)* ∧ *wantstobuy(w, z)*∧*(hasrequiredmoney(w, m)* ∨ *hasgoodcreditrating(w))* —> *cansell(y,w,z);*
4. *give(John,Mary,Book-1);*
5. *give(John,Chris,Book-2);*
6. *wantstobuy(Walter,Book-2);*
7. *hasrequiredmoney(Walter,Money);*
8. *hasgoodcreditrating(Walter);*

Our system uses the above rule base and makes inferences shown below.

1. *own(Mary,Book-1);*
2. *donate(Mary,Book-1);*
3. *own(Chris,Book-2);*
4. *cansell(Chris,Walter,Book-2);*

3   Forward Reasoning Using Connectionist System

In this paper we see how to accomplish the forward reasoning for predicate calculus facts and rules using neural networks which operate on coarse coded distributed representations. Each fact of predicate p_i is represented by a vector $\mathbf{v_{ij}}$. The vector $\mathbf{v_{ij}}$ is a k dimensional vector which stores the coarse coded representation of predicate fact. The different instantiations of predicate p_i are each represented by separate vector $\mathbf{v_{ij}}$ where j varies from *1* to *m* where *m* is the number of vectors in predicate p_i table.

We describe here, briefly with an example how forward reasoning using localist representations [6],[7] is made using a connectionist system. Let us consider the *rule 1:give(x,y,z)*—> *own(y,z)* from the knowledge base. The localist pattern for the LHS of rule 1 can be written as *0001 001   001 001   1*. The first 4 bit value denotes the predicate *give*, the next 3 bit value denotes an object getting bound to variable *x* '*John*', the next 3 bit value denotes an object getting bound to variable *y*, '*Mary*' and the next value denotes, '*Book-1*'. The last bit indicates the truth value of predicate *give*. This instantiation will activate rule *1* and make variables on the right hand side of the rule '*y*' and '*z*' be assigned the values '*001*' and '*001*' representing the objects '*Mary*' and ' *Book-1*' respectively. Because of the rule activation the localist pattern representation for RHS will be *0010 001 001 1* denoting *own(Mary,Book-1)*. This triggers the rules whose left hand sides match RHS of rule 1 and through this   forward chaining, forward reasoning using localist representations is accomplished. In *Table 1* and *2* below we show samples of localist vectors for some of the predicates in the rule base.

Table 1. Shows a sample of localist tuples used by predicate *give*

S.No of Tuple	Predicate 'id' code	Localist Value of x	Localist Value of y	Localist Value of z	Truth Value of Predicate
215	00001000000	0000100000	0000100000	0000010000	00001

Table 2. Shows a sample of localist tuples used by predicate *Cansell*

S.No of Tuple	Predicate 'id' code	Localist Value of y	Localist Value of w	Localist Value of z	Truth Value of Predicate
Tuple no. 46	00000001000	0000000010	0000100000	0000010000	00001

3.1 Obtaining Coarse-Coded Distributed Representations from Localist Representations

Consider the following tuple from the localist representation table of predicate give(x,y,z), ' *00001000000 0000000001 0000000010 0000000010 00001*'.

   We view the above vector as being kept in overlapping coarse zones of length of 4 consecutive bits and encode the zone as *1* if there is at least one *1* bit in that zone or else as *0*. We then consider next coarse zone and encode it as *1* or *0* following above method. We do this process left to right starting from the left most bit. We do this encoding process for above localist tuple to get the following coarse-coded tuple

   ' *01111000000 0000001111 0000011110 0000011110 01111*'.

Coarse-coding can be applied when the number of 1's in the original string is sufficiently sparse. If the number of 1's in the original string is not sufficiently sparse then coarse-coded string when decoded will not yield the original string. This is the reason we have chosen a 5 bit string to denote the truth value of predicate(in which first 4 bits were kept as zeros). The reason the coarse-coding could be applied successfully to our reasoning problem is that localist representations of instantiated predicates were sufficiently sparse with regard to distribution of 1's. Coarse-coding increases the information capacity [4] by increasing the number of units active at a time compared to localist codes which have sparsely populated 1's. The amount of information conveyed by a unit that has a probability *p* of being '*1*' is $-p\log(p) - (1-p)\log(1-p)$. We obtain the coarse-coded representations of tuples for all the *predicates* in the *rule base* using the above described method. We show here a sample of coarse-coded representations of the tuples for predicates *give* and *cansell* in the rule base.

Table 3. Shows a sample of coarse-code representation of data tuples used by predicate *Give*

S.No of Tuple	Predicate 'id' code	Value of x	Value of y	Value of z	Truth Value of Predicate
Tuple no. 215	01111000000	0111100000	0111100000	0011110000	01111

Table 4. Shows a sample of coarse-code representation of data tuples used by predicate *Cansell*

S.No of Tuple	Predicate 'id' code	Value of y	Value of w	Value of z	Truth Value of Predicate
Tuple no. 46	00001111000	0000011110	0111100000	0011110000	01111

3.2 Organization of Neural Networks in the Connectionist Reasoning System

The neural networks shown accomplish the forward reasoning using the coarse-coded tuples. They generate inferences by firing rules from the rule base. Consider the neural networks shown in figure 1. When impressed on its inputs with one of the vectors v_g from the predicate table *give* the network 1 generates on its outputs a vector v_o from the predicate table *own*. This way the rule *give(x,y,z)* —> *own(y,z)* was processed. This in turn impresses on the inputs of network 2 to generate a vector v_d on its outputs. This processed the rule *own(y,z)* —> *donate (y,z).* These vectors are in coarse-coded form and denote a predicate fact. So we see the rules 1 and 2 are getting activated in a forward chaining fashion.
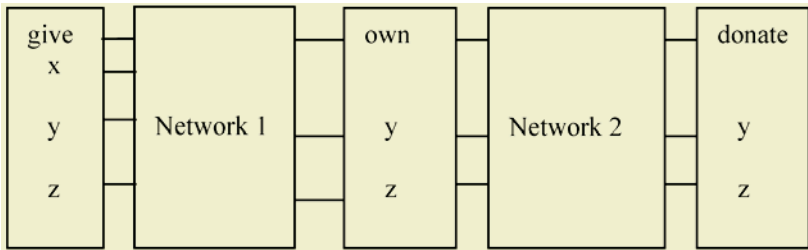
Fig. 1. Neural Networks for processing rules 1 and 2

3.3 Variable Binding During Processing of the Complex Rule

Consider the following complex rule which involves multiple conjunctions.

own(y,z) Λ *wantstobuy(w,z)* Λ (*hasrequiredmoney(w,m)* $\lor$ *hasgoodcreditrating(w)*) —> *cansell(y,w,z);*

This rule is processed by the connectionst architecture shown in figure 2. We use the vectors v_o, v_w, v_h and v_g from the predicate tables *own, wantstobuy, hasrequiredmoney* and *hasgoodcreditrating* respectively. Though these are coarse-coded tuples their structure has the format of predicate code p, value of variable *1*, variable *2,....*variable n and predicate truth value T / F. These constituents are distinguishable and hence they can be used directly. These constituents are in coarse-coded form. We use these constituents to implement the complex rule under consideration. Component z is taken from both v_o and v_w and given as inputs to network 5. This network generates truth value T or F depending on whether the values of variable z given to it are equal or unequal respectively. Similarly component w is taken both from vectors, v_w and v_h and given as inputs to network 6. This network generates truth value of T or F depending upon whether the values of variable w given to it are equal or unequal. Similarly component w is taken both from vectors, v_w and v_g and given as inputs to network 7. This network generates truth value of T or F depending upon whether the values of variable w given as inputs to it are equal or unequal. These truth values from outputs of network 6 and network 7 are given as inputs of network 8 which outputs T if either or both (inclusive or) of the truth values on its inputs are true else outputs F. The truth values from outputs of network 5 and network 8 are given as inputs to network 9 which outputs T if both of the truth values on its inputs are true else outputs F. The predicate code components of the vectors v_o, v_w, v_h and v_g are given as inputs

to neural network *10* which outputs predicate code *p* for *cansell*. The values of *y, w* and *z* are passed on to the output lines as shown in figure 2 from the vectors v_o, v_w and v_o respectively. If network *9* output is 'T' the values of *y, w* and *z* are accepted as belonging to vector v_c of the predicate table of *cansell*. Using this method we had processed a complex rule. The *variable binding problem* has been solved as described above while processing the above complex rule which is involving multiple conjunctions and a disjunction. The disjunction was meaning an inclusive OR operation. Our task was to check whether the variable *w* belonging to v_w and *either or both* of v_h and v_g are binding to same value. We had accomplished these with networks 6, 7 and 8 respectively. Similarly, we had to check whether variable *z* belonging to v_o and v_w are bound to same value. We had accomplished this with network 5.
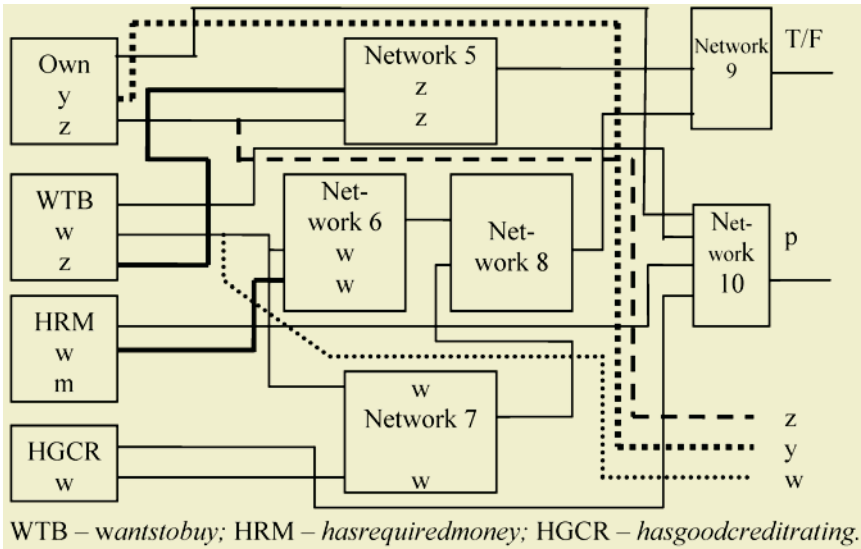
Fig. 2. Neural Networks for processing rule 3

We have solved the *variable binding problem* faced while implementing multiple conjunctions and a disjunction in a complex rule using *coarse-coded representations* without the need to decode them into localist representations. We have accomplished the variable binding task here using a *divide and conquer* strategy and distributed the total task to a set of neural networks which together accomplished the same. This approach enables us to perform variable binding in more complex rules which involve more number of conjunctions and disjunctions in them.

4   Testing

The performance of the above coarse-coded reasoning system was compared   for error tolerance under noise conditions with a localist representation based reasoning system (which was having identical number of input, hidden and output units for its neural networks). Subset of the training patterns were made test patterns after introducing *1* bit error at a random location in each pattern. In the test 1, neural network 1

with 46 input units, 40 hidden units and 36 output units was trained with 216 patterns. A subset consisting of 108 of them were made test patterns. In the test 2, neural network 11 (implementing rule 1 for larger data) with 66 input units, 56 hidden units and 52 output units was trained with 750 patterns. A subset consisting of 300 of the training patterns were made test patterns. Results are as shown in table 5.

Table 5. Shows the details of tests

	No. of Training Patterns	No. of Test Patterns	No. of Patterns Corrected	No. of Patterns not Corrected
Localist reasoning system	216	108	60	48
	750	300	207	93
Coarse-coded reasoning system	216	108	89	19
	750	300	274	26

In tests 3 and 4, Neural network 11 was tested with coarse-coded patterns for generalization on unseen test patterns as shown in table 6.

Table 6. Shows the details of tests

Tests	Training Patterns	Unseen Test Patterns	Correctly Generalized	Not Correctly Generalized
3	650	350	342	8
4	700	300	300	0

5   Conclusions

We have tested a connectionist forward chaining reasoning system using distributed coarse-coded representations on a given reasoning task. The system has successfully performed the given reasoning task. The coarse-coded reasoning system was found to be much more fault tolerant to errors compared to localist reasoning system as was indicated by tests performed. The system has displayed good generalization ability on unseen test patterns. We have solved the variable binding problem faced while implementing multiple conjunctions and a disjunction in a complex rule using coarse-coded representations without the need to decode them into localist representations.

References

1. Shastri, L. (1999). Advances in SHRUTI: a neurally motivated model of relational knowledge representation and rapid inferencing using temporal synchrony *Applied Intelligence,* 11 (1), 79-108 .
2. Sun,R. (1992) On variable binding in connectionist networks. *Connection Science*, 4, 93-124.
3. T.J. Van Gelder, Defining ' distributed representation', *Connection Science* 4 (3 and 4) (1992) 175-192.
4. Hinton,G.E., J.L. McClelland and D.E.Rumelhart. (1986).Distributed representations. In D.E.Rumelhart and J.L.McClelland, editors, *Parallel Distributed Processing,* Vol.1. Cambridge,MA. MIT Press.

5. Shastri, L. and V. Ajjanagadde. (1993) From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings. *Behavioral and Brain Sciences,* 16(3), 417-494.
6. Browne, A., & Sun, R. (1999). Connectionist variable binding. *Expert systems: The International Journal of Knowledge Engineering and Neural Networks*, 16(3), 189- 207.
7. A. Browne, R.Sun. Connectionist inference models, *Neural Networks* 14 (2001)   1331-1355
8. Russel, S & Norvig, P. *Artificial Intelligence a Modern Approach,* Second Edition, Delhi: Pearson Education, 2003.
9. Haykins, S. *Neural Networks, a comprehensive foundation,* Second edition, New Jersey: Prentice hall,1999.

A General-Purpose Forward Deduction Engine for Modal Logics

Shinsuke Nara, Takashi Omi, Yuichi Goto, and Jingde Cheng

Department of Information and Computer Sciences
Saitama University , Saitama, 338-8570, Japan
{nara,omi,gotoh,cheng}@aise.ics.saitama-u.ac.jp

Abstract. Many advanced knowledge-based systems with purposes of discovery and/or prediction, such as anticipatory reasoning-reacting systems, legal reasoning systems, geographical information system, and so on, require a forward deduction engine for modal logics as an important component. However, until now, there is no forward deduction engine with the general-purpose for modal logics. This paper presents a general-purpose forward deduction engine for modal logics. We show some essential requirements and functions for a general-purpose forward deduction engine for modal logics, present our implementation of the forward deduction engine, and show possible applications based on our forward deduction engine.

1 Introduction

A forward deduction engine is an automated deduction system which deduces new conclusions from given premises automatically by applying logical inference rules (inference rules for short) to the premises and previously deduced results repeatedly until some previously specified termination conditions are satisfied [1]. Many advanced knowledge-based systems with purposes of discovery and/or prediction, such as anticipatory reasoning-reacting systems [2, 3], legal reasoning systems [9], geographical information system, agent-oriented systems, theorem finding systems and so on, require a forward deduction engine for modal logics as an important component. However, until now, there is no forward deduction engine with the general-purpose for modal logics.

This paper presents a general-purpose forward deduction engine for modal logics. We show some essential requirements and functions for a general-purpose forward deduction engine for modal logics, present our implementation of the forward deduction engine, and show possible applications based on our forward deduction engine.

2 Automated Forward Deduction Based on Modal Logics

Automated forward deduction is a process of deducing new conclusions from premises automatically by applications inference rules to the premises and pre-

viously deduced conclusions repeatedly until some previously specified termination conditions are satisfied [1]. The process of automated forward deduction is as follows.

1. Initialization process: premises, inference rules and termination conditions are taken as an input.
2. Taking an inference rule up process : it takes an inference rule up from the set of inference rules.
3. Matching process : it judges whether the inference rule which is took up at taking an inference rule up process can apply to premises and previously deduced conclusions which are took up from the set of premises or not.
4. Deduction process: it deduces a conclusion from the inference rule and the premises and previously deduced conclusions which are succeed at matching process.
5. Duplication checking process: it compares the conclusion deduced at deduction process with premises and previously deduced conclusions to check whether it is duplicate or not.
6. Adding process: it adds the conclusions which was judged to be new at duplication checking process to the set of premises.
7. Outputting process: it outputs all deduced conclusions.

Difference between automated forward deduction and automated forward deduction based on modal logics is whether it based on modal logics or not.

Modal logic is the logic of necessity and possibility, of 'must be' and 'may be' [4, 7]. Necessity and possibility may be interpreted in various ways. For examples, necessity is interpreted as necessity truth in alethic modal logic [4]; necessity is interpreted as moral or normative necessity in deontic logic [6]; necessity is interpreted as what is known or believed to be true in epistemic logic [8]; necessity is interpreted as to what always has been or to what henceforth always will be true in temporal logic [11].

One of the difference between modal logics and nonmodal logics is whether they contains modal operators which denotes certain modal expressions as their vocabulary or not. Moreover different modal logic has different kind of modal operators, and the number of modal operators is also different in each modal logics.

Other difference among between modal logics and nonmodal logics is whether they contains inference rules concerning with a certain modal operator or not. Furthermore different kind of inference rules are defined in each modal logics, and the number of inference rules defined in each modal logics are also different.

3   Requirements and Functions for General-Purpose Forward Deduction Engine for Modal Logics

We define requirements for a general-purpose forward deduction engine for modal logics. The requirements are classified by two elements. One is requirements for

a general-purpose forward deduction engine. The other is requirements for a general-purpose forward deduction engine for modal logics. We name the requirements for a general-purpose forward deduction engine as the general requirements, and name the requirements for a general-purpose forward deduction engine for modal logics as the particular requirements.

In this paper, we define the particular requirements because Cheng already defined the general requirements [1]. Briefly, we describe the general requirements; a general-purpose forward deduction engine must deduce conclusions by applying inference rules to premises and previously deduced conclusions; it must be able to take premises as input from users; it should not serve conclusions which are duplicate to premises and previously deduced conclusions; it should terminate its execution by certain termination condition; it should provide a way to customize its termination condition.

Requirement 1 *A general-purpose forward deduction engine for modal logics must provide a way to use various modal operators defined in each modal logics which users can satisfy.*

Requirement 2 *A general-purpose forward deduction engine for modal logics must provide a way to use various inference rules defined in each modal logics which users can satisfy.*

A general-purpose forward deduction engine for modal logics must provide modal operators and inference rules which are defined in modal logics because differences between a general-purpose forward deduction engine and it for modal logics are whether it can use modal operators and the inference rules.

Moreover a general-purpose forward deduction engine for modal logics provide a way to use various modal operators and various inference rules defined in each modal logics for users, because there are a lot of modal logics and modal operators and inference rules are defined in each modal logics. We can not know what modal logics users want to use, i.e, we can not know what modal operators and the inference rules users want to use, therefore we can not satisfy the requirements by implementing some specified modal operators and inference rules.

Furthermore, a general-purpose forward deduction engine for modal logics should be able to be used even when new extensions of modal logic are proposed because such new extensions of modal logic must have advantage for other modal logics which was proposed before and users must want to use the logics.

In order to satisfy **Requirement 1**, a general-purpose forward deduction engine for modal logics provide enough number of modal operators for users and we provide **customizing modal operators function** which make correspondence between a certain modal operator and characters. one of the differences among various modal logics is the number of modal operators because modal operators does not have fixed interpretations; same operators are interpreted different way in each logic system [12].

In order to satisfy **Requirement 2**, we design **customizing inference rules function** where users customize inference rules. I guess that we can not

provide enough number of inference rules and enough kind of inference rules which users satisfy by a way of implementing some specified inference rules because a lot of modal logics exist and a lot of inference rules were also defined in these modal logics, moreover we can not know which inference rules users want to use, furthermore, even if new extensions of modal logic is provide, users can not use forward deduction engine with it immediately, therefore we should provide a way of customizing inference rules where users can make inference rules which user want to use.

From two of the general requirements which is that a general-purpose forward deduction engine should terminate its execution by certain termination condition and it should provide a way to customize its termination condition, we should define a termination condition for forward deduction engine for modal logics. Cheng proposed a method which is to measure a degree of nested a implication which is a kind of logical operators to limit the range of candidates for "new knowledge" [1]. The method can be used as termination condition for a forward deduction engine [1].

We modify the methods suitable for our purpose because the method may not work well as termination condition of a forward deduction engine for modal logics. In modal logics, there is an inference rule which join modal monadic operators to a formula. If the forward deduction engine does not have the limit method for joining modal monadic operators to a formula, then the forward deduction engine may not terminate its execution.

Definition 1 *Let $deg_\Delta(A) = n$ denote that the degree of nested a monadic or binary operator Δ in a formula A is n. A and B are formulas.*

1. *If there is no occurrence of Δ in A, then $deg_\Delta(A) = 0$*
2. *If $*$ is a monadic operator other than Δ, then $deg_\Delta(*A) = deg_\Delta(A)$.*
3. *If $\circ$ is a binary operator other than Δ, then*
   $deg_\Delta(A \circ B) = max(deg_\Delta(A), deg_\Delta(B))$.
4. *$deg_\Delta(\forall xA) = deg_\Delta(A)$*
5. *If Δ is a monadic operator then $deg_\Delta(\Delta A) = deg_\Delta(A) + 1$*
6. *If Δ is a binary operator then $deg_\Delta(A\Delta B) = max(deg_\Delta(A), deg_\Delta(B)) + 1$*

Our termination condition is to deduce all formulas which can be deduced from given premises and inference rules and which are not exceeding some degree of some operators which a user specifies.

4   Implementation and Experiments

According to the functions defined in section 3, we implement a general-purpose forward deduction engine for modal logics by improving an automated forward deduction system for general-purpose entailment calculus (EnCal for short) [1]. EnCal is a forward deduction engine based on nonmodal logics. Our forward deduction engine is based on propositional modal logics.

In order to provide enough number of modal operators to users, we investigate the number of operators in 7 modal logics, and we find 15 monadic operators and

6 binary operators, therefore our forward deduction engine provides 31 binary operators and monadic operators, moreover in our forward deduction engine, a user defines which characters correspond to logical and/or modal binary and/or monadic operators by **customizing modal operators function**.

In order to implement a **customizing inference rules function**, we investigate inference rules. An inference rule is formulated from some well formed formula schemata which are patterns of formulas [10]. In these well formed formula schemata, not less than one well formed formula schemata are premises of the inference rule and not less than one well formed formula schemata are conclusions of the inference rule. Applying an inference rule is pattern matching a well formed formula schemata which is a premise of the inference rule and formula which is a premise of a forward deduction engine and a conclusion is deduced by applying results of pattern matching. In our forward deduction engine, inference rules are inputted as a series of some well formed formula schemata and by applying pattern matching to well formed formula schemata and formula, our forward deduction engine deduces conclusions corresponding to each inference rules.

In order to terminate execution of forward deduction engine with limit of degree, our forward deduction engine takes operators and limit of degree corresponding to the operators and our forward deduction engine counts degree of conclusions which are deduced by our forward deduction engine. If the degree of conclusions is not less than the limit of degree, then they were not added into a set of premises. If the degree of conclusions is not more than the limit of degree, then they were added into a set of premises. Our forward deduction engine deduces conclusions until all conclusions which is not more than the limit of degree.

We deduced some set of logical theorems of 6 systems of propositional modal logic from axioms of each system. The systems of which we deduced logical theorems were the systems K which is a minimal alethic modal logic , D which is a minimal deontic logic, K_t is a minimal temporal logic, $S4$ which is a minimal doxastic logic and $K4$ which is minimal epistemic logic. In our experiments, we deduced all formulas of which $deg_\rightarrow(A)$ are not exceeding 2 and $deg_\diamond(A)$ are not exceeding 2 in each logic system. The number of deduced conclusions of K are 215, D are 751, K_t are 1227, $S4$ are 1154. $K4$ are 278. In these logic systems, one or two modal operators are defined and three or four inference rules are defined. If we can not specify "$deg_\diamond(A)$ are not exceeding 2" as termination condition, Our forward deduced engine with these logic systems can not terminate its execution because these logic systems contains necessitation or temporal generalization which are inference rules joining modal monadic operators to a formula.

5   Possible Applications

Our forward deduction engine can be a useful component for developers of advanced knowledge-based systems which requires conclusions including modal expressions. For example, in anticipatory reasoning-reacting systems which is a

new type new type of reactive system and which can detects and predicts omens of attacks anticipatory, our forward deduction engine based on temporal relevant logic can be used as a anticipatory reasoning engine [3]. If legal reasoning systems which deduce legal conclusions for given case by statutory laws and old judicial precedents [9] contain our forward deduction engine based on deontic logic as a judicial precedents generating engine then the legal reasoning systems can automatically generate new judicial precedents from old judicial precedents and statutory laws.

On the other hand, our forward deduction engine can be a useful tool for researchers. Our forward deduction engine can be used as an important component in theorem finding systems in various fields. Our forward deduction engine also a useful tool for logicians which propose new logic systems. the logicians can easily make an example to show how to use the logic systems. logicians can get a lot of logical theorems in various logic systems easily and uses the logical theorems for various purposes.

6   Concluding Remarks

In this paper, we defined essential requirements for a general-purpose forward deduction engine for modal logics. We also designed essential functions for a general-purpose forward deduction engine for modal logics. We presented our implementation of our general-purpose forward deduction engine for modal logics too.

Our forward deduction engine can be used in various advanced information systems, such as legal reasoning systems, anticipatory reasoning-reacting systems and agent-oriented systems in various fields because our forward deduction engine deduce conclusions including various modal expressions which were required in various fields by customizing inference rules and customizing modal operators. Our forward deduction engine also can be a useful tool for researchers because our forward deduction engine can be used as an important component in theorem finding systems in various fields.

As future works, we will design and implement a general forward deduction engine based on first order predicate logics because, for practical use, a forward deduction engine should be enriched vocabulary of modal logics. We will also improve performance of a forward deduction engine because a forward deduction engine is useless at all if it can not satisfy a requirement of acceptability of time [5]. We will also have to make an example to show effectiveness of our forward deduction engine.

References

1. Cheng, J.: EnCal: An automated forward deduction system for general–purpose entailment calculus. In: Terashima, N., Altman, E., (eds.): Advanced IT Tools, IFIP World Conference on IT Tools, IFIP96 – 14th World Computer Congress. Chapman & Hall (1996) 507–517

2. Cheng, J.: Anticipatory reasoning-reacting systems. In: Proc. International Conference on Systems,Development and Self-organization (ICSDS'2002), Beijing, China (2002) 161–165
3. Cheng, J.: Temporal relevant logic as the logical basis of anticipatory reasoning-reacting systems. In: Dubois, D.M., (ed.): Computing Anticipatory Systems: CASYS 2003 - Sixth International Conference. AIP Conference Proceedings (2004) 362–375
4. Cresswell, M.J.: Modal logic. In: Goble, L., (ed.): The Blackwell Guide to Philosophical Logic. Blackwell (2001) 136–158
5. Goto, Y., Nara, S., Cheng, J.: Efficient anticipatory reasoning for anticipatory systems with requirements of high reliability and high security. In: Dubois, D.M., (ed.): International Journal of Computing Anticipatory Systems. Volume 14., CHAOS (2004) 156–171
6. Hilpinen, R.: Deontic logic. In: Goble, L., (ed.): The Blackwell Guide to Philosophical Logic. Blackwell (2001) 159–182
7. Hughes, G.E., Cresswell, M.J.: A New Introduction to Modal Logic. Routledge (1996)
8. Meyer, J.J.C.: Epistemic logic. In: Goble, L., (ed.): The Blackwell Guide to Philosophical Logic. Blackwell (2001) 183–202
9. Nitta, K., Ohtake, Y., Maeda, S., Ono, M., Ohsaki, H., Sakane, K.: Helic-II: A legal reasoning system on the parallel inference machine. In: Proceedings of the International Conference on Fifth Generation Computer Systems. Volume 2. (1992) 1115–1124
10. Nolt, J., Rohatyn, D.: The propositional calculus. In: Aliano, J., (ed.): Schaum's Outline of Theory and Problems of Logic. McGraw–Hill (1988) 39–65
11. Venema, Y.: Temporal logic. In: Goble, L., (ed.): The Blackwell Guide to Philosophical Logic. Blackwell (2001) 203–223
12. Popkorn, S.: The modal language. In: First Step in Modal Logic. Cambridge University press (1994) 13–19

A Deductive Semantic Brokering System

Grigoris Antoniou[1], Thomas Skylogiannis[1], Antonis Bikakis[1], and Nick Bassiliades[2]

[1] Computer Science Department, University of Crete, Greece
Institute of Computer Science, FORTH, Greece
{antoniou,dogjohn,bikakis}@ics.forth.gr
[2] Department of Informatics, Aristotle University of Thessaloniki, Greece
nbassili@csd.auth.gr

Abstract. In this paper we study the brokering and matchmaking problem in the tourism domain, that is, how a requester's requirements and preferences can be matched against a set of offerings collected by a broker. The proposed solution uses the Semantic Web standard of RDF to represent the offerings, and a deductive logical language for expressing the requirements and preferences. We motivate and explain the approach we propose, and report on a prototypical implementation exhibiting the described functionality in a multi-agent environment.

1 Introduction

E-Commerce describes the revolution that is currently transforming the way business is conducted through the use of information technology, and in particular the World Wide Web. The 2nd generation of e-Commerce will be realized through the use of automated methods of information technology. Web users will be represented by *software agents*. According to [8], there is an increasing use of software agents for all the aspects of e-Commerce.

As software agents start to engage in e-commerce, new issues arise. Information must be organized in a way that is accessible by both humans and machines. Additionally, machines must be able to access, process and interpret the information in the same way. This vision is consistent with the Semantic Web initiative 4], which enriches the current Web through the use of machine-processable information about the meaning (semantics) of information content. This way, the meaning of displayed information is accessible not only to humans, but also to software agents.

The focus of the present work is on semantic-based electronic markets. Such markets help both service providers and requesters to match their interests. The key operations in such markets are to: (a) Identify appropriate services that satisfy user requirements; and (b) Select the best service based on the user's preferences.

How to address these questions using Semantic Web technology is the main focus of the present work. Previous works related to our approach include [10, 13, 7, 11 and 5]. Our work distinguishes itself from previous efforts in a novel combination of standard Semantic Web technology with a nonmonotonic rules language that allows one to express requirements and preferences.

2 Proposed Solution

2.1 General Approach

The three basic roles that we identify in our brokering system are the *Service Requester*, the *Service Provider*, and the *Broker*. Directory Facilitator is a service, which agents use to find each other and register the protocols and the ontologies that they use. The technical solution we provide is based on the following key ideas:

- Service requesters, service providers and brokers are represented by software agents that run on the JADE multi-agent platform.
- The requirements of the service requester are represented in defeasible logic, using rules and priorities. These requirements include both indispensable requirements that must be met for a service to be acceptable, and soft requirements (preferences) that can be used to select among the potentially acceptable offerings.
- The offerings or advertisements are represented in a certain semi-structured format using the Semantic Web standard language RDF for describing Web resources.
- The terminology shared by providers, requesters and brokers is organized in an ontology using the Semantic Web standard of RDF Schema.
- The broker is also a software agent and has special knowledge both for the declarative language and the advertisement format. It also has the ability to perform semantic checks to the information it receives.
- When the broker receives a request it matches the request to the advertisements by running the request specification against the available offerings, making use of information provided by the shared ontology, as required. Then the requester's preferences are applied to select the most suitable offering(s).
- For the persistent storage of advertisements, an RDF repository, and particularly ICS-FORTH RDF Suite [1], is used.

2.2 Description of Offerings

The offerings are described in RDF, the standard Semantic Web language for representing factual statements. This choice (a) supports interoperability among agents and applications, and (b) facilitates the easy publication, collection and combination in decentralized dynamic settings. The offerings are enriched through reference to shared ontologies, e.g. of the tourism domain or geographical terms. We assume that these ontologies are expressed in RDF Schema. We have chosen this language over the use of OWL because at present it is not clear how the deductive capabilities of OWL and rule systems can be combined. We could certainly use most features of OWL Lite, given that they can be expressed using rules 6].

2.3 Description of Requests and Preferences

We have chosen defeasible logic to represent requesters' requirements and preferences because it satisfies the above criteria. In particular,

- It is a formal language with well-understood meaning, thus it is also predictable and explainable.

- It is designed to be executable; implementations are described in [9]. It is also scalable, as demonstrated in the same paper, where it is shown that 100,000 rules can be processed efficiently, due to its low computational complexity.
- It is expressive, as demonstrated by the use of rules in various areas of information technology. In fact, among the logical systems, it is rule-based systems that have been best integrated in mainstream IT.
- Finally, it is suitable for expressing requirements and preferences in our setting. This is so, because it supports the natural representation of important features: (a) Rules with exceptions are a useful feature in our problem. For example, a general rule may specify acceptable offerings, while more specific rules may describe cases in which the general rule should not apply and the offering should not be accepted. (b) Priorities are an integral part of defeasible logic, and are useful for expressing user preferences for selecting the most appropriate offerings from the set of the acceptable offerings.

The expressing capabilities of defeasible logic are exhibited in the following example: "Antonis has the following preferences about his holiday travel package: He wants to depart from Athens and considers that the hotel at the place of vacation must offer breakfast. In addition, he would like either the existence of a swimming pool at the hotel to relax all the day, or a car equipped with A/C, to make daily excursions at the island. However, Antonis believes that if there is no parking area at the hotel, the car is useless, because he adds to him extra effort and fatigue. Lastly, if the tickets for his transportation to the island are not included in the travel package, he is not willing to accept it…". This verbal description of Antonis' hard requirements about acceptable offers can be modeled through the following defeasible logic rules:

r_1: *from(X,athens),includesResort(X,Y),breakfast(Y,true)* $\Rightarrow$ *accept(X)*
r_2: *from(X,athens),includesResort(X,Y),swimmingPool(Y,true)* $\Rightarrow$ *accept(X)*
r_3: *from(X,athens), includesService(X,Z), hasVehicle(Z,W),*
    vehicleAC(W,true) $\Rightarrow$ *accept(X)*
r_4: *includesResort(X,Y),parking(Y,false)* $\Rightarrow$ *~accept(X)*
r_5: *~includesTransportation(X,Z)* $\Rightarrow$ *~accept(X)*
        $r_4 > r_3, r_1 > r_4, r_2 > r_4, r_5 > r_1, r_5 > r_2, r_5 > r_3$

3 Brokering System Implementation

3.1 Multi-agent Framework

The agent framework we used for the development of our system is JADE (jade.cselt.it). JADE is an open-source middleware for the development of distributed multi-agent applications. It is Java-based and compliant with the FIPA specifications (www.fipa.org). It provides libraries for agent discovery, communication and interaction, based on FIPA standards.

From the functional point of view, JADE provides the basic services necessary to distributed peer-to-peer applications in the fixed and mobile environment. JADE allows each agent to dynamically discover other agents and communicate with them according to the peer-to-peer paradigm. From the application point of view, each agent is identified by a unique name and provides a set of services. It can register and

modify its services and/or search for agents providing given services, it can control its life cycle and, in particular, communicate with all other peers.
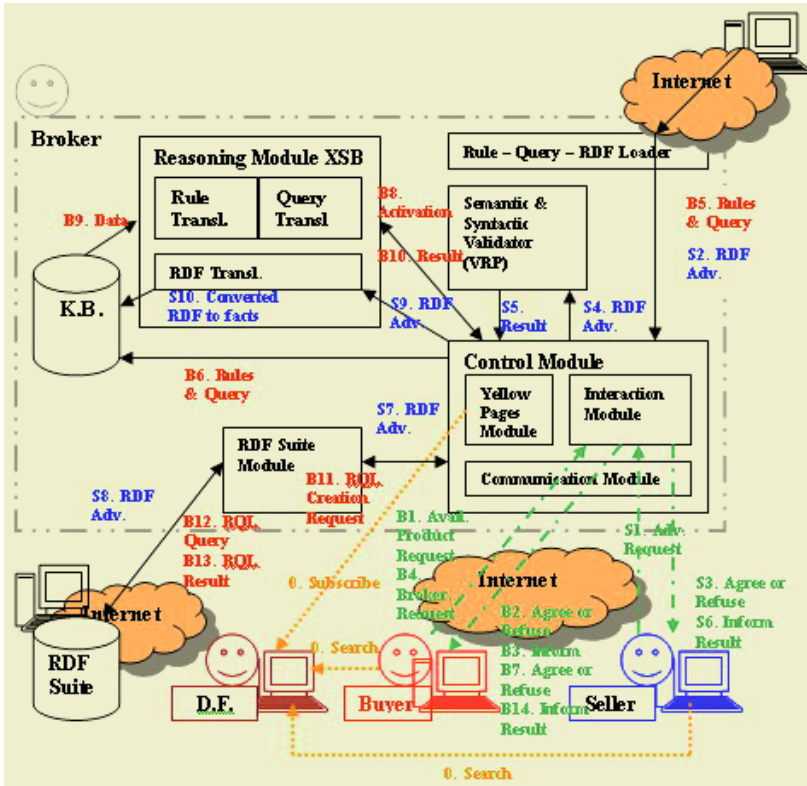
Fig. 1. The Brokering System Architecture

3.2   System Architecture and Modules

The architecture (Fig. 1) of the broker consists of five main modules: (a) reasoning module, (b) control module, (c) semantic and syntactic validator, (d) RDF Suite module, and (e) rule-query-RDF loader module. Reasoning and control modules consist of other sub-modules as one can see in Fig.1 which depicts the overall system architecture. The other three modules are stand-alone. Finally, the control module is responsible for the coordination of all the other modules.

RDF Translator: The role of the RDF translator is to transform the RDF statements into logical facts, and the RDFS statements into logical facts and rules. This transformation allows the RDF/S information to be processed by the rules provided by the Service Requester. The RDF data are transformed into logical facts of the form: *Predicate(Subject,Object)*. To capture the semantics of RDF Schema constructs, we create the following rules.

   a*: C(X):- rdf:type(X,C).*
   b*: C(X):- rdfs:subClassOf(Sc,C),Sc(X).*

c*: P(X,Y):- rdfs:subPropertyOf(Sp,P),Sp(X,Y).*
d*: D(X):- rdfs:domain(P,D),P(X,Z).*
e*: R(Z):- rdfs:range(P,R),P(X,Z).*

All the above rules are created at compile-time, i.e. before the actual querying takes place. Therefore, although the above rules at first sight seem second-order because they contain variables in place of predicate names, they are actually first-order rules, i.e. predicate names are constant at run-time.

Semantic-Syntactic Validator: This module is an embedded version of VRP[14], a parser for validating RDF descriptions. Upon receipt of an advertisement, its RDF description is checked by this module. Among others, the tests performed are: class hierarchy loops, property hierarchy loops, domain/range of subproperties, source/target resources of properties and types of resources.

Interaction and Communication Modules: The communication module is responsible for sensing the network and notifying the control module when an external event occurs. In order to decide the course of action based on the incoming message's type, the broker agent extracts the message from the queue and examines its type, i.e. whether it is a "Broker Request", "Advertise Request" etc. Accordingly it activates the interaction module. The interaction module consists of different interaction protocols that extend the standard FIPA Request interaction protocol.

RDF Suite Module: The RDF Suite module is responsible for all the actions related with the handling of the advertisements and the domain ontology. The most important functions of this module are:

- Initial upload of RDFS ontology and RDF instances into the RDF repository.
- Update of the RDF repository.
- Preparation of RQL queries and forwarding to the RDF Suite.
- Receipt of RQL queries' results.

Rule-Query-RDF Loader: The role of this module is to download the files, which contain the rules and the query of the user, in defeasible logic format. It also downloads the service providers' advertisements (RDF descriptions). Finally, it implements methods for file handling. For the implementation of this module we used the API of Java for File Management and the API for Networking.

Reasoning Module: The role of the Reasoning Module is to apply the queries to files, which contain the facts and the rules, and to evaluate the answer. When the Service Requester makes a query, the Reasoning Module compiles the files containing the facts and the rules, and applies the query to the compiled files. The answer is sent to the Control Module. XSB Prolog is used as the compiler and reasoning module for our system. We made this choice, as XSB supports the well-founded semantics of logic programs through the use of tabled predicates, and its sk_not negation operator.

Rule Parser and Translator: The Rule Parser is responsible for checking the validity of the defeasible rules, which are submitted by the Service Requester. The rules are considered to be valid, if they follow the standard syntax of defeasible logic, as described in [2]. If they are valid, the theories are passed to the Logic Translator.

The Rule Translator is responsible for transforming the rules submitted by the Service Requester using into Prolog rules that emulate the semantics of defeasible logic. The method for this translation is described in detail in [3].

Query Translator: Its role is to translate the users' queries into valid Prolog queries, that can be applied to the files representing the rules and the facts.

4 Conclusions and Future Work

In this paper we studied the brokering and matchmaking problem in the tourism domain, that is, how a requester's requirements and preferences can be matched against a set of offerings collected by a broker. The proposed solution uses the Semantic Web standard of RDF to represent the offerings, and a deductive logical language for expressing the requirements and preferences. We motivated and explained the approach we propose, and reported on a prototypical implementation exhibiting the described functionality in a multi-agent environment.

Our approach has obvious *advantages* compared to other approaches. Particularly, (a) we do not provide a fixed algorithm for brokering but it is the user who specifies the algorithm on the basis of its preferences. (b) The architecture we provide is highly reusable. The system can be applied in any domain only with the addition of a new ontology and new rules which capture the preferences. (c) Using JADE, we exploit the advantages of peer-to-peer systems (i.e. travel agencies and broker as peers) and also make use of FIPA standards for agent communication and discovery. (d) We use a highly expressive language for preferences specification with interesting features, such as conflicting rules and priorities of rules. (e) We use RDF for the expression of advertisements. This choice supports interoperability among agents and applications and facilitates the easy publication, collection and combination in decentralized dynamic settings. (f) We allow for permanent storing of advertisements with the use of the RDF Suite repository. The main *drawbacks* of the current implementation are: (a) The advertisements cannot be removed automatically when they expire. (b) The syntax of the defeasible logic may appear too complex for many users, and should be supported by, possibly graphical, authoring tools.

In the future we intend to extend the described work in various directions: (i) Add advertisement removal functionality; (ii) Implement graphical user interfaces for the integrated system. (iii) Integrate the current brokering system with the negotiation system proposed in [12]. In our current implementation, a service requester agent is able to find potential products or services and potential service providers. We intend to extend our system to support direct involvement of the service requester in negotiation with the service provider for the resulted product, as soon as the brokering stage has been completed. Finally, as a long-term goal we intend to extend the semantic brokering approach presented in this paper to brokering for general purpose semantic web services, providing matchmaking between Web Service advertisements and requests described in OWL-S.

References

1. S. Alexaki, V. Christophides, G. Karvounarakis, D. Plexousakis and K. Tolle (2001). The ICS-FORTH RDFSuite: Managing Voluminous RDF Description Bases. In *Proc. 2nd International Workshop on the Semantic Web*, Hongkong, May 1, 2001.
2. G. Antoniou, D. Billington, G. Governatori and M.J. Maher (2001). Representation results for defeasible logic. *ACM Transactions on Computational Logic* 2, 2 (2001): 255 – 287.

3. G. Antoniou, A. Bikakis (2004). "A System for Non-Monotonic Rules on the Web». Workshop on Rules and Rule Markup Languages for the Semantic Web (RuleML 2004), G. Antoniou, H. Boley (Ed.), Springer-Verlag, Hiroshima, Japan, 8 Nov. 2004.
4. T. Berners-Lee (1999). *Weaving the Web*. Harper 1999.
5. Y-C. Chen, W-T. Hsu, P-H. Hung (2003). Towards Web Automation by Integrating Semantic Web and Web Services. In *Proc. 12th International WWW Conference.*
6. B. N. Grosof, I. Horrocks, R. Volz and S. Decker (2003). Description Logic Programs: Combining Logic Programs with Description Logic". In: *Proc. 12th Intl. Conf. on the World Wide Web (WWW-2003)*, ACM Press.
7. L. Li and I. Horrocks (2003). A Software Framework For Matchmaking Based on Semantic Web Technology. In *Proc. 12th International Conference on WWW*, ACM 2003.
8. P. Maes, R.H. Guttman and A.G. Moukas (1999). Agents That Buy and Sell. *Communications of the ACM* Vol. 42, No. 3 March 1999.
9. M.J. Maher, A. Rock, G. Antoniou, D. Billington and T. Miller (2001). Efficient Defeasible Reasoning Systems. *International Journal of Tools with Artificial Intelligence* 10,4 (2001): 483—501.
10. M. Nodine, W. Bohrer, A. Hee Hiong Ngu (1998). Semantic Brokering over Dynamic Heterogeneous Data Sources in InfoSleuth. In *Proc. 15th International Conference on Data Engineering*, IEEE Computer Society.
11. M. Paolucci, T. Kawamura, T.R. Payne, K. Sycara (2002). Semantic Matching of Web Services Capabilities. In *Proc. 1st International Semantic Web Conference (ISWC-2002).*
12. T. Skylogiannis, G. Antoniou, N. Bassiliades, G. Governatori (2005) "DR-NEGOTIATE-A System for Automated Agent Negotiation with Defeasible Logic-Based Strategies ". In proceedings of the IEEE international conference on e-Technology,e-Commerce and e-Service (EEE05). Hong Kong, China April 2005.
13. D. Trastour, C. Bartolini, J. Gonzalez- Castillo (2001). A Semantic Web Approach to Service Description for Matchmaking of Services. HP Technical Report. 2001.
14. VRP. The ICS-FORTH Validating Rdf Parser - VRP (2004). Available at: http://139.91.183.30:9090/RDF/VRP/.

Uncertainty Management in Logic Programming: Simple and Effective Top-Down Query Answering

Umberto Straccia

ISTI - CNR, Via G. Moruzzi, 1 I-56124 Pisa (PI), Italy

Abstract. We present a simple, yet general top-down query answering procedure for logic programs managing uncertainty. The main features are: (i) the certainty values are taken from a certainty lattice; (ii) computable functions may appear in the rule bodies to manipulate certainties; and (iii) we solve the problem by a reduction to an equational systems, for which we device a top-down procedure.

1 Introduction

The management of uncertainty in deduction systems is an important issue whenever the real world information to be represented is of imperfect nature (which is likely the rule rather than an exception). Classical logic programming, with its advantage of modularity and its powerful top-down and bottom-up query processing techniques, has attracted the attention of researchers and numerous frameworks have been proposed towards the management of uncertainty. Essentially, they differ in the underlying notion of uncertainty[1] (e.g. probability theory [10, 11], fuzzy set theory [13], multi-valued logic [3, 6, 7, 9], possibilistic logic [5]) and how uncertainty values, associated to rules and facts, are managed. Roughly, these frameworks can be classified into *annotation based* (AB) and *implication based* (IB). In the AB approach (e.g. [6, 11]), a rule is of the form $A: f(\beta_1, \ldots, \beta_n) \leftarrow B_1:\beta_1, \ldots, B_n:\beta_n$, which asserts "the certainty of atom A is at least (or is in) $f(\beta_1, \ldots, \beta_n)$, whenever the certainty of atom B_i is at least (or is in) β_i, $1 \leq i \leq n$". Here f is an n-ary computable function and β_i is either a constant or a variable ranging over an appropriate certainty domain. In the IB approach (see [3, 7] for a more detailed comparison between the two approaches), a rule is of the form $A \xleftarrow{\alpha} B_1, ..., B_n$, which says that the certainty associated with the implication $B_1 \wedge ... \wedge B_n \rightarrow A$ is α. Computationally, given an assignment v of certainties to the B_i, the certainty of A is computed by taking the "conjunction" of the certainties $v(B_i)$ and then somehow "propagating" it to the rule head. The truth-values are taken from a certainty lattice. More recently, [3, 7, 13] show that most of the frameworks can be embedded into the IB framework (some exceptions deal with probability theory). Usually, in order to answer to a query in such frameworks, we have to compute the whole intended model by a bottom-up fixed-point computation and then answer with the evaluation of the query in this model. This always requires to compute a whole model, even if not all the atoms truth is required to determine the answer. To the best of our knowledge the only work presenting top-down procedures are [4, 6, 7, 13].

---

[1] See e.g. [12] for an extensive list of references.

In this paper we present a general, simple and effective top-down query answering procedure for logic programs over lattices in the IB framework, which generalizes the above cited works. The main features are: (i) the certainty values are taken from a certainty lattice; (ii) computable functions may appear in the rule bodies to manage these certainties values; and (iii) we solve the problem by a reduction to an equational systems over lattices, for which we device a top-down procedure, which to the best of our knowledge is novel.

We proceed as follows. In the next section, we will briefly recall some preliminary definitions. Section 3 is the main part of this work, where we present our top-down query procedure and the computational complexity analysis, while Section 4 concludes.

2   Preliminaries

Certainty Lattice. A *certainty lattice* is a complete lattice $\mathcal{L} = \langle L, \preceq \rangle$, with L a countable set of certainty values, bottom $\bot$, top element $\top$, meet $\wedge$ and join $\vee$. The main idea is that an statement $P(a)$, rather than being interpreted as either true or false, will be mapped into a certainty value c in L. The intended meaning is that c indicates to which extend (how *certain* it is that) $P(a)$ is true. Typical certainty lattices are the following. (i) Classical 0-1: $\mathcal{L}_{\{0,1\}}$ corresponds to the classical truth-space, where 0 stands for 'false', while 1 stands for 'true'. (ii) Fuzzy: $\mathcal{L}_{[0,1]_{\mathbb{Q}}}$, which relies on the unit real interval, is quite frequently used as certainty lattice. (iii) Four-valued: another frequent certainty lattice is Belnap's $\mathcal{FOUR}$ [1], where L is $\{f, t, u, i\}$ with $f \preceq u \preceq t$ and $f \preceq i \preceq t$. Here, u stands for 'unknown', whereas i stands for inconsistency. We denote the lattice as $\mathcal{L}_B$. (iv) Many-valued: $L = \langle \{0, \frac{1}{n-1}, \dots \frac{n-2}{n-1}, 1\}, \leq \rangle$, n positive integer. A special case is $\mathcal{L}_4$, where L is $\{f, lf, lt, t\}$ with $f \preceq lf \preceq lt \preceq t$. Here, lf stands for 'likely false', whereas lt stands for 'likely true'. (v) Belief-Doubt: a further popular lattice allows us to reason about *belief and doubt*. Indeed, the idea is to take any lattice $\mathcal{L}$, and to consider the cartesian product $\mathcal{L} \times \mathcal{L}$. For any pair $(b, d) \in \mathcal{L} \times \mathcal{L}$, b indicates the degree of *belief* a reasoning agent has about a sentence s, while d indicates the degree of *doubt* the agent has about s. The order on $\mathcal{L} \times \mathcal{L}$ is determined by $(b, d) \preceq (b', d')$ iff $b \preceq b'$ and $d' \preceq d$, i.e. belief goes up, while doubt goes down. The minimal element is (f, t) (no belief, maximal doubt), while the maximal element is (t, f) (maximal belief, no doubt). We indicate this lattice with $\bar{\mathcal{L}}$.

In a complete lattice $\mathcal{L} = \langle L, \preceq \rangle$, a function $f \colon L \to L$ is *monotone*, if $\forall x, y \in L$, $x \preceq y$ implies $f(x) \preceq f(y)$. A *fixed-point* of f is an element $x \in L$ such that $f(x) = x$. The basic tool for studying fixed-points of functions on lattices is the well-known Knaster-Tarski theorem. Let f be a monotone function on a complete lattice $\langle L, \preceq \rangle$. Then f has a fixed-point, the set of fixed-points of f is a complete lattice and, thus, f has a *least* fixed-point. The *least* fixed-point can be obtained by iterating f over $\bot$, i.e. is the limit of the non-decreasing sequence $y_0, \dots, y_i, y_{i+1}, \dots, y_\lambda, \dots$, where for a successor ordinal $i \geq 0$, $y_0 = \bot$, $y_{i+1} = f(y_i)$, while for a limit ordinal λ, $y_\lambda = lub_\preceq \{y_i \colon i < \lambda\}$. We denote the least fixed-point by $\mathrm{lfp}(f)$. For ease, we will specify the initial condition y_0 and the next iteration step y_{i+1} only, while the condition on the limit is implicit.

Logic Programs. Fix a lattice $\mathcal{L} = \langle L, \preceq \rangle$. We extend standard logic programs [8] to the case where *arbitrary computable functions* $f \in \mathcal{F}$ are allowed in rule bodies to

manipulate the certainty values. In this paper we assume that $\mathcal{F}$ is a family of continuous n-ary functions $f\colon \mathcal{L}^n \to \mathcal{L}$. That is, for any monotone chain $x_0, x_1, \ldots$ of values in L, $f(\vee_i x_i) = \vee_i f(x_i)$. The n-ary case $n > 1$ is similar. We assume that the standard functions $\wedge$ and $\vee$ belong to $\mathcal{F}$. Notably, $\wedge$ and $\vee$ are both continuous. For reasons of space, we limit our attention to propositional logic programs. The first order case can be handled by grounding. There exists free software (e.g. Lparse), which transforms a logic program with variables into one with propositional variables only. So, consider an alphabet of propositional letters. An *atom*, denoted A is a propositional letter. A *formula*, φ, is an expression built up from the atoms, the certainty values $c \in L$ of the lattice and the functions $f \in \mathcal{F}$. Note that members of the lattice may appear in a formula, as well as functions: e.g. in $\mathcal{L}_{[0,1]_{\mathbb{Q}}}$, $\varphi = \min(p, q) \cdot \max(r, 0.7) + v$ is a formula, where p, q, r and v are atoms. The intuition here is that the truth value of the formula $\min(p, q) \cdot \max(r, 0.7) + v$ is obtained by determining the truth value of p, q, r and v and then to apply the arithmetic functions to determine the value of φ. A *rule* is of the form $A \leftarrow \varphi$, where A is an atom and φ is a formula. The atom A is called the *head*, and the formula φ is called the *body*. A *logic program*, denoted with $\mathcal{P}$, is a finite set of rules. The *Herbrand base* of $\mathcal{P}$ (denoted $B_{\mathcal{P}}$) is the set of atoms occurring in $\mathcal{P}$. Given $\mathcal{P}$, the set $\mathcal{P}^*$ is constructed as follows; (i) if an atom A is not head of any rule in $\mathcal{P}^*$, then add the rule $A \leftarrow \bot$ to $\mathcal{P}^*$ [2]; and (ii) replace several rules in $\mathcal{P}^*$ having same head, $A \leftarrow \varphi_1, A \leftarrow \varphi_2, \ldots$ with $A \leftarrow \varphi_1 \vee \varphi_2 \vee \ldots$. Note that in $\mathcal{P}^*$, each atom appears in the head of *exactly one* rule.

Example 1 ([9]) *Consider $\mathcal{L}_{[0,1]_{\mathbb{Q}}}$, where $\wedge = \min$ and $\vee = \max$. Consider an insurance company, which has information about its customers used to determine the risk coefficient of each customer. Suppose a value of the risk coefficient is already known, but has to be re-evaluated (the client is a new client and his risk coefficient is given by his precedent insurance company). The company may have: (i) data grouped into a set of facts $\{(\texttt{Experience(john)} \leftarrow 0.7, (\texttt{Risk(john)} \leftarrow 0.5, (\texttt{Sport\_car(john)} \leftarrow 0.8\}$; and (ii) a set of rules, which after grounding are:*

$$\begin{array}{ll}
\texttt{Good\_driver(john)} \leftarrow & \texttt{Experience(john)} \wedge (0.5 \cdot \texttt{Risk(john)}) \\
\texttt{Risk(john)} & \leftarrow 0.8 \cdot \texttt{Young(john)} \\
\texttt{Risk(john)} & \leftarrow 0.8 \cdot \texttt{Sport\_car(john)} \\
\texttt{Risk(john)} & \leftarrow \texttt{Experience(john)} \wedge (0.5 \cdot \texttt{Good\_driver(john)})
\end{array}$$

Interpretations. An *interpretation I of a logic program* on the lattice $\mathcal{L} = \langle L, \preceq \rangle$ is a mapping from atoms to members of L. I is extended from atoms to formulae as follows: (i) for $c \in L$, $I(c) = c$; (ii) for formulae φ and φ', $I(\varphi \wedge \varphi') = I(\varphi) \wedge I(\varphi')$, and similarly for $\vee$; and (iii) for formulae $f(\varphi)$, $I(f(\varphi)) = f(I(\varphi))$, and similarly for n-ary functions. The ordering $\preceq$ is extended from $\mathcal{L}$ to the set $\mathcal{I}(\mathcal{L})$ of all interpretations point-wise: (i) $I_1 \preceq I_2$ iff $I_1(A) \preceq I_2(A)$, for every ground atom A. We define $(I_1 \wedge I_2)(A) = I_1(A) \wedge I_2(A)$ and similarly for $\vee$. With $\mathbf{I}_\bot$ we denote the bottom interpretation under $\preceq$ (it maps any atom into $\bot$). It is easy to see that $\langle \mathcal{I}(\mathcal{L}), \preceq \rangle$ is a complete lattice as well.

Models. An interpretation I is a *model* of a logic program $\mathcal{P}$, denoted by $I \models \mathcal{P}$, iff for all $A \leftarrow \varphi \in \mathcal{P}^*$, $I(\varphi) \preceq I(A)$ holds.

---

[2] It is a standard practice in logic programming to consider such atoms as *false*.

Query. A *query*, denoted q, is an expression of the form $?A$ (*query atom*), intended as a question about the truth of the atom A in the minimal model of $\mathcal{P}$ (see below). We also allow a query to be a *set* $\{?A_1, \ldots, ?A_n\}$ of query atoms. In that latter case we ask about the truth of all the atoms A_i in the minimal model of $\mathcal{P}$.

Semantics of Logic Programs. The semantics of a logic program $\mathcal{P}$ is determined by the least model of $\mathcal{P}$, $M_{\mathcal{P}} = \min\{I : I \models \mathcal{P}\}$. The *existence and uniqueness* of $M_{\mathcal{P}}$ is guaranteed by the fixed-point characterization, by means of the *immediate consequence operator* $\Phi_{\mathcal{P}}$. For an interpretation I, for any ground atom A, $\Phi_{\mathcal{P}}(I)(A) = I(\varphi)$, where $A \leftarrow \varphi \in \mathcal{P}^*$. We can show that the function $\Phi_{\mathcal{P}}$ is continuous over $\mathcal{I}(\mathcal{L})$, the set of fixed-points of $\Phi_{\mathcal{P}}$ is a complete lattice under $\preceq$ and, thus, $\Phi_{\mathcal{P}}$ has a least fixed-point and I is a model of a program $\mathcal{P}$ iff I is a fixed-point of $\Phi_{\mathcal{P}}$. Therefore, the minimal model of $\mathcal{P}$ coincides with the least fixed-point of $\Phi_{\mathcal{P}}$, which can be computed in the usual way by iterating $\Phi_{\mathcal{P}}$ over $\mathtt{I}_{\perp}$ and is attained after at most ω (the least limit ordinal) iterations.

Example 2 *Consider* $\mathcal{L}_{[0,1]_{\mathbb{Q}}}$, *the function* $f(x) = \frac{x+a}{2}$ ($0 < a \leq 1, a \in \mathbb{Q}$), *and* $\mathcal{P} = \{A \leftarrow f(A)\}$. *Then the minimal model is attained after* ω *steps of* $\Phi_{\mathcal{P}}$ *iterations starting from* $\mathtt{I}_{\perp}(A) = 0$ *and is* $M_{\mathcal{P}}(A) = a$.

Example 3 *Consider Example 1. It turns out that by a bottom-up computation the minimal mode is* $M_{\mathcal{P}}$, *where (for ease, we use first letters only)* $M_{\mathcal{P}}(\mathtt{R(j)}) = 0.64$, $M_{\mathcal{P}}(\mathtt{S(j)}) = 0.8$, $M_{\mathcal{P}}(\mathtt{Y(j)}) = 0$, $M_{\mathcal{P}}(\mathtt{G(j)}) = 0.32$, $M_{\mathcal{P}}(\mathtt{E(j)}) = 0.7$.

3 Top-Down Query Answering

Given a logic program $\mathcal{P}$, one way to answer to a query $?A$ is to compute the minimal model $M_{\mathcal{P}}$ of $\mathcal{P}$ by a bottom-up fixed-point computation and then answer with $M_{\mathcal{P}}(A)$. This always requires to compute a whole model, even if in order to determine $M_{\mathcal{P}}(A)$, not all the atom's truth is required. Our goal is to present a general, simple, yet effective top-down method, which relies on the computation of just a part of the minimal model. Essentially, we will try to determine the value of a single atom by investigating only a part of the program $\mathcal{P}$. Our method is based on a transformation of a program into a system of equations of monotonic functions over lattices for which we compute the least fixed-point in a top-down style. The idea is the following. Let $\mathcal{L} = \langle L, \preceq \rangle$ be a lattice and let $\mathcal{P}$ be a logic program. Consider the Herbrand base $B_{\mathcal{P}} = \{A_1, \ldots, A_n\}$ of $\mathcal{P}$ and consider $\mathcal{P}^*$. Let us associate to each atom $A_i \in B_{\mathcal{P}}$ a variable x_i, which will take a value in the domain L (sometimes, we will refer to that variable with x_A as well). An interpretation I may be seen as an assignment of lattice values to the variables $x_1, \ldots, x_n$. For the immediate consequence operator $\Phi_{\mathcal{P}}$, a fixed-point is such that $I = \Phi_{\mathcal{P}}(I)$, i.e. for all atoms $A_i \in B_{\mathcal{P}}$, $I(A_i) = \Phi_{\mathcal{P}}(I)(A_i)$. Therefore, we may identify the fixed-points of $\Phi_{\mathcal{P}}$ as the solutions over $\mathcal{L}$ of the system of equations of the following form:

$$x_1 = f_1(x_{1_1}, \ldots, x_{1_{a_1}}),$$
$$\vdots \tag{1}$$
$$x_n = f_n(x_{n_1}, \ldots, x_{n_{a_n}}),$$

where for $1 \leq i \leq n$, $1 \leq k \leq a_i$, we have $1 \leq i_k \leq n$. Each variable x_{i_k} will take a value in the domain L, each (continuous) function f_i determines the value of x_i

(i.e. A_i) given an assignment $I(A_{i_k})$ to each of the a_i variables x_{i_k}. The function f_i implements $\Phi_{\mathcal{P}}(I)(A_i)$. For instance, by considering the logic program in Example 1, the fixed-points of the $\Phi_{\mathcal{P}}$ operator are the solutions over a lattice of the system of equations

$$
\begin{aligned}
&x_{\text{E(j)}} = 0.7 \,, \ x_{\text{S(j)}} = 0.8 \,, \ x_{\text{Y(j)}} = 0 \,, \ x_{\text{G(j)}} = \min\{x_{\text{E(j)}}, 0.5 \cdot x_{\text{R(j)}}\}, \\
&x_{\text{R(j)}} = \max\{0.5, 0.8 \cdot x_{\text{Y(j)}}, 0.8 \cdot x_{\text{S(j)}}, \min\{x_{\text{E(j)}}, 0.5 \cdot x_{\text{G(j)}}\}\} \,.
\end{aligned}
\tag{2}
$$

It is easily verified that the least solution corresponds to the minimal model of $\mathcal{P}$. Therefore, our general approach for query answering is as follows: given a logic program $\mathcal{P}$, translate it into an equational system as (1) and then compute the answer in a top-down manner. Formally, let $\mathcal{P}$ be a logic program and consider $\mathcal{P}^*$. As already pointed out, each atom appears exactly once in the head of a rule in $\mathcal{P}^*$. The system of equations that we build from $\mathcal{P}^*$ is straightforward. Assign to each atom A a variable x_A and substitute in $\mathcal{P}^*$ each occurrence of A with x_A. Finally, substitute each occurrence of $\leftarrow$ with $=$ and let $\mathcal{S}(\mathcal{P})$ be the resulting equational system (see Equation 2). The answer of a query variable $?A$ w.r.t. a logic program $\mathcal{P}$ is computed by the algorithm $Solve(\mathcal{P}, ?A)$. It first computes $\mathcal{S}(\mathcal{P})$ and then calls $Solve(\mathcal{S}(\mathcal{P}), \{x_A\})$, which will solve the equational system answering with the value for x_A. Therefore, query answering in logic programs reduces to query answering in equational monotone systems of the form (1), which we address next. We refer to the monotone system as in Equation (1) as the tuple $\mathcal{S} = \langle \mathcal{L}, V, \boldsymbol{f} \rangle$, where $\mathcal{L}$ is a lattice, $V = \{x_1, ..., x_n\}$ are the variables and $\boldsymbol{f} = \langle f_1, ..., f_n \rangle$ is the tuple of functions. As it is well known, a monotonic equation system as (1) has a least solution, lfp($\boldsymbol{f}$), which can be computed by a bottom-up evaluation. Indeed, the least fixed-point of $\boldsymbol{f}$ is given as the least upper bound of the monotone sequence, $\boldsymbol{y}_0, \dots, \boldsymbol{y}_i, \dots$, where $\boldsymbol{y}_0 = \bot$ and $\boldsymbol{y}_{i+1} = \boldsymbol{f}(\boldsymbol{y}_i)$.

Our top-down procedure needs some auxiliary functions. $\mathbf{s}(x)$ denotes the set of *sons* of x, i.e. $\mathbf{s}(x_i) = \{x_{i_1}, \dots, x_{i_{a_i}}\}$ (the set of variables appearing in the right hand side of the definition of x_i). $\mathbf{p}(x)$ denotes the set of *parents* of x, i.e. the set $\mathbf{p}(x) = \{x_i : x \in \mathbf{s}(x_i)\}$ (the set of variables depending on the value of x). In the general case, we assume that each function $f_i : L^{a_i} \mapsto L$ in Equation (1) is monotone. We also use f_x in place of f_i, for $x = x_i$. Informally our algorithm works as follows. Assume we are interested in the value of x_0 in the least fixed-point of the system. We associate to each variable x_i a marking $\mathbf{v}(x_i)$ denoting the current value of x_i (the mapping $\mathbf{v}$ contains the current value associated to the variables). Initially, $\mathbf{v}(x_i)$ is $\bot$. We start with putting x_0 in the *active* list of variables $\mathbf{A}$, for which we evaluate whether the current value of the variable is identical to whatever its right-hand side evaluates to. When evaluating a right-hand side it might of course turn out that we do indeed need a better value of some sons, which will assumed to have the value $\bot$ and put them on the list of active nodes to be examined. In doing so we keep track of the dependencies between variables, and whenever it turns out that a variable changes its value (actually, it can only increase) all variables that might depend on this variable are put in the active set to be examined. At some point (even if cyclic definitions are present) the active list will become empty and we have actually found part of the fixed-point, sufficient to determine the value of the query x_0. The algorithm is given below.

Procedure $Solve(\mathcal{S}, Q)$
Input: monotonic system $\mathcal{S} = \langle \mathcal{L}, V, \boldsymbol{f} \rangle$, where $Q \subseteq V$ is the set of query variables;
Output: A set $B \subseteq V$, with $Q \subseteq B$ such that the mapping v equals $\mathrm{lfp}(f)$ on B.
1. $\mathtt{A} := Q$, $\mathtt{dg} := Q$, $\mathtt{in} := \emptyset$, **for all** $x \in V$ **do** $\mathtt{v}(x) = \bot$, $\mathtt{exp}(x) = 0$
2. **while** $\mathtt{A} \neq \emptyset$ **do**
3.   **select** $x_i \in \mathtt{A}$, $\mathtt{A} := \mathtt{A} \setminus \{x_i\}$, $\mathtt{dg} := \mathtt{dg} \cup \mathtt{s}(x_i)$
4. $r := f_i(\mathtt{v}(x_{i_1}), ..., \mathtt{v}(x_{i_{a_i}}))$
5.   **if** $r \succ \mathtt{v}(x_i)$ **then** $\mathtt{v}(x_i) := r$, $\mathtt{A} := \mathtt{A} \cup (\mathtt{p}(x_i) \cap \mathtt{dg})$ **fi**
6.   **if not** $\mathtt{exp}(x_i)$ **then** $\mathtt{exp}(x_i) = 1$, $\mathtt{A} := \mathtt{A} \cup (\mathtt{s}(x_i) \setminus \mathtt{in})$, $\mathtt{in} := \mathtt{in} \cup \mathtt{s}(x_i)$ **fi**
   od

The variable dg collects the variables that may influence the value of the query variables, the array variable exp traces the equations that has been "expanded" (body variables are put into the active list), while in keeps track of the variables that have been put into the active list so far due to an expansion (to avoid, to put the same variable multiple times in the active list due to function body expansion). The attentive reader will notice that the *Solve* procedure is related to the so-called *tabulation* procedures, like [2, 4]. Indeed, it is a generalization of it to arbitrary monotone equational systems over lattices.

Example 4 *Consider Example 1 and query variable $x_{R(j)}$ (we ask for the risk coefficient of John). Below is a sequence of $Solve(\mathcal{S}, \{x_{R(j)}\})$ computation. Each line is a sequence of steps in the 'while loop'. What is left unchanged is not reported.*
1. $\mathtt{A} := \{x_{R(j)}\}$, $x_i := x_{R(j)}$, $\mathtt{A} := \emptyset$, $\mathtt{dg} := \{x_{R(j)}, x_{Y(j)}, x_{S(j)}, x_{E(j)}, x_{G(j)}\}$,
   $r := 0.5$, $\mathtt{v}(x_{R(j)}) := 0.5$, $\mathtt{A} := \{x_{G(j)}\}$, $\mathtt{exp}(x_{R(j)}) := 1$, $\mathtt{A} := \{x_{Y(j)}, x_{S(j)}, x_{E(j)}, x_{G(j)}\}$,
   $\mathtt{in} := \{x_{Y(j)}, x_{S(j)}, x_{E(j)}, x_{G(j)}\}$
2. $x_i := x_{Y(j)}$, $\mathtt{A} := \{x_{S(j)}, x_{E(j)}, x_{G(j)}\}$, $r := 0$, $\mathtt{exp}(x_{Y(j)}) := 1$
3. $x_i := x_{S(j)}$, $\mathtt{A} := \{x_{E(j)}, x_{G(j)}\}$, $r := 0.8$, $\mathtt{v}(x_{S(j)}) := 0.8$, $\mathtt{A} := \{x_{E(j)}, x_{G(j)}, x_{R(j)}\}$,
   $\mathtt{exp}(x_{S(j)}) := 1$
4. $x_i := x_{E(j)}$, $\mathtt{A} := \{x_{G(j)}, x_{R(j)}\}$, $r := 0.7$, $\mathtt{v}(x_{E(j)}) := 0.7$, $\mathtt{exp}(x_{E(j)}) := 1$
5. $x_i := x_{G(j)}$, $\mathtt{A} := \{x_{R(j)}\}$, $r := 0.25$, $\mathtt{v}(x_{G(j)}) := 0.25$, $\mathtt{exp}(x_{G(j)}) := 1$,
   $\mathtt{in} := \{x_{Y(j)}, x_{S(j)}, x_{E(j)}, x_{G(j)}, x_{R(j)}\}$
6. $x_i := x_{R(j)}$, $\mathtt{A} := \emptyset$, $r := 0.64$, $\mathtt{v}(x_{R(j)}) := 0.64$, $\mathtt{A} := \{x_{G(j)}\}$
7. $x_i := x_{G(j)}$, $\mathtt{A} := \emptyset$, $r := 0.32$, $\mathtt{v}(x_{G(j)}) := 0.32$, $\mathtt{A} := \{x_{R(j)}\}$
8. $x_i := x_{G(j)}$, $\mathtt{A} := \emptyset$, $r := 0.64$
10. $\mathtt{stop}$. $\mathtt{return}$ $\mathtt{v}($ *in particular,* $\mathtt{v}(x_{R(j)}) = 0.64)$
The fact that only a part of the model is computed becomes evident, as the computation does not change if we add any program $\mathcal{P}'$ to $\mathcal{P}$ not containing atoms of $\mathcal{P}$, while a bottom-up computation will consider $\mathcal{P}'$ as well.

Given $\mathcal{S} = \langle \mathcal{L}, V, \boldsymbol{f} \rangle$, where $\mathcal{L} = \langle L, \preceq \rangle$, let $h(\mathcal{L})$ be the *height* of the truth-value set L, i.e. the length of the longest strictly increasing chain in L minus 1, where the length of a chain $v_1, ..., v_\alpha, ...$ is the cardinal $|\{v_1, ..., v_\alpha, ...\}|$. The *cardinal* of a countable set X is the least ordinal α such that α and X are *equipollent*, i.e. there is a bijection from α to X. For instance, $h(\mathcal{FOUR}) = 2$, while $h(\mathcal{L}_{[0,1]_\mathbb{Q}}) = \omega$. It can be shown that the above algorithms answer correctly.

Proposition 5 *Given monotone $\mathcal{S} = \langle \mathcal{L}, V, \boldsymbol{f} \rangle$, then there is a limit ordinal λ such that after $|\lambda|$ steps $Solve(\mathcal{S}, Q)$ determines a set $B \subseteq V$, with $Q \subseteq B$ such that the*

mapping v *equals* $lfp(\boldsymbol{f})$ *on* B, *i.e.* $\mathbf{v}_{|B} = lfp(\boldsymbol{f})_{|B}$. *As a consequence, let* $\mathcal{P}$ *and* $?A$ *be a logic program and a query, respectively. Then* $M_{\mathcal{P}}(A) = Solve(\mathcal{P}, \{?A\})$ [3].

From a computational point of view, by means of appropriate data structures, the operations on A, v, dg, in, exp, p and s can be performed in constant time. Therefore, Step 1. is $O(|V|)$, all other steps, except Step 2. and Step 4. are $O(1)$. Let $c(f_x)$ be the maximal cost of evaluating function f_x on its arguments, so Step 4. is $O(c(f_x))$. It remains to determine the number of loops of Step 2. In case the height $h(\mathcal{L})$ of the lattice $\mathcal{L}$ is finite, observe that any variable is increasing in the $\preceq$ order as it enters in the A list (Step 5.), except it enters due to Step 6., which may happen one time only. Therefore, each variable x_i will appear in A at most $a_i \cdot h(\mathcal{L}) + 1$ times, where a_i is the arity of f_i, as a variable is only re-entered into A if one of its son gets an increased value (which for each son only can happen $h(\mathcal{L})$ times), plus the additional entry due to Step 6. As a consequence, the worst-case complexity is $O(\sum_{x_i \in V}(c(f_i) \cdot (a_i \cdot h(\mathcal{L}) + 1)))$. Therefore:

Proposition 6 *Given monotone* $\mathcal{S} = \langle \mathcal{L}, V, \boldsymbol{f} \rangle$, *where the computing cost of each function in* $\boldsymbol{f}$ *is bounded by* c, *the arity bounded by* a, *and the height is bounded by* h, *then the worst-case complexity of the algorithm* $Solve$ *is* $O(|V|cah)$.

In case the height of a lattice is not finite, the computation may not terminate after a finite number of steps (see Example 2). Fortunately, under reasonable assumptions on the functions, we may guarantee the termination of $Solve$ (see [12]). For instance, a condition that guarantees the termination of $Solve$ is inspired directly by [3]. On lattices, we say that a function $f: L^n \to L$ is *bounded* iff $f(x_1, \ldots, x_n) \preceq \wedge_i x_i$. Now, consider a monotone system of equations $\mathcal{S} = \langle \mathcal{L}, V, \boldsymbol{f} \rangle$. We say that $\boldsymbol{f}$ is *bounded* iff each f_i is a composition of functions, each of which is either bounded, or a constant in L or one of $\vee$ and $\wedge$. For instance, the function in Example 2 is not bounded, while $f(x, y) = \max(0, x + y - 1) \wedge 0.3$ over $\mathcal{L}_{[0,1]_{\mathbb{Q}}}$ is. It can be shown that

Proposition 7 *Given monotone* $\mathcal{S} = \langle \mathcal{L}, V, \boldsymbol{f} \rangle$ *with* $\boldsymbol{f}$ *bounded, then* $Solve$ *terminates.*

Concerning the special case were the equational system is directly obtained from the translation of a logic program, we can avoid the cost of translating $\mathcal{P}$ into $\mathcal{S}(\mathcal{P})$ as we can directly operate on $\mathcal{P}$. So the cost $O(|\mathcal{P}|)$ can be avoided. In case the height of the lattice is finite, from Proposition 6 it follows immediately that the worst-case complexity for top-down query answering is $O(|B_{\mathcal{P}}|cah)$. Furthermore, often the cost of computing each of the functions of $\boldsymbol{f}_{\mathcal{P}}$ is in $O(1)$. By observing that $|B_{\mathcal{P}}|a$ is in $O(|\mathcal{P}|)$ we immediately have that in this case the complexity is $O(|\mathcal{P}|h)$. It follows that over the lattice $\mathcal{FOUR}$ ($h = 2$) the top-down algorithm works in *linear time*. Moreover, if the height is a fixed parameter, i.e. a constant, we can conclude that the additional expressive power of logic programs over lattices (with functions with constant cost) does not increase the computational complexity of classical propositional logic programs, which is *linear* The computational complexity of the case where the height of the lattice is not finite is determined by Proposition 7. In general, the continuity of the functions in $\mathcal{S}(\mathcal{P})$ guarantees the termination after at most ω steps.

---

[3] The extension to a set of query atoms is straightforward.

4 Conclusions

We have presented a simple, general, yet effective top-down algorithm to answer queries for logic programs over lattices with arbitrary continues functions in the body to manipulate uncertainty values. We believe that its interest relies on its easiness for an effective implementation and the fact that many approaches to uncertainty management in logic programming are based on lattices, respectively.

References

1. N. D. Belnap. A useful four-valued logic. In G. Epstein and J. Michael Dunn, editors, *Modern uses of multiple-valued logic*, pages 5–37. Reidel, Dordrecht, NL, 1977.
2. W. Chen and D. S. Warren. Tabled evaluation with delaying for general logic programs. *Journal of the ACM*, 43(1):20–74, 1996.
3. C. Viegas Damásio, J. Medina, and M. Ojeda Aciego. Sorted multi-adjoint logic programs: Termination results and applications. In LNCS 3229, pages 252–265. Springer Verlag, 2004.
4. C. Viegas Damásio, J. Medina, and M. Ojeda Aciego. A tabulation proof procedure for residuated logic programming. In *Proc. of European Conf. on Artificial Intelligence (ECAI-04)*, 2004.
5. D. Dubois, J. Lang, and H. Prade. Towards possibilistic logic programming. In *Proc. of the 8th Int. Conf. on Logic Programming (ICLP-91)*, pages 581–595. The MIT Press, 1991.
6. M. Kifer and V.S. Subrahmanian. Theory of generalized annotated logic programming and its applications. *Journal of Logic Programming*, 12:335–367, 1992.
7. L. V.S. Lakshmanan and N. Shiri. A parametric approach to deductive databases with uncertainty. *IEEE Transactions on Knowledge and Data Engineering*, 13(4):554–570, 2001.
8. J. W. Lloyd. *Foundations of Logic Programming*. Springer, Heidelberg, RG, 1987.
9. Y. Loyer and U. Straccia. The approximate well-founded semantics for logic programs with uncertainty. In LNCS 2747, pages 541–550, 2003. Springer-Verlag.
10. T. Lukasiewicz. Probabilistic logic programming. In *Proc. of the 13th European Conf. on Artificial Intelligence (ECAI-98)*, pages 388–392, 1998.
11. R. Ng and V.S. Subrahmanian. Probabilistic logic programming. *Information and Computation*, 101(2):150–201, 1993.
12. U. Straccia. Top-down query answering for logic programs over bilattices. Technical Report 2004-TR-62, Istituto di Scienza e Tecnologie dell'Informazione, CNR, Pisa, Italy, 2004.
13. P. Vojtáš. Fuzzy logic programming. *Fuzzy Sets and Systems*, 124:361–370, 2004.

Specifying Distributed Authorization with Delegation Using Logic Programming

Shujing Wang and Yan Zhang

University of Western Sydney, Australia
{shwang,yan}@cit.uws.edu.au

Abstract. Trust management is a promising approach for the authorization in distributed environment. There are two key issues for a trust management system: how to design high-level policy language and how to solve the compliance-checking problem [3, 4]. We adopt this approach to deal with distributed authorization with delegation. In this paper, we propose an *authorization language* $\mathcal{AL}$, a human-understandable high level language to specify various authorization policies. Language $\mathcal{AL}$ has rich expressive power which can not only specify delegation, and threshold structures addressed in previous approaches, but also represent structured resources and privileges, positive and negative authorizations, separation of duty, incomplete information reasoning and partial authorization and delegation. We define the semantics of $\mathcal{AL}$ through logic programming with answer set semantics and through an authorization scenario we demonstrate the application of language $\mathcal{AL}$.

1 Introduction

Access control is an important topic in computer security research. It provides availability, integrity and confidentiality services for information systems. The access control process includes identification, authentification and authorization. With the development of Internet, there are increasing applications that require distributed authorization decisions. For example, in the application of electronic commerce, many organizations use the Internet (or large Intranets) to connect offices, branches, databases, and customers around the world. One essential problem among those distributed applications is how to make authorization decisions, which is significantly different from that in centralized systems or even in distributed systems which are closed or relatively small. In these traditional scenarios, the authorizer owns or controls the resources, and each entity in the system has a unique identity. Based on the identity and access control policies, the authorizer is able to make his/her authorization decision. In distributed authorization scenarios, however, there are more entities in the system, which can be both authorizers and requesters, and probably are unknown to each other. Quite often, there is no central authority that everyone trusts. Because the authorizer does not know the requester directly, he/she has to use the information from the third parties who know the requester better. He/She trusts these third parties only for certain things to certain degrees. The trust and delegation issues make distributed authorization different from traditional access control scenarios.

In recent years, the trust management approach, which was initially proposed by Blaze *et al.* in [3], has received a great attention by many researchers [3, 4, 7]. Under this approach public keys are viewed as entities to be authorized and the authorization can be delegated to third parties by credentials or certificates. This approach frames the authorization decision as follows: "Does the set C of *credentials* prove that the *request r complies with* the local security *policy P*?" from which we can see that there are at least two key issues for a trust management system: (1) Designing a high-level policy language to specify the security policy, credentials, and request. Better it is if the language has richer expressive power and is more human-understandable; (2) Finding well theory foundation for checking proof of compliance.

In our research, we view the problem of a language for representing authorization policy and credentials as a knowledge representation problem. Logic programming approach has been proved very successful in knowledge representation. Some research using logic programming in centralized access control systems has been well developed [2, 5], where underlying languages can support multiple access-control policies and achieve separation of policies from enforcement mechanisms. But their work focuses on centralized systems, and can not be used in distributed systems. Delegation Logic [7], developed by Li *et al.*, is an approach in distributed systems along this line. However the D1LP is based on Definite ordinary logic program, which is less expressive and flexible, and cannot deal with some important issues such as negative authorization, and nonmonotonic reasoning. D2LP extends D1LP to have the nonmonotonic features and bases its syntax and semantics on GCLP (Generalized Courteous Logic Programs). Since it was only briefly mentioned in [6], it was not clear yet how D2LP can handle nonmonotonic reasoning in distributed authorization. In our research, we design a language $\mathcal{AL}$, a nonmontonic language, which is based on Answer Set Programming. We adopt the delegation with depth control and static and dynamic threshold structure from DL approach. Compared to previous work in trust management systems, our language is able to specify positive and negative authorization, the request from conjunctive subjects, structured resources and privileges, incomplete information reasoning, and partial delegation and authorization. The reasons we choose Answer Set Programming as the foundation of language $\mathcal{AL}$ are as follows: (1) Through negation as failure, Answer Set Programming implements nonmonotonic reasoning which is reasoning about incomplete information. A language with nonmontonic feature is easy to specify security policies which is close to the natural language. For example, many systems permit a login request only if they do not find that the requester inputs the password wrong over consecutive three times; (2) The highly efficient solvers for Answer Set Programming have been implemented, such as `Smodels`, `dlv` etc. This is an important reason that Answer Set Programming has been widely applied in product configuration, planning, cryptanalysis, etc. We need to indicate that `Smodels` supports some extended literals such as constraint literal and conditional literal which are particularly useful to express the static and dynamic threshold structures.

The rest of this paper is organized as follows. Section 2 presents the syntax and expressive features of language $\mathcal{AL}$. Section 3 provides the formal definition of the semantics of $\mathcal{AL}$ and shows how to use language $\mathcal{AL}$ through an authorization scenario. Finally, Section 4 concludes the paper.

2 An Authorization Language $\mathcal{AL}$

In this section, we define the syntax of the authorization language $\mathcal{AL}$ and illustrate its expressiveness via some examples. Language $\mathcal{AL}$ consists of *entities, atoms, thresholds, statements, rules* and *queries*.

Entities. In distributed systems, the entities include *subjects* who are authorizers who own or control resources and requesters who make requests, *objects* which are resources and services provided by authorizers, and *privileges* which are actions executed on objects. We define constant entities starting with a lower-case character and variable entities starting with an upper-case character for subjects, objects and privileges. We provide a special subject, $local$. It is the local authorizer which makes the authorization decision based on local policy and credentials from trusted subjects.

Atoms. An atom is a function symbol with n arguments, generally one, two or three constant or variable entities to express a logical relationship between them. There are three types of atoms:

Relation atoms. A relation atom is a 2-ary function symbol and expresses the relationship of two entities. We provide three relation atoms, *neq, eq*, and *below*. The atoms *neq* and *eq* denote two same type entities equal or not equal, and *below* to denote the hierarchy structure for objects and privileges. In most realistic systems, the data information is organized using hierarchy structure, such as file systems and object oriented database system. For example, $below(ftp, pub\text{-}services)$ denotes that ftp is one of $pub\text{-}services$.

Assert atoms. This type of atoms, denoted by $exp(a_1, \ldots, a_n)$, is an application dependant function symbol with n arguments, usually one, two or three constant or variable entities and states the property of the subjects, or the relationship between entities. It is a kind of flexible atoms in language $\mathcal{AL}$. For example, $isaTutor(alice)$ denotes that $alice$ is a tutor.

Auth atoms. An auth atom is of the form, $right(sign, priv, obj)$, in which *sign* is +, -, or $\square$. It states $positive$ or $negative$ privilege executed on the $object$ based on its arguments, *sign, obj*, and *priv*. When an auth atom is used in delegation statement, the $sign$ is $\square$ to denote both positive and negative authorizations.

Threshold Structure. There are two types of threshold structures, static threshold and dynamic threshold. The static threshold is of the form, $sthd(k, [s_1, s_2, \ldots, s_n])$, where k is the threshold value, $[s_1, s_2, \ldots, s_n]$ is the static threshold pool, and we require $k \leq n$ and $s_i \neq s_j \ for \ 1 \leq i \neq j \leq n$. This structure states that we choose k subjects from the threshold pool. The dynamic threshold structure is of the form, $dthd \ (k, S, sub \ assert \ exp(\ldots, S, \ldots))$, where S is a subject variable and we require that S is one argument of assert atom exp. This structure denotes we choose k subjects who satisfy the assert statement.

Statements. There are four types of statements, *relation statement, assert statement, delegation statement*, and *auth statement*. Only the local authorizer can issue the *rela-*

tion statement to denote structured resources and privileges. We provide the following examples for them respectively.

Relation statement: *local* says *below*(*alice*, *postgraduate*).
Assert statement: *hrM* asserts *isStaff*(*Alice*).
Delegation statement: *sa* delegates *right*($\Box$, *access*, *ftp*) to *ipB* with depth 2.
Auth statement: *sa* grants *right*(+, *access*, *ftp*) to *ipB*.

Rules. A rule is a local authorization policy or a credential from other subjects, in which the basic unit is a statement. It is of the form,

⟨*head-statement*⟩ *if* ⟨*list-of-body-statement*⟩
     with absence ⟨*list-of-body-statement*⟩.

The issuer of the rule is the issuer of the head statement. Then we limit the issuer in head statements is just a single subject while it can be a complex structure in body statements. We present the following examples to demonstrate expressive power of $\mathcal{AL}$.

Partial delegation and authorization: A firewall system protects the *allServices*, including *ssh, ftp*, and *http*. The administrator permits *ipA* to access all the services except *ssh* and delegates this right to *ipB* and allow it redelegated within 2 steps.

local grants *right*(+, *access*, *X*) to *ipA* if
     local says *below*(*X*, *allServices*), *local* says *neq*(*X*, *ssh*).
local delegates *right*($\Box$, *access*, *X*) with depth 2 to *ipB* if
     local says *below*(*X*, *allServices*), *local* says *neq*(*X*, *ssh*).

Separation of duty: A company chooses to have multiparty control for emergency key recovery. If a key needs to be recovered, three persons are required to present their individual PINs. They are from different departments, *managerA*, a member of management, *auditorB*, an individual from auditing department, and *techC*, one individual from IT department.

local grants *right*(+, *recovery*, *k*) to [*managerA*, *auditorB*, *techC*].

Negative authorization: In a firewall system, the administrator *sa* does not permit *ipB* to access the *ftp* services.

sa grants *right*(−, *access*, *ftp*) to *ipB*.

Incomplete information reasoning: In a firewall system, the administrator *sa* permit a *person* to access the *mysql* service if the human resource manager *hrM* asserts the person is a *staff* and not in holiday.

sa grants *right*(+, *access*, *mysql*) to *X* if
     hrM asserts isStaff(*X*), with absence *hrM asserts inHoliday*(*X*).

Query

Language $\mathcal{AL}$ supports single subject query and group subject query. They are of the forms,

sub requests *right*(+, *p*, *o*), and
[$s_1, s_2, \ldots, s_n$] requests *right*(+, *p*, *o*).

Through group subject query, we implement *separation of duty* which is an important security concept. It ensures that a critical task cannot be carried out by one subject. If

we grant an authorization to a group subject, we permit it only when the subjects in the group request the authorization at the same time.

3 Semantics of $\mathcal{AL}$

We define the semantics for language $\mathcal{AL}$ through translating it to logic programs, evaluating answer sets of them using *Smodels* and answering the authorization queries based on answer sets. Readers are referred to [1] for Answer Set Programming and [8] for *Smodels*. In this section, we first give the formal definition for *domain description* $\mathcal{D}_{\mathcal{AL}}$ of language $\mathcal{AL}$ and how to answer queries $\mathcal{Q}_{\mathcal{AL}}$ which are requests in language $\mathcal{AL}$. In our full paper [9], we define function $TransRules(\mathcal{D}_{\mathcal{AL}})$ to translate $\mathcal{D}_{\mathcal{AL}}$ into answer set program $\mathcal{P}$, and function $TransQuery(\mathcal{Q}_{\mathcal{AL}})$ to translate query $\mathcal{Q}_{\mathcal{AL}}$ into answer set program Π and ground literals $\varphi(+)$ and $\varphi(-)$. We use $\varphi(+)$ to denote positive right and $\varphi(-)$ to denote negative right. We solve a query based on $\mathcal{P}$, Π and φ via *Smodels*. Note that program $\mathcal{P}$ includes not only rules translated from $\mathcal{D}_{\mathcal{AL}}$ directly, but also propagation rules, and conflict resolution rules predefined because we permit both positive and negative authorizations in our language. We also argue that through logic programming we can specify different propagation and conflict resolution policies easily [9].

An answer set program may have one, more than one, or no answer sets at all. For a given program Π and a ground atom φ, we say Π *entails* φ, denoted by $\Pi \models \varphi$, iff φ is in every answer set of Π.

Definition 1. *A* domain description $\mathcal{D}_{\mathcal{AL}}$ *of language* $\mathcal{AL}$ *is a finite set of rules.*

Definition 2. *Given a domain description* $\mathcal{D}_{\mathcal{AL}}$ *and a query* $\mathcal{Q}_{\mathcal{AL}}$ *of language* $\mathcal{AL}$, *there are* $TransRules(\mathcal{D}_{\mathcal{AL}}) = \mathcal{P}$ *and* $TransQuery(\mathcal{Q}_{\mathcal{AL}}) = \Pi \cup \varphi(+) \cup \varphi(-)$. *We say that query* $\mathcal{Q}_{\mathcal{AL}}$ *is* permitted, denied, *or* unknown *by the domain description* $\mathcal{D}_{\mathcal{AL}}$ *iff* $(\mathcal{P} \cup \Pi) \models \varphi(+)$, $(\mathcal{P} \cup \Pi) \models \varphi(-)$, *or* $(\mathcal{P} \cup \Pi) \not\models \varphi(+)$ *and* $(\mathcal{P} \cup \Pi) \not\models \varphi(-)$ *respectively.*

Now we represent a specific authorization scenario to demonstrate how to define the semantics of language $\mathcal{AL}$.

Scenario: A server provides the services including $http$, ftp, $mysql$, and $smtp$. It sets up a group for them, called $services$. The server trusts the security officer so to assign all the services to related people. The security officer so grants the services to $staff$. The service $mysql$ can not be accessed if the staff is in holiday. Officer so can get information of staff from the human resource manager hrM. The policy and credentials are described as follows.

> $local$ $says$ $below(http, services)$.
> $local$ $says$ $below(ftp, services)$.
> $local$ $says$ $below(mysql, services)$.
> $local$ $says$ $below(smtp, services)$.
> $local$ $delegates$ $right(\Box, access, services)$ to so with depth 3.
> so grants $right(+, access, services)$ to X if hrM asserts $isStaff(X)$.

so grants $right(-, access, mysql)$ to X
    if hrM asserts $isStaff(X)$, hrM asserts $inHoliday(X)$.
hrM asserts $isStaff(alice)$.
hrM asserts $isStaff(bob)$.
hrM asserts $inHoliday(alice)$.

After translating above domain description, we get a logic program. Through *Smodels*, we compute one and only one answer set for it. Here we demonstrate part authorization information for $Alice$ in the answer set as example,

$auth(local, alice, right(+, access, http), 2)$,
$auth(so, alice, right(+, access, http), 1)$,
$auth(local, alice, right(-, access, mysql), 2)$,
$auth(so, alice, right(-, access, mysql), 1)$,
$auth(local, alice, right(+, access, mysql), 2)$,
$auth(so, alice, right(+, access, mysql), 1)$,
$grant(alice, right(+, access, ftp))$,
$grant(alice, right(+, access, smtp))$,
$grant(alice, right(+, access, http))$,
$grant(alice, right(-, access, mysql))$.

For Alice, there is both positive and negative authorization information for the service $mysql$ with the same step 2. According to the following conflict resolving and decision rule,
    $grant(X, right(-, P, O)) \leftarrow$
        $auth(local, X, right(+, P, O), T_1)$,
        $auth(local, X, right(-, P, O), T_2)$,
        $not\ neg\_far(X,\ right(-, P, O), T_2)$,

if the step of negative authorization is less or equal to that of positive authorization, we grant negative authorization to Alice. In the answer set, because there are both positive and negative authorizations for service $mysql$ to Alice, we grant Alice negative authorization for service $mysql$.

The predicates $auth$ and $delegation$ are helpful for us to find the authorization path. We have

$auth(so, alice, right(-, access, mysql), 1)$,
$delegate(local, so, right(\square, access, services), 3, 1)$ and
$below(mysql, services)$,

From the rule,
    $auth(local, X, right(-, access, O), T + 1) : -$
        $delegate(local, so, right(\square, access, services), 3, 1)$,
        $auth(so, X, right(-, access, O), T)$,
        $below(O, services)$.
we conclude $auth(local, alice, right(-, access, mysql), 2)$ is in the answer set. We can also find that the authorization passes from $local$ to so, and then from so to $alice$.

4   Conclusion

In this paper, we developed an expressive authorization language $\mathcal{AL}$ to specify the distributed authorization with delegation. We used Answer Set Programming as a foundational basis for its semantics. As we have showed earlier, $\mathcal{AL}$ has a rich expressive power which can represent positive and negative authorization, structured resources and privileges, partial authorization and delegation, and separation of duty. We should indicate that our formulation also has implementation advantages due to recent development of Answer Set Programming technology in AI community[1]. The scenario in section 3 has been fully implemented through Answer Set Programming.

References

1. C. Baral. Knowledge Representation, Reasoning and Declarative Problem Solving. Cambridge University Press, 2003.
2. E. Bertino, F. Buccafurri, E. Ferrari, and P. Rullo. A Logical Framework for Reasoning on Data Access Control Policies. In *Proceedings of the 12th IEEE Computer Security Foundations Workshop(CSFW-12)*, pages 175-189, IEEE Computer Society Press, Los Alamitos, CA, 1999.
3. M. Blaze, J. Feigenbaum, and J. Lacy. Decentralized Trust Management. In *Proceedings of the Symposium on Security and Privacy*, IEEE Computer Society Press, Los Alamitos,1996, pages 164-173.
4. M. Blaze, J. Feigenbaum, and M. Strauss. Compliance-checking in the PolicyMaker trust management system. In *Proceedings of Second International Conference on Financial Cryptography (FC'98)*, volume 1465 of Lecture Notes in Computer Science, pages 254-274. Springer, 1998.
5. S. Jajodia, P. Samarati, and V. S. Subrahmanian. Flexible Support for Multiple Access Control Policies. In *ACM Transactions on Database Systems*, Vol.26, No.2, June 2001, Pages 214-260.
6. N. Li, J. Feigenbaum, and B.N. Grosof. A logic-based knowledge representation for authorization with delegation (extended abstract). In *Proceedings of the IEEE Computer Security Foundations Workshop (CSFW-12)*(June). IEEE Computer Society Press, Los Alamitos, Calif., pages 162-174.
7. N. Li, B. N. Grosof, and J. Feigenbaum. Delegation Logic: A logic-based approach to distributed authorization. In *ACM Transactions on Information and System Security (TISSEC)*, February 2003.
8. T. Syrjänen. Lparse 1.0 User's Mannual. http://www.tcs.hut.fi/Software/smodels.
9. S. Wang, and Y. Zhang. Handling Distributed Authorization with Delegation through Answer Set Programming(manuscript). 2005.

---

[1] Please refer to http://www.tcs.hut.fi/Software/smodels/index.html

The Design and Implementation of SAMIR

Fabio Zambetta[1] and Fabio Abbattista[2]

[1] RMIT University, Melbourne VIC 30001, Australia
fabio@cs.rmit.edu.au
[2] Universitá degli Studi di Bari, 70100 Bari, Italy
fabio@di.uniba.it

Abstract. This paper describes the architecture and the implementation of the SAMIR (Scenographic Agents Mimic Intelligent Reasoning) system, designed to implement 3D conversational agents. SAMIR uses an XCS classifier system to learn natural facial postures to be shown by the agents, during their dialogues with users. An administration program can be run by *agent administrators* in order to fine-tune the agent knowledge base and emotive responses. It results in a flexible approach to dialogue and facial expressions management, fitting the needs of different web applications such as e-shops or digital libraries.

1 Introduction

Intelligent learning agents are software components required, for example, by web applications, military training applications or videogames. When used in a web context, they are mainly devoted to replacing classical WYSIWYG interfaces, which are often confusing for casual users, with reactive and/or pro-active virtual clerks. This sort of advisors can converse with them and fulfill their specific needs, finding information on their behalf without pressing a lot of buttons and selecting several menus. These systems use an animated 2D/3D look-and-feel and since that, they are often referred to as Embodied Conversational Agents, representing a face or an entire body, which serves to manifest the personality and the intelligence of the virtual persona. It enhances users trust into an interaction scheme which is clearly mimicking a face-to-face conversation, as reported in [4].

The Interface project [13] features an MPEG-4 facial animation player completely written in Java, built upon the Shout3D rendering engine. Some example applications were provided online, enabling users to chat with a chatterbot improved with a synthetic face and some audio output. In these samples, the user can discuss with the dialogue software by typing text, whilst the server includes a dialogue manager to generate an appropriate answer with emotions, and a Text-To-Speech process to construct audio output and the phonemes.

Military systems have been using intelligent agents for quite a while, and the Steve agent was used in many simulations of peacekeeping scenarios, as reported in [10]. It is remarkable that Steve's behavior is not scripted: It consists of a set of general, domain independent capabilities operating over a declarative representation of domain tasks. Steve can be applied to a new domain simply

by feeding new declarative knowledge of the virtual world at hand (i.e., its objects and especially their spatial properties), and the description of the tasks to be completed. Task description uses a relatively standard hierarchical plan representation, which is clearly based on goals and their causal relationships.

Many different videogames use so-called *bots*, specific kind of intelligent agents, mainly devoted to executing relevant actions in a game, be it a RTS (Real Time Strategic), a FPS (First Person Shooter), or something else. Quake bots are a very well known example in this field, being firstly used by the best-selling Quake videogame, and since that moment on, by a considerable amount of hobbyists and researchers around the world who were interested in implementing strategies and smart game-play [8].

SAMIR is focused on web solutions based on light and efficient clients. SAMIR digital assistants are smart web agents, using an XCS classifier system to learn correct sensors/effectors pairs, integrated with a customizable 3D look [2]. The remainder of the paper is organized as follows. The next section describes the architecture of SAMIR. In the third, fourth and fifth sections, the three main modules of SAMIR are detailed. Some experimental results are given in the sixth section. Finally, conclusions are sketched in the last section.

2   Overview of the System

The SAMIR (Scenographic Agents Mimic Intelligent Reasoning) system is a framework for building Embodied Conversational Agents running on the web [1]. SAMIR has been designed in order to fulfill the following specific requirements:

- SAMIR's agents should be able to run on the web, avoiding specific plug-ins.

- SAMIR should be available to heterogeneous clients, running on different operating systems, using at least the 1.1 version of the Java language.

- Agents should be embodied by 3D faces, rendered and animated at a high framerate on the client machine (with optional OpenGL acceleration).

- Last but not least, agents should mimic a convincing and, to some extent, intelligent behavior.

Figure 1 depicts the standard implementation of the SAMIR system. The main scenario for SAMIR is an e-commerce website: The user types some requests in a text box that will be captured by the DMS (Dialogue Management System) Client, which will send it to the Server Dispatcher, on the server side of the application. The Server Dispatcher is in charge of controlling messages on the server side, and will forward users requests both to the BeMan (Behavior Manager) module and to the DMS Server. The former will process the request to decide what kind of expression should be assumed by the 3D face, while the latter will decide the next answer to be given to the user. The DMS Client will analyze data producing the answer to be shown to users in a text area, and finally SAMIR face will be animated according to the new expression generated by the system.
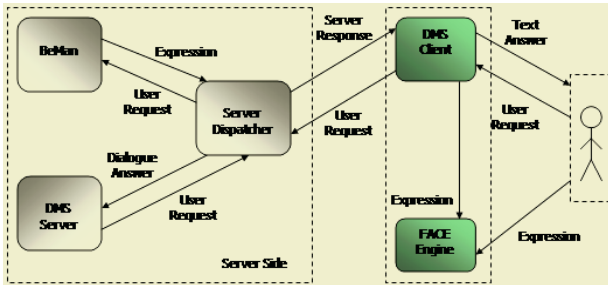
Fig. 1. The SAMIR architecture.

3 The Animation Module

The FACE (Facial Animation Compact Engine) module heavily relies on SACs (Standard Anatomic Components).

SACs are face regions, acting as objects offering different services. Those include different types of low-level deformations useful for the animation process, or face remodeling needed in order to customize 3D faces. We decided to use mainly a morph-target base technique because, in our opinion, it represented the best compromise between speed and accuracy. FACE was conceived keeping in mind lightness and performance so that it supports a variable number of morph targets: For example we have used either 12 high-level ones or the entire low-level FAP set, in order to comply with the MPEG-4 standard [9]. Using just a small set of high-level parameters might be extremely useful when debugging the behavior module because it gets rid of long sets of facial parameters. An unlimited number of timelines can be used allocating one channel for the stimulus-response expressions, another one for eye-lid non-conscious reflexes, and so on.

4 The Dialogue Management System

Two software modules, the DMS Client and the DMS Server, are the main components of the Dialogue Management System (DMS), a client/server application capable of managing user inputs and properly process them in order to obtain an adequate response to user requests and needs.

The DMS Client captures user inputs to be sent to the DMS Server, which analyzes the current input text typed by the user and looks for certain classes of words that are particularly relevant in the domain addressed by the system. The result of the task performed by the server side of the DMS consists of a textual response containing the result of the search issued by the user, or an invitation to be more precise in formulating requests. This task is accomplished through the use of the ALICE chatterbot open source software [3]. All the knowledge of the system is enclosed in the AIML dialogues, which are XML files storing a set of pre-defined categories. An AIML category is an elementary unit structured in two

main parts: A pattern section matching user input, and a corresponding template section containing an adequate response to user requests and/or actions (i.e., a JavaScript execution). The pattern matching algorithm used by ALICE finds a sort of morphological distance (i.e., based on the similarity of words) between the text typed by the user and the list of possible patterns in the system. To assist users during their navigation in a bookshop web site, for example, required the definition of a set of AIML categories resembling a dialogue in a real bookstore. Successful examples of book requests issued to the Amazon bookshop web site are: *I want a book written by Sepulveda*, *I am looking for a book entitled Journey and whose author is Celine*, *I am searching for all books written by Henry Miller and published after 1970*, *I am interested in a book about horror*. Equivalent interrogative forms are also allowed.

5   The Behavior Generator

The Behavior Manager (BeMan) aims at managing the consistency between the facial expression of the character and the conversation tone. It relies on soft computing approaches, namely a Learning Classifier System (LCS). Learning Classifier Systems (LCS) represent a machine learning paradigm introduced by Holland in 1976 [7] which tries to combine the best of reinforcement learning techniques [12] and genetic algorithms [6].

The learning module of SAMIR has been implemented through an XCS [14], a specific kind of LCS that differs in some aspects from the traditional Holland's framework. At discrete time intervals, the agent observes a state of the environment, takes an action, observes a new state and finally receives an immediate reward. In our case the environment is the web site where SAMIR is embedded in, the actions are represented by the verbal (simple text) and non-verbal (facial expressions) responses of the character and the reward should be an estimate of the correctness of SAMIR behaviour. In such an environment, the repertoire of possible actions and states perceivable by the agent is very limited: In fact, the agent could only detect the clickstreams or textual requests. Given such a hard constraint, to find a way to infer the proper behaviour for the agent, we decided to map the universe of the dialogue for some applicative scenario (some e-commerce web sites, a peculiar entertainment application, etc.).

This mapping is expressed by specific classifiers i.e., rules in a classical **IF** *input* **THEN** *output* form: In the current implementation of the system the input is a string coded on a ternary alphabet (0, 1 and #, the don't care symbol which represent a wildcard allowing for the presence of 0 or 1), whilst the output is coded on a simple binary alphabet. The input string, represented by a 6 digits string, indicates whether a peculiar category was detected or not. As in the well-known domain of text categorization, user's phrases are inspected, in order to decide what subset of the given semantic categories will match them. In our scenario, for example, most of the experimentation performed so far used *compliment*, *insult*, *order*, *greeting*, *vulgarity* and *question*. For each category a vocabulary of fitting terms is defined (a Bag-of-words [11]), which will be detect

in the users' requests. The verbal responses of the agent are grouped together in categories, while the non-verbal responses (e.g., the facial expressions assumed by the 3D face) are expressed using Ekman's model of universal expressions [5]: Six expressions, which are common to the whole human race, are linearly combined in order to achieve a large set of possible expressions. We employed 7 bits to code the value of each expression (values are in the [0,127] range), so that a 42 digits string constitutes the output part of a classifier. A genetic algorithm which recombines the existing rules can learn new ones, which are necessary in order to avoid a very standard generation of behaviours.

An example of a classifier is portrayed below (note that blanks are just used for the sake of easier reading):

IF 1000## **THEN** 0110010 0000000 0000000 0000000 1100100 0000000

6   Experimental Results

Here, we are just describing the last set of experiments performed so far, with our Agent Administration Program (AAP): It is a server side application, used to let agents learn a set of meaningful rules at the agent administrator's wish. For other experimental results related to different aspects of SAMIR, one can take a look for example at [2].

The core idea of the AAP is a reward procedure which can give credit based on the administrator intervention: The XCS learns the rules selected by the administrator: Clearly, different sets of rules might be defined by different administrators. We present here some results obtained using the administration program over two period of fifteen minutes both: Our program should be able to generate meaningful rules in a few minutes, because otherwise there would be no point in using it. The importance of our Administrative program is twofold, because it enables to test and detect errors occurred in the dialogue design phase, and it can generate and select effective rules for the SAMIR system, by means of a fast and interactive process.

Figure 2 shows the growth of rules in the system during the experiment, whilst figure 3 reports the number of rules which achieved very good rewards, ending up as the most relevant rules. It should be noted here, that the duration
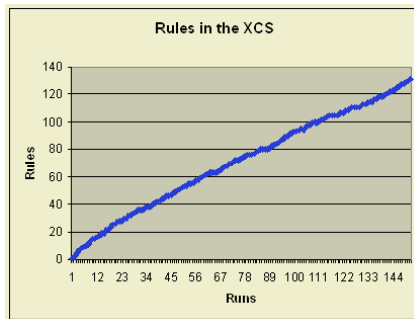
Fig. 2. The growth of rules in the system, during the experiment.
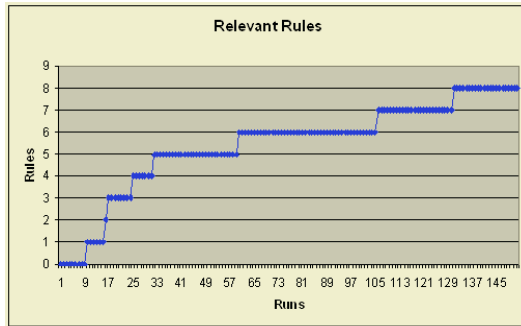
Fig. 3. The number of relevant rules i.e., the ones which were highly rewarded.

Fig. 4. The action part of a relevant rule in the system.

of the experiment summed up to 30 minutes, the number of runs (a "run" here is a request/answer couple) was 152, the final number of rules in the system 131, and the number or relevant rules 8. These numbers are quite different from the ones usually obtained with maze or classification problems, but SAMIR uses a different definition of a run: Usually a classifier system is executed in a batch procedure, in order to learn, for example, all the useful rules until we reach a goal state (i.e., the LCS has learnt a satisfactory policy in that run). In our case a run is represented by a request/answer couple, and the learning process could go on indefinitely if we did not stop it: There is not an explicit condition for convergence, but the administrator can stop and restore the process at any given time. We are reporting just a rule here for the sake of brevity, but we are giving some more details to explain the learning process in the XCS. One of the most rewarded rules during the experiment was **100000->0000111 1101100 0001010 00001110011111 1010011**. It can be translated in a more readable form: **IF** *greetings* **THEN** (*sadness=7, joy=108, anger=10, disgust=7, fear=31, surprise=83*). This is a very good rule for the system, because if user's greetings are perceived, mainly joy and surprise will be shown by the 3D face. The values of the other expressions are so small that they are not perceivable in the expression (see figure 4), while the fear value seems to reinforce the surprise one, because both the expressions cause the 3D face to open its mouth.

7   Conclusions

In this paper we presented the architecture and the implementation of the
SAMIR system, born to create 3D agents acting as advisors for web applica-
tions. Specifically, we explained how theAAP tool can learn rules by means of
the XCS, to generate facial expressions coherent with the conversation going on
with users. We reported some experimental results that clearly show that the
AAP can be profitably used to modify the knowledge base of an agent in an
interactive fashion.

References

1. F. Abbattista, A. Paradiso, G. Semeraro, and F. Zambetta. An agent that learns to
   support users of a web site. In M. Roy, R. Koeppen, S. Ovaska, T. Furuhashi, and
   Hoffmann F., editors, *Soft Computing and Industry: Recent Applications*, pages
   489–496. Springer, 2002.
2. F. Abbattista, A. Paradiso, G. Semeraro, and F. Zambetta. An agent that learns
   to support users of a web site. volume 4, pages 1–12. February 2004.
3. ALICE AI Foundation. http://www.alicebot.org.
4. J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors. *Embodied Conversa-
   tional Agents*. MIT Press, Cambridge, 2000.
5. P. Ekman. *Emotion in the human face*. Cambridge University Press, Cambridge,
   1982.
6. D. A. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learn-
   ing*. Addison-Wesley, 1989.
7. J. H. Holland. Adaptation. In R. Rosen and F. M. Snell, editors, *Progress in
   theoretical biology*. Plenum, New York, 1976.
8. J. E. Laird. It knows what you're going to do: Adding anticipation to a Quakebot.
   In *International Conference on Autonomous Agents*, pages 385–392, 2001.
9. MPEG-4 standard specification. http://mpeg.telecomitalialab.com/standards/
   mpeg-4/mpeg-4.htm.
10. J. Rickel and W. L. Johnson. STEVE: A pedagogical agent for virtual reality. In
    International Conference on Autonomous Agents, pages 332–333, 1998.
11. Fabrizio Sebastiani. Machine learning in automated text categorization. *ACM
    Computing Surveys*, pages 1–47, 2002.
12. R. S. Sutton and A. G. Barto. *Reinforcement learning: an introduction*. MIT Press,
    1998.
13. The INTERFACE Project at MiraLab. http://www.miralab.unige.ch/.
14. S. Wilson. Classifier fitness based on accuracy. *Evolutionary Computation*,
    3(2):149–175, 1995.

Intelligent Data Analysis, Decision Making and Modelling Adaptive Financial Systems Using Hierarchical Neural Networks

Masoud Mohammadian and Mark Kingham

School of Information Sciences & Engineering
University of Canberra, Canberra, Australia 2616
`masoud.mohammadian@canberra.edu.au`

Abstract. In this paper an intelligent Hierarchical Neural Network system for prediction and modelling of interest rates in Australia is developed. A Hierarchical Neural Network system is developed to model and predict three months (quarterly) interest rate fluctuations. The system is further trained to model and predict interest rates for six months and one-year periods. The proposed system is developed with first four and then five hierarchical Neural Network to model and predict interest rates. Conclusions on the accuracy of prediction using hierarchical neural networks are also reported.

1 Introduction

Time series data are a special form of data where past values in the series may influence future values. Trends, cycles and non-stationary behaviour in data are often used in predictive models. Such models attempt to recognise patterns and non-linear relationships in the time series data. However due to the nature of data linear models are inaccurate and not suitable. There has been a great interest in nonlinear modelling techniques as well as techniques from Artificial Intelligence fields such as Neural Networks (NNs), Fuzzy Logic (FL) and Genetic Algorithms (GA). The techniques have been successfully used in the place of the complex mathematical systems for forecasting of time series (Azoff, 1994, Bauer 1994, Goldberg 1989, Cox 1993). These new techniques are capable of responding quickly and efficiently to the uncertainty and ambiguity of the system. Neural network has been found useful when the process is either difficult to predict or difficult to model by conventional methods. Neural networks' modelling has numerous practical applications in control, prediction and inference. Neural networks (Azzof, 1994, Grossberg, 1988) can be trained in an adaptive manner to map past and future values of a time series, and thereby extract hidden structure and relationships governing the data.

Investors and governments alike are interested in the ability to predict future interest rate fluctuations from current economic data. Investors are trying to maximise their gains on the capital markets, while government departments need to know the current position of the economy and where it is likely to be in the near future for the well being of a country's people.

Next section the development of a hierarchical neural network system that can be used to predict the fluctuations of quarterly, half yearly and yearly interest rates in Australia. Comparison of the results from single neural networks and proposed hierarchical neural networks system is made. Results of simulations are presented and conclusions and further research directions are given in the last section of the paper.

2   Neural Networks for Prediction of Interest Rates

To predict fluctuations in the interest rate, a neural network system was created. There are a number of steps to perform to create the neural network system. These are: identifying the inputs and outputs for neural network system, pre-processing data if required, and splitting data into training and test suites, defining a neural network system to predict the interest using training data and using the developed neural network on test data to evaluate the accuracy of the prediction of the system.

2.1   Identify the Inputs and Outputs for Neural Network System

To design a neural network system, the actual inputs and outputs must first be determined. There are a number of possible indicators that could be used to predict the interest rate. Some of the main economic indicators released by the Australian Government are: Interest Rate which is the indicator being predicted. The Interest Rate used here is the Australian Commonwealth government 10-year treasury bonds, Job Vacancies is where a position is available for immediate filling or for which recruitment action has been taken, The Unemployment Rate is the percentage of the labour force actively looking for work in the country, Gross Domestic Product (GDP) is an average aggregate measure of the value of economic production in a given period, The Consumer Price Index (CPI) is a general indicator of the rate of change in prices paid by consumers for goods and services, Household Saving Ratio is the ratio of household income saved to households disposable income, Home Loans measure the supply of finance for home loans, not the demand for housing, Average Weekly Earnings is the average amount of wages that a full time worker takes home before any taxes, Current Account is the sum of the balances on merchandise trade, services trade, income and unrequited transfers, Trade Weighted Index measures changes in Australian currency relative to the currencies of our main trading partners, RBA Commodity Price Index provides an early indication of trends in Australia's export Prices, All Industrial Index provides an indication of price movements on the Australian Stock Market, Company Profits are defined as net operating profits or losses before income tax, New Motor Vehicles is the number of new vehicles registered in Australia.

The current interest rate is included in the input indicators to the system as the predicted interest rate is highly dependent on the current rate as there is only likely to be a small fluctuation in the interest rate. The current interest rate gives the neural network system an indication as to the expected "ball park" area of the predicted rate.

2.2   Pre-process Data

In most time series predictions, there is some pre-processing of the data so that it is in a format that the system can use. This may be where data is normalised so it fits within certain boundaries, formatted into an appropriate form for the neural networks system to use. It is also where decisions on how the data is represented are made. In our system, the change from one quarter to the next is used for the GDP and CPI indicators, while the interest rate is the actual reported rate from the Australian Bureau of Statistics. Once the data has been pre-processed, it must be split into some

groups for the training and testing of the system. For this system, the first 2/3^{rds} of the data was assigned to the training set while the other 1/3rd was assigned to the test set. The system uses the training set to train the neural network system. The neural network system is then tested on the test set.

2.3   Hierarchical Neural Network for Prediction of Interest Rates

In this section, a neural network is created to predict the following quarter's interest rate in Australia. The system uses the economic indicators described earlier. A hierarchical neural network system is also used to predict the following quarter's interest rate (as shown in figure 1). A single neural network system is also created. In this neural network all the input parameters were presented and an interest rate prediction was made. In order for the neural network to use the economic data for predicting the following quarter's interest rate, a number of pre-processing steps must be performed. This allows the data to be presented to the neural network in a format that it can easily work with. Data presented to the neural network must fall within certain ranges (usually 0 to +1 or -1 to +1) due to the fact that the network uses a Sigmoid Activation function in its middle (or hidden) layers. The neural network system formats the data for processing where the difference from the current quarter to the previous quarter is used as the data for the input. The change from one quarter to the next is used by all the indicators except the interest rate, where the actual interest rate is used. However there can still be a large range between the smallest and largest values. To reduce this into a more useable range, the data is modified by as follows: *New Data = (current data - Mean) / standard deviation* .The new data that has been calculated represents the distance from the mean value as a fraction of the standard deviation. This gives a good variability to the data, with only a few values that are out of the 0 to +1 or -1 to +1 range. The next step in the pre-processing stage is to squash the data so that it falls between the required range of 0 to +1 for this simulation. For this system, a Sigmoid squashing function is performed. The equation for this step is:

$$Squash\ data = 1 / (1 + exp(-Norm\ Data))) \qquad (1)$$

As well as using the indicators as inputs for the neural network, we also present data to the system that relates to the rate of change in the data, which is the second derivative of the data set. This accents changes in the data set between one quarter and the next. The equation for this is:

$$mov\ diff = (current\ val - previous\ val) / (current\ val + previous\ val) \qquad (2)$$

The above equation is performed on the original data (before any pre-processing steps are performed) and will give a value between -1 and 1. The result from this equation becomes an additional input to the system. Therefore, for each input into the system, an additional input is created, doubling the number of input parameters to the neural network.

2.4   Hierarchical Neural Network System

The hierarchical neural network structure is formed by having the most influential inputs as the system variables in the first level of the hierarchy, the next important

inputs in the second layer, and so on. The first level of the hierarchy gives an approximate output, which is then modified by the second level and so on. This is repeated for all succeeding levels of the hierarchy. The method used in this paper is to split the inputs into a number of related groups. These inputs in these groups are related because they have some common connection between the inputs, such as dealing with employment, or imports and exports. This changes the hierarchy into a two level hierarchy, with the outputs from all the groups in the top layer giving their results as inputs into the bottom layer (Figure 1). Using the economic indicators already indicated we can develop five separate groups. These groups are as follows: Country Group -This group contains Gross Domestic Product and Consumer Price Index, Employment Group - This group contains the Unemployment rate and the Job Vacancies indicators, Savings Group - This group contains House Hold Saving Ratio, Home Loans and Average Weekly Earnings, Company Group - This group contains All Industrial Index, Company Profit and New Motor Vehicles indicators, Foreign Group - This group contains Current Account, Trade Weight Index and also the RBA Commodity Index. These five groups each produce a predicted interest rate for the next quarter. These are then fed into the next layer of the hierarchy where the final predicted interest rate is found, as shown in Figure 1.
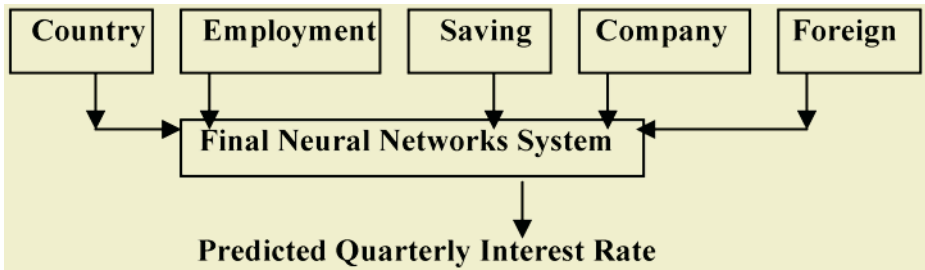
Fig. 1. Hierarchical neural network system for interest rate prediction

The final neural network system uses the predicted interest rate from the five above groups to produce a final interest rate prediction. In the following sections, we first create each of the neural network systems required for the top layer of the hierarchy. We then combine first 4 and then all 5 groups together to form the final Hierarchical neural network system to predict the quarterly interest rate in Australia. As a measure of how well a groups predict the following quarters interest rate, we calculate the average error of the system for the training set and tests sets. This is calculated using the following formula:

$$E = \frac{\sum_{i=1}^{n} abs(Pi - Ai)}{n} \qquad (3)$$

where E is the average error, Pi is the Predicted interest rate at time period i, Ai is the actual interest rate for the quarter and n is the number of quarters predicted. The table below shows the results for the Average error for all groups.

	Test Average	Test Average	Overall Average
Country	1.023	1.023	0.591
Company	0.228	1.290	0.548
Employment	0.352	0.742	0.471
Savings	0.378	0.923	0.534
Foreign	0.387	0.714	0.487

The average error of the Company group shows that there was a slight decrease in the average error of the simulation when compared to the Country group. However, the test set average error is larger than that of the Country group. The performance of the employment group network system for predicting the interest rate on the test data looks slightly better than that achieved by the neural network in the Country or Company groups. The neural network system for foreign group has a number of fluctuations in interest rate that is not accurately predicted by the neural network system. In one quarter (quarter 46), there is a difference between the actual and predicted interest rate of more than three percent. However, the rest of the quarters perform better than this and compare favorably with previously generated neural network system for other groups.

2.5 Building a Hierarchy Neural Network System by Combining Groups

After creating the above neural network systems, we combine them so that we can utilise the information they present and obtain better predictions of each quarters interest rate than any of the neural network systems previously created. A hierarchical neural network system was created by combining first four and then all five of the neural network system from the previous section. The way these groups were combined to form the Hierarchical neural network system. The combined 4 groups consist of: Company group, Country group, Employment group and Saving group. The combined 5 groups consist of: Company group, Country group, Employment group and Saving group and Foreign group. A back-propagation neural network is used with two hidden layers, each consisting of 20 neurons, output layer consists of one node. This was found to produce a quicker and more accurate result than using a single hidden layer. Sigmoid learning is used to predict the following quarters interest rate. The error tolerance was set to 0.0001, the Learning Parameter (Beta) was set to 0.6, momentum (alpha) and Noise Factor were both set to 0. The neural network was trained for 10000 cycles.

	Average Training Error	Average Test Error	Average Overall Error
Four group hierarchical neural network	0.318	0.681	0.428
Five group hierarchical neural network	0.354	0.607	0.431
All indicators single neural network	0.039	0.880	0.296

The Training Average Error, which is the average error, recorded during the training period of 40 quarter's, ranges from a high value of 0.318 for the four groups

down to 0.039 for the All Indicator Neural Network. The All Indicator Neural Network was able to learn the training data almost perfectly. The Test Average Error is the average error recorded during the test period, which is where the system is presented with inputs that it has not been trained on. These compare favourably with all indicator neural network systems had disappointing test average error results.

2.6 Long Term Predictions

A hierarchical neural network system for predicting interest rates, six months (half yearly) from the current quarter is developed followed by a hierarchical neural network system that predicts the interest rate one year ahead of the current quarter.

6 Month Predictions	Training Error	Test Error	Overall Error
Four group hierarchical neural network	0.196	0.794	0.378
Five group hierarchical neural network	0.197	0.813	0.383
All indicators single neural network	0.079	1.421	0.480
One Year Predictions			
Four group hierarchical neural network	0.110	1.248	0.456
Five group hierarchical neural network	0.139	1.114	0.435
All indicators single neural network	0.054	1.320	0.439

The structure and grouping of the inputs and neural network systems in the for the hierarchical neural network system for long term prediction is the same as the hierarchical neural network system described above for quarterly interest rate prediction. From the figures given in the table above, it can be seen that the results for long-term predictions produce similar results to the short-term predictions (i.e. quarterly). Table above shows a comparison between the hierarchical neural network predictions and a single neural network prediction for 6 months and one year. From these results, it can be seen that the hierarchical neural network systems have a much better Test Average Error when compared to the all indicator single neural network systems.

From these results, we can conclude that using 14 economic indicators and training the system for 40 quarters, the hierarchical systems provide much better prediction results than the all indicators single neural network systems.

3   Conclusions

This paper has presented a method in which a hierarchical neural networks system can be used to model and predict the fluctuations of the Australian Interest Rate using Australian economic data. From simulation results it was found that the hierarchical neural networks system is capable of making accurate predictions of the following quarter's interest rate. The results from the hierarchical neural networks were compared to a neural network that used all the indicators as inputs. Long-term predictions

for six months and one year from the current quarter were then undertaken with the hierarchical neural networks systems proving to be more accurate in their predictions than the neural networks systems. These results were found to be similar to those obtained when quarterly interest rates were predicted. Having a time lag for some economic indicators may increase prediction accuracy. Finally the structure of the hierarchical neural network system may affect the performance of the system. This is the subject of our further research.

References

1. Azzof, E. M. (1994). Neural Networks Time Series Forecasting of Financial Markets, New York: John Wiley & Sons
2. Bauer, R. J. (1994). Genetic Algorithms and Investment Strategies. John Wiley & Sons, NewYork.
3. Cox, E. (1993). "Adaptive fuzzy systems". IEEE Spectrum, February, pg 27-31.
4. Goldberg, D. (1989) Genetic Algorithms in Search, Optimisation and Machine Learning. Reading, Massachusetts: Addison Wesley.
5. Grossberg, S. (1988). "Nonlinear Neural Networks: Principles, mechanisms, and achitectures", Neural Networks. Ed 1 pg 17-68.
6. Zadeh, L. (1965). Fuzzy Sets. Inf. Control, vol 8, pg 338-353.

Electricity Load Prediction
Using Hierarchical Fuzzy Logic Systems

Masoud Mohammadian and Ric Jentzsch

School of Information Sciences and Engineering
University of Canberra, Canberra, Australia

Abstract. Electricity load forecasting has been the subject of research over the past several years by researchers and practitioners in academia and industry. This is due to its very important role for effective and economic operation of power stations. In this paper an intelligent hierarchical fuzzy logic system using genetic algorithms for the prediction and modelling of electricity consumption is developed. A hierarchical fuzzy logic system is developed to model and predict daily electricity load fluctuations. The system is further trained to model and predict electricity consumption for daily peak.

1   Introduction

The field of modelling and prediction continues to be one of the principle areas of scientific investigation. There are a numerous real life applications requiring modelling and prediction of the systems behaviours. A large number of techniques for modelling prediction each with their advantages/disadvantages are suggested by researchers overs years (Mohammadian 2000, Bolzern and Fronza. (1986) Ho, Hsu and Yang. (1992)). Time series is a special form of data where past values in the series may influence future values, depending on the presence of underlying deterministic forces. Linear models (based on regression techniques) have been the basis of traditional statistical forecasting models and are not accurate for complex time series. The accuracy of short-term load forecasting in electric power systems has been of great interest due to its strong influence on economic operation of power stations and related industries using the electricity for their operations.

   Large organisations and governments alike are interested in the ability to predict future electricity load fluctuations from existing data. Large organisations are trying to maximise their gains on the capital markets, while government departments need to know the current position of the usage and availability of electricity (i.e energy) in the near future for the well being of a country's people and economy. Large errors in forecast values may force the running of too many power units in power stations causing high reserves. On the other hand lower forecasting values may lead to failure to meet the appropriate electricity requirements. In both cases the operational cost is increased and it creates economical problems, which are undesirable in the current economic competitive world. Electricity consumption follows a complex and nonlinear fashion with various factors affecting the system. A large number of methods to model these relationships have been suggested such as time series, regression models, state space models, and simple intelligent systems (Raju, and Zhou. (1993), Bolzern and Fronza. (1986) Ho, Hsu and Yang. (1992)). Recently techniques from Artificial

Intelligence fields such as Neural Networks (NNs), Fuzzy Logic (FL) and Genetic Algorithms (GAs) have been successfully used in the place of the complex mathematical systems for forecasting of time series (Azoff, 1994).

2   Hybrid Intelligent Systems

Genetic Algorithms and fuzzy logic have a number of characteristics in common. Both are fairly fast and can be used when dealing with nonlinear systems. Fuzzy logic is well suited to storing expert's knowledge in the form of fuzzy rules. For well-defined systems, constructing the fuzzy rules is simple, but for more complex systems where there is a large number of rules, developing a fuzzy knowledge base is difficult and very time-consuming task using trial and error techniques. A genetic algorithm can be used to learn the fuzzy rules of a fuzzy logic system. We can this FL-GA system. Genetic algorithms (GAs) are powerful search algorithms based on the mechanism of natural selection and use operations of reproduction, crossover, and mutation on a population of strings. A set (population) of possible solutions, in this case, a coding of the fuzzy rules of a Fuzzy Logic (FL) system, represented as a string of numbers.

2.1   Encoding and Decoding of Fuzzy Rules by GAs

First the input parameters of the FL system is divided into fuzzy sets. Assume that the FL system has two inputs α and β and a single output δ. Assume also that the inputs and output of the system is divided into 5 fuzzy sets. Therefore a maximum of twenty-five fuzzy rules can be written for the FL system. The consequent for each fuzzy rule is determined by genetic evolution. In order to do so, the output fuzzy sets are encoded. It is not necessary to encode the input fuzzy sets because the input fuzzy sets are static and do not change. The consequent of each fuzzy rule can be any one of the five output fuzzy sets. Assume that the parameter (δ) has five fuzzy sets with the following fuzzy linguistic variable: NB (Negative Big), NS (Negative Small), ZE (Zero), PS (Positive Small), and PB (Positive Big). The output fuzzy sets are encoded by assigning 1 = NB (Negative Big), 2 = NS (Negative Small), 3 = ZE (Zero), 4 = PS (Positive Small), and 5 = PB (Positive Big).

GAs randomly encodes each output fuzzy set into a number ranging from 1 to 5 for all possible combinations of the input fuzzy variables. A string encoded this way can be represented as a string. Each individual string is then decoded into the output linguistic terms. The set of fuzzy rules thus developed is evaluated by the fuzzy logic system based upon a fitness value, which is specific to the system. At the end of each generation, two copies of the best performing strings from the parent generation are included in the next generation to ensure that the best performing strings are not lost. GAs then performs the process of selection, crossover and mutation on the rest of the individual strings. Selection and crossover are the same as a simple GAs while the mutation operation is modified. For mutation an allele is selected at random and it is replaced by a random number ranging from 1 to 5. The process of selection, crossover and mutation are repeated for a number of generations till a satisfactory fuzzy rule base is obtained. We define a satisfactory rule base as one whose fitness value differs from the desired output of the system by a very small value.

3   Fuzzy Logic and Genetic Algorithms
   for Prediction of Electricity Consumption

To predict fluctuations in the electricity load, a hierarchical fuzzy logic system was created. The fuzzy knowledge base (FKB) of this fuzzy logic system was discovered by the use of a GAs. There are a number of steps to perform to create the FKB of the fuzzy logic system these are: identifying the inputs and outputs, pre-processing data if required, and split into training and test suites, defining fuzzy sets and fuzzy membership functions for input and output parameters of the system, setting parameters for genetic algorithm (mutation and crossover), finding an "optimum" FKB for a fuzzy logic predictor using the GAs with training data, using the developed FKB for fuzzy logic on test data to evaluate the accuracy of the prediction of the system.

3.1   Identify the Inputs and Outputs

There are a number of possible indicators that could be used to predict the electricity load. These indicators are: Electricity load which is the indicator being predicted. The electricity load used here is the past electricity consumption (hourly), Predicted Minimum Temperature is the predicted minimum temperature, Predicted Maximum Temperature is the predicted maximum temperature, Actual Minimum Temperature is the actual predicted minimum temperature, Actual Maximum Temperature is the actual predicted maximum temperature, Season is one of the four seasons in the year, Day of the week is one of the seven days of the week, Holiday is one of several public holidays in the year, Time of day is divided here in 48 parts each consisting of 30 minutes.

   The current electricity load is included in the input indicators to the system as the predicted electricity load is highly dependent on the current rate as there is only likely to be a fluctuation in the electricity load from current electricity load. The current electricity load gives the FL-GA system an indication as to the expected "ball park" area of the predicted electricity load (in fact, this gives an indication as to what part of the fuzzy knowledge base to look at).

3.2   Pre-process Data

There are a number of ways in which the raw data from the above indicators could be represented. In this system, the change from one hour to the next is used for the electricity load and temperature indicators. Once the data has been pre-processed, it must be split into some groups for the training and testing of the system. For this system, the first $2/3^{rds}$ of the data was assigned to the training set while the other $1/3^{rd}$ was assigned to the test set. The system uses the training set to learn the Fuzzy Knowledge Base (FKB) for the fuzzy logic system using the GAs. The fuzzy logic system is then tested on the test set using the "best" FKB obtained.

3.3   Fuzzy Membership Functions

For each fuzzy set a fuzzy membership function is assigned. The membership functions used in this study are triangular membership functions. Each input and output parameter is divided into five fuzzy sets as this provided the necessary precision of

the indicators for the system without increasing the number of fuzzy rules required by the FKB to a computationally expensive amount. The linguistic values assigned for the input and output fuzzy sets are: PB - Positive Big, PS - Positive Small, ZE - Zero, NS - Negative Small, NB - Negative Big. The first step in finding the fuzzy sets and membership functions is to find the minimum and maximum range of the data for the indicator. The minimum and the maximum point become the mid point (centre) for the first and last membership function respectively. The range is then divided into 4 equal parts, which become the mid point for the membership functions in between the first and last sets. It was decided that the membership functions have an equal overlap of $1/3^{rd}$ size of the membership function as this was found to give good results over the entire range of data. Using the above method, the fuzzy membership functions for electricity consumption would be found. The output for the fuzzy logic system is the same indicator as per the input electricity load, we use the same fuzzy sets and membership functions of electricity load for the input electricity load parameter.

3.4 Genetic Algorithm Parameters

In order for the genetic algorithm to learn the mapping for the FKB, it must encode its strings to represent the rules in the FKB. Instead of using a binary string as suggested by Goldberg (Goldberg, 1989), the string for the genetic algorithm is encoded using the integer numbers that represent the fuzzy sets. The genetic algorithm had the following parameters: population size = 500, crossover rate = 0.6, mutation rate = 0.01, number of generations= 5000. These figures were decided on after a number of trials and were found to give the best results overall. The genetic algorithm used an elitist strategy so that the best population generated was not lost. Fitness of each chromosome was calculated as the sum of the absolute differences from the predicted hour and the actual hour electricity load. The aim is to minimise the difference between the predicted electricity load and actual electricity load. The fitness was subtracted from a fitness amount, which was decided to be 400000 as it is unlikely that the error amount would be higher than 400000 for the data used for training of the system. The fitness of the system is calculated by the following formula:

$$fitness = {}_{400000} - \sum_{i=0}^{n} abs(PL_i - L_{i+1}) \tag{1}$$

where PL_i is the predicted electricity load, i is the current hour and L_{i+1} is the actual electricity load for the next hour. The genetic algorithm is then run to find the FKB for the system. The genetic algorithm randomly creates an initial population of strings that represent possible FKB's solutions. Using the processes of selection, crossover and mutation, the genetic algorithm is able to find a satisfactory FKB.

3.5 Hierarchical Fuzzy Logic System

In fuzzy systems, there is a direct relationship between the number fuzzy sets of input parameters of the system and the size of the FKB. Kosko and Isaka, (Kosko, B & Isaka, S., 1993) calls this the "Curse of Dimensionallity". The "curse" in this instance is that there is exponential growth in the size of the FKB.

$$k = m^n \tag{2}$$

where k is the number of rules in the FKB, m is the number of fuzzy sets for each input and n is the number of inputs into the fuzzy system. As the number of fuzzy sets of input parameters increase, the number of rules increases exponentially. The hierarchical fuzzy logic structure is formed by having the most influential inputs as the system variables in the first level of the hierarchy, the next important inputs in the second layer, and so on. If the hierarchical fuzzy logic structure contains n system input parameters and L number of hierarchical levels with n_i the number of variables contained in the ith level, the total number of rules k is then determined by equation 3.

$$k = \sum_{i-1}^{L} m^{ni} \tag{3}$$

In this equation m is the number of fuzzy sets. The above equation means that by using a hierarchical fuzzy logic structure, the number of fuzzy rules for the system is reduced to a linear function of the number of system variables n, instead of an exponential function of n as is the conventional case (Raju et al, 1991). A hierarchical fuzzy logic system was used for electricity load prediction. The related inputs are grouped together because they have some common connection among the inputs, such as temperature, time of day etc. This changes the hierarchy into a two level hierarchy, with the outputs from all the groups in the top layer giving their results as inputs into the bottom layer (See Fig 1). These groups are as follows: Predicted Temperature Group -This group contains Electricity Load, Predicted Minimum Temperature, Predicted Maximum Temperature, Time of day, Actual Temperature Group - This group contains Electricity Load, Actual Minimum Temperature, Actual Maximum Temperature, Time of day, Season day Group -This group contains: Electricity Load, Season (a value from 1 to 4 representing each season), Day of the week (two values, one for weekdays and zero representing weekend), Public Holiday (two values, one representing a public holidays and zero representing a working day), Time of day. A (traditional) fuzzy logic system would have a large FKB consisting of over 31,250,000 rules!

As can be imagined, this would require large computing power to not only train the fuzzy logic system with a genetic algorithm, but also large storage and run-time costs when the system is operational. Using a hierarchical fuzzy logic structure (Raju, G.V.S., Zhou, J. 1993), it is possible to overcome this problem. The three groups created for the electricity load prediction each produce a predicted electricity load. These are then fed into the next layer of the hierarchy where the final predicted electricity load is found (see Fig 1).
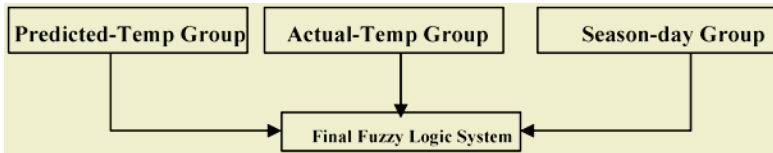
Fig. 1. Hierarchical Fuzzy Logic system for Electricity Load prediction

The total number of rules for the hierarchical fuzzy logic system is: 1455. This is a significant reduction in the number of rules required if we used a (traditional) fuzzy

logic system with a single FKB. Using the inputs from each group the hybrid fuzzy logic and GAs system was used to create three FKB for the fuzzy logic systems (one for each group) to predict the electricity load. The results for the training and testing average error for each fuzzy logic system (of each group) is given below:

	Training Average	Test Average
Predicted Temperature Group	0.793	0.891
Actual Temperature Group	0.653	0.724
Season day Group	0.602	0.842

By combining all three fuzzy logic systems from the above groups a hierarchical fuzzy logic system was created (Fig 1) and then used to predict the electricity load. Hybrid fuzzy logic and genetic algorithm systems was created to learn the FKB for the fuzzy logic systems in the hierarchical fuzzy logic system. It was also used to find the rules used by the final fuzzy logic system to create a hierarchical fuzzy logic system.

The results for hierarchical fuzzy logic system for daily electricity load prediction is given below.

	Training Average	Test Average
Hierarchical Fuzzy Logic System	0.589	0.742

The system is further trained to model and predict electricity consumption for daily peak using a hierarchical fuzzy logic system. The results were found to be similar to those obtained when daily electricity consumption (See below):

	Training Average	Test Average
Hierarchical Fuzzy Logic System	0.689	0.798

4 Conclusion

This paper has presented a method in which a hierarchical fuzzy logic system can be used to model and predict the fluctuations of the electricity load using past electricity usage data. Using a hierarchical fuzzy logic system the number of fuzzy rules in FKB is reduced significantly hence computational times is decreased resulting in a more efficient system. FL-GA system was used to learn the FKBs for each FL system as well as the mapping between different FL systems in the hierarchical fuzzy logic system. From simulation results it was found that the hierarchical fuzzy logic system is capable of making accurate predictions of the electricity load. The hierarchical fuzzy logic systems can also be in parallel. This results in a speed increase during the training period. Using hierarchical fuzzy logic system the reasoning behind a prediction can be explained (either to the customer or senior management).

References

1. Bolzern, P. and Fronza, G.: Role of Weather Inputs in Short-term Forecasting of Electric Load, Electric Power Energy Systems 8(1), (1986) 42-46

2. Dahdashti, A. S.: Forecasting of hourly load by pattern recognition: A deterministic approach, IEEE Trans Power App Systems, 101 (1982) 3290-3294.
3. Goldberg, D.: Genetic Algorithms in Search, Optimisation and Machine Learning. Reading, Addison Wesley, Massachusetts (1989)
4. Ho, K. L. Hsu, Y. Y. and Yang, C. C.: Short term load forecasting using a multi layer neural network with an adaptive learning algorithm, IEEE Trans Power Systems, (1992) 141-149.
5. Kosko, B. & Isaka, S.: Fuzzy Logic, Scientific American. July, (1993) 76-81.
6. Mohammadian, M. (ed.): Advances in Intelligent Systems: Theory and Applications, IOS Press, Amsterdam (2000)
7. Raju, G. V. S. and Zhou, J.: Adaptive Hierarchical Fuzzy Controller, IEEE Transactions on Systems, Man & Cybernetics, Vol 23. **4** (1993) 973-980
8. Zadeh, L.: Fuzzy Sets. Inf. Control, Vol 8. (1965) 338-353.

MEDADVIS: A Medical Advisory System

Zul Waker Al-Kabir, Kim Le, and Dharmendra Sharma

School of Information Sciences and Engineering
University of Canberra, ACT 2601, Australia
z.al-kabir@student.canberra.edu.au
{kim.le,dharmendra.sharma}@canberra.edu.au

Abstract. An advisory system has always been a central need for medical practitioners and specialists for the last few decades. Most of current medical retrieval systems are based on passive databases. Actually, the data stored in any information storage system is a rich source of knowledge, which needs appropriate techniques to discover. This paper introduces the design of a well-structured knowledge-base system that holds patient medical records. The proposed system will be equipped with data mining and AI techniques such as statistics, neural network, fuzzy logic, generic algorithm, etc., so that it becomes an active distributed medical advisory system.

1 Introduction

With advances in soft computing technologies, human expert reasoning and knowledge can be modelled and implemented in a distributed computer system so that it can think and act like a human expert [1–3]. However, in medical applications, where a decision may have fatality effects, machine experts can only play a supportive role. The role becomes very essential in situations that have numerous cases to analyse such as lung tuberculosis and cancer examination based on X-rays: An artificial medical expert can sort patients into different categories depending on their severity levels, so that medical specialists can schedule their examination lists. A medical advisory system is also useful and efficient for quick proposal treatments in simple but common sicknesses as headache, cold, etc.

A number of advisory systems are available for medical and educational purposes. DXPlain, Gideon and IIiad [4] are three online decision support systems that provide possible diagnoses depending on the laboratory reports and symptoms. These systems store a huge amount of medical information on various diseases, but none of them really focuses on the capability of data mining.

This paper is an introduction to a Web-based design of an active Medical Advisory (MEDADVIS) system that is a knowledge base equipped with data mining and X-ray image analysis tools to make it much more supportive and intelligent. The paper is organized as follows. Section 2 outlines the structure of the MEDADVIS system and the operations of its functional units. Section 3 presents some progress on the development of MEDADVIS. The paper ends with a brief discussion on progress difficulty.

2   Structure of MEDADVIS

MEDADVIS is a distributed Web-based advisory system. A full MEDADVIS server has four layers: User access, System memory, Intelligent toolbox and External resources, each with some functional units as summarised in Figure 1.
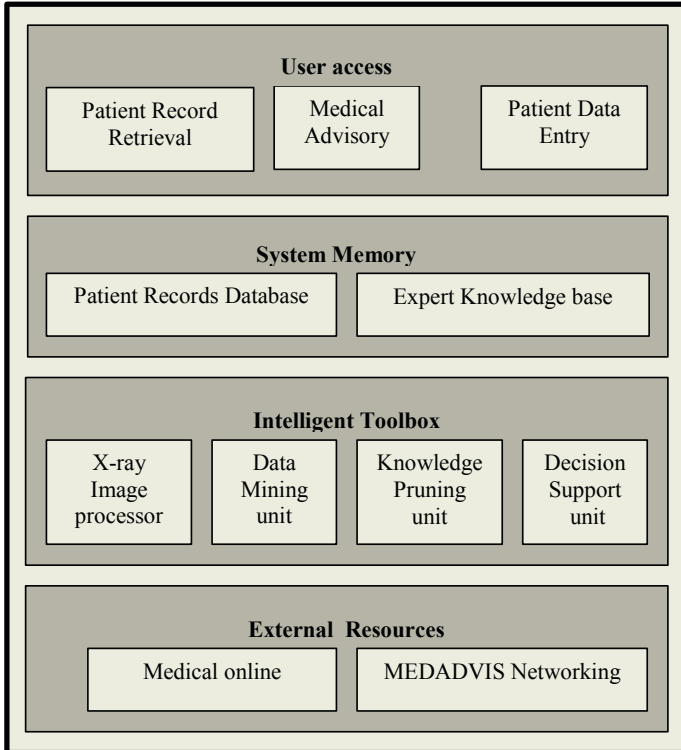
Fig. 1. Structure of a Full MEDADVIS Distributed Web Server

A simple MEDADVIS host may have only essential units to keep patients' records and to link to other MEDADVIS servers. This arrangement enables the host to work at full capability but may be at a slower speed. Task distribution among MEDADVIS servers is a problem that needs consideration for accessibility improvement.

In its initial stage, MEDADVIS focuses on some specific diseases: lung cancer and tuberculosis, and some simple but common sicknesses like cold, headache. An X-ray image processor is a need for lung diseases. Data mining techniques are used to build expert knowledge so that users can make self-examination and obtain treatment advice in cases of simple sicknesses.

a)   User Access

This layer plays the role similar to that of the human sensory system. Authorised users, including patients, and staff from hospitals, clinics, government offices, etc., can use any Web browser to access a MEDADVIS server. Patients can retrieve their

own medical records, including X-ray images. Medical staff can access, for both read and write operations, any patient' record to enter data, e.g. X-ray images, medical examination results, etc. Both patients and medical staff can access the Medical Advisory unit to get advices. The unit can send patients messages, by email or phone, such as special or periodic checking appointments. The unit also enables medical staff to invoke specific tasks in the Decision Support unit (in Intelligent Toolbox layer).

b)   System Memory

The System Memory module plays the role similar to that of human conscious. It is to hold patient records and expert knowledge. A patient's record may be stored in some hosts, one at a central server and some at others depending on recent accesses. The consistency in a distributed storage system is a problem needed to consider. A patient record contains personal details, and medical history of the patient including medical examination reports, medication prescriptions, comments, etc. It also stores relevant findings (smoking habit, sleeping status, food diet, weight, height etc.) that may be useful to trace some diseases like lung cancer, headache, high blood pressure, etc. Furthermore, lung X-ray images are also accumulated with some specific features (e.g. lung, heart status and bronchi statuses, etc.) extracted by the X-ray Image processor (in the Intelligent Toolbox layer).

The Expert Knowledge base stores both medical expertise and reasoning strategies. Medical knowledge is stored in a central server, which may be different from the patients' records central server. Partial knowledge may be distributed to some other servers depending on disease classification, symptoms, etc. Medical knowledge may be presented in the form of rules, which are built from medical theory, specialists' advices, examination practices, etc. Moreover, some intelligent tools may be used to extract rules from patients' records.

c)   Intelligent Toolbox

The Intelligent Toolbox layer plays a role similar to that of human sub-conscious. The tools are operated off line, i.e., whenever a server has available computation time, the tools will be operated. Similar to the cases of highly intelligent persons who can partly control their sub-conscious, some authorised users can invoke MEDADVIS intelligent tools. CGI (Common Gate Interface) and Web service technologies may enable this capability. Some MEDADVIS soft-computing intelligent tools are: X-Ray Image processor; and Data Mining, Knowledge Pruning and Decision Support units. The X-ray Image processor extracts higher-level information from X-ray images of a patient and put them in the patient' record. The Data Mining unit, equipped with soft-computing techniques like statistics, artificial neural network, Fuzzy logic, Generic algorithm, etc. [5-6], enriches the expert knowledge base with additional rules that are hidden in patients' data. The Knowledge Pruning unit enhances the expert knowledge base with the addition of rules' weights. Important rules that have strong effect on decision will have heavier weights, and rules with negligent weights are excluded from the knowledge base. The Decision Support unit is the most intelligent one of MEDADVIS. It examines a patient record and gives necessary advice to the patient, e.g. special and routine checking schedule, preliminary treat-

ment and medication, proposals to see medical general practitioners or specialists if necessary.

The development of the intelligent toolbox is the main task of MEDADVIS project, of which the development progress is report in Section 3.

d) External Resources

Following ancient Indian psychology, human has a very deep layer of sub-conscious, which might be the source of some psychic phenomena like extrasensory or telepathy. Similarly, MEDADVIS has links to online external resources: pharmacy library, hospitals, clinic, diagnostic centres etc. Online pharmacy libraries are important resources that contain medication information, such as new medicine, new diseases and syndromes, which are used by the Decision Support unit. The MEDADVIS Networking unit supplies necessary linkages to make MEDADVIS becoming a distributed system. MEDADVIS servers use both uni-casting and multi-casting techniques for communication to each other based on popular network protocols, such as TCP/IP [7-9]. The development of the MEDADVIS Networking unit is another important task.

3   MEDADVIS Project Development

The MEDADVIS project has been under development at the School of Information Sciences and Engineering, University of Canberra, with some contribution from medical specialist and clinics. The project involves research activities from students at different levels, undergraduate (as final year projects), master and Ph.D., as well as academic staff. This section outlines some project progress and future works.

a) X-Ray Image Processor

An X-ray image is processed through several stages, from image enhancement, segmentation and feature extract [10-11]. There are various enhancement techniques from noise removal, histogram equalising, etc. The enhanced image will then be passed to a segmentation stage, where suspected areas will be labelled and their edges will be identified.

The image processor extracts high-level information from X-ray images. It consults the Expert Knowledge base to find features that have significant role for examination decision. Our current interest is to analyse lung X-ray images for pre-scanning lung cancer and tuberculosis, which are the causes of death of many people around the world [12-14], especially in under-developed countries where there is a shortage of medical specialists. In our previous work [15-16], a segmentation approach was proposed to isolate suspected regions of lung tissues (Figure 2).

Six features were extracted from a lung X-ray image, and then supplied to a fuzzy-rule-based classifier, which sorts an X-ray into one of the five suspecting levels L0 (healthy), L1,…, L4 (most serious). The classifier was tuned on 50 X-Ray images obtained from the MEDPIX online library [17]. The training set consists of 10 normal lung images and 40 X-ray images with various stages of tuberculosis. Based on the examination results given by medical experts, we tuned the fuzzy rules and fuzzy sets to reduce false negatives at the cost of increasing false positives. However, to obtain zero false negative, a high false positive result was obtained (Table 1).
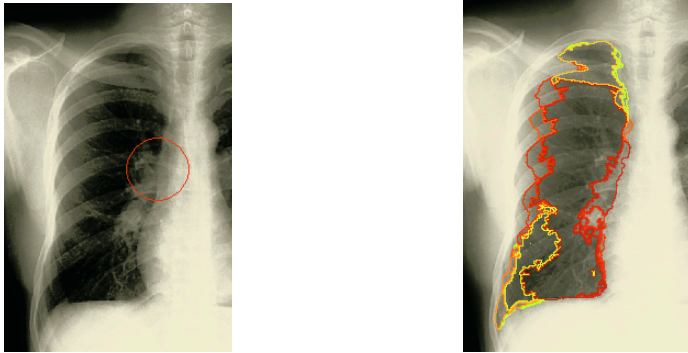
(a) Original X-ray with marked disease area
(b) X-ray with moderate segmentation

Fig. 2. Segmentation of Human Lung X-rays for Suspected Regions

Table 1. Tuning X-ray Classifier for Tuberculosis

Decision	L0	L1	L2	L3	L4
Real experts	1	8	1	28	12
Fuzzy system	0	1	5	12	32
Tuning error					
From Ln − 2	0	0	1	3	0
From Ln − 1	0	0	4	1	20
From Ln +	0	0	0	0	0
Total correct classification	0	1	0	8	12

Our current work aims to find additional features and fuzzy rules to improve the quality of the classifier. The focus will be on finding edges of different parts on a lung X-ray image (e.g. lung, bronchi, trachea, rib, heart, diaphragm, etc.), as well as on identifying nodule, consolidation, etc. We intend to use some technologies like binary morphological processing, erosion techniques [18], etc. Some prospected features are: the sizes, shape and grey levels (brightness) of different objects found in the lung image and the texture of suspected regions (Figure 3).
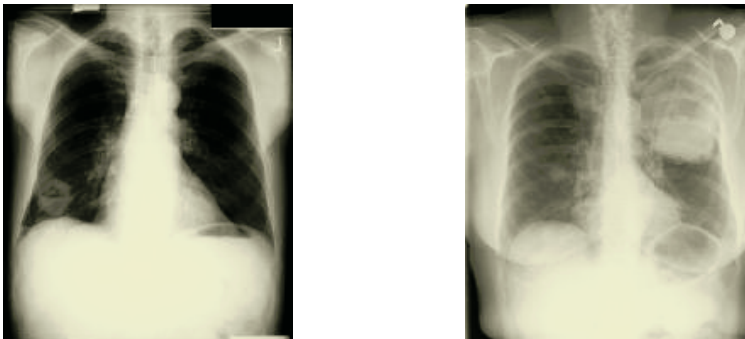
(a) Non–small cell lung cancer. A cavitating right lower lobe squamous cell carcinoma.
(b) Small cell Lung cancer, A large mass is noted in the left mid lung with opacity extending to the upper lung.

Fig. 3. X-ray images of infected lungs [17]

b) Data Mining, Knowledge Pruning and Decision Support

The huge amount of Patients' records in MEDADVIS draws our attention to apply data mining techniques to enrich its expert knowledge with hidden rules. The development of the Data Mining, the Knowledge Pruning and the Decision Support units requires in-depth level of analysis. However, our initial aim focuses on some simple sicknesses, with intention on using some advances techniques such as statistics, neural network, fuzzy logic, etc.

4   Discussion

MEDADVIS is a continuing and long-term project. We are confident that such a system like MEDADVIS is a necessity for our society, both in developing and developed countries. Its development involves various study fields like computer engineering, soft computing, medical sciences, psychology, etc. This paper is just an outline of the MEDADVIS design with some progress reports. Numerous facilities can be added to the system to make it become a more intelligent, friendly, and human-like medical advisor. Advances in techniques like speech recognition, voice and sympathy human face synthesis, etc., make the creation of such a medical advisor possible. The development needs time and the cooperation of members with various backgrounds especially ones from medical sciences. Any contribution to the project is warmly and grateful welcome.

Acknowledgement

References

1. Russel S. and Norvig P.: Artificial Intelligence, A Modern Approach, Second Edition, Pearson Education, 2003.
2. Bratko I.: Prolog Programming for Artificial Intelligence, Third Edition, Pearson Education, 2001.
3. Robinson S., Edwards S. and Yongfa W.: An expert systems approach to simulating the human decision maker, http://portal.acm.org/citation.cfm, accessed on 10 March 2005.
4. Coiera E.: Open Medical- Knowledge Management for Medical Care, AI Systems in Medical Practice, http://www.openclinical.org/aisinpractice.html, accessed on 7 March 2005.

5. Han J., Kamber M.: Data Mining Concepts and Techniques, Morgan Kaufmann Publishers, 2001
6. Fayyad U. M., el al: Advances in knowledge discovery and Data Mining, AAAI press/ The MIT Press, 1996
7. Comer E. D.: Internetworking with TCP/IP Principles, Protocols, and Architectures, Volume 1, 4th Edition, Prentice Hall, 2000.
8. Peterson, L.L. and Davie, B.S.: Computer Networks, A System Approach, 3rd Edition, Morgan Kaufmann, 2003.
9. Makofske D. and Almeroth K.: Multicast Sockets, Practical Guide for Programmers, Morgan Kaufmann, 2003.
10. Russ, J.C.: The Image Processing and Analysis Cookbook, Reindeer Graphics, Inc. 2001.
11. Baxes A. G., Digital Image Processing –Principles and Applications, John Wiley & Sons, 1994.
12. Briggs G.: Pocket Guide to chest X-rays, McGraw Hill, 2004
13. Corne J., el al.: Chest X-Ray Made Easy, Churchill Livingstone, 2002
14. Sharma S., Lung Cancer, Non Small Cell, http://www.emedicine.com/radio/, August 30 2004, Accessed on 27 February 2005.
15. Watman, C. and Le, K., "Gravity Segmentation of Human Lungs from X-ray Images for Sickness Classification", Fourth International Conference on Intelligent Technologies (InTech'03), pp. 428-434. 2003.
16. Watman C. and Le K., Fuzzy Analysis of X-ray Images for Automated Disease Examination, KES 2004, LNAI 3214, pp. 491-497, 2004
17. MEDPIX Website: http://rad.usuhs.mil/medpix/medpix.html
18. Beucher, S. And Meyer, F., The Morphological Approach of Segmentation: The Watershed Transformation, Mathematical Morphology in Image Processing, E. Dougherty, ed., pp. 43-481, New York: Marcel Dekker, 1992.

Towards Adaptive Clustering
in Self-monitoring Multi-agent Networks

Piraveenan Mahendra rajah[1], Mikhail Prokopenko[2], Peter Wang[2], and Don Price[3]

[1] University of Adelaide, Adelaide, SA 5000, Australia
[2] CSIRO Information and Communication Technology Centre
mikhail.prokopenko@csiro.au
[3] CSIRO Industrial Physics
Locked bag 17, North Ryde 1670, Australia

Abstract. A Decentralised Adaptive Clustering (DAC) algorithm for self-monitoring impact sensing networks is presented within the context of CSIRO-NASA Ageless Aero-space Vehicle project. DAC algorithm is contrasted with a Fixed-order Centralised Adaptive Clustering (FCAC) algorithm, developed to evaluate the comparative performance. A number of simulation experiments is described, with a focus on the scalability and convergence rate of the clustering algorithm. Results show that DAC algorithm scales well with increasing network and data sizes and is robust to dynamics of the sensor-data flux.

1 Introduction

Structural health management (SHM) is expected to play a critical role in the development and exploitation of future aerospace systems, operating in harsh working environments and responding to various forms of damage and possible manufacturing and/or assembly process variations. Ultimately, large numbers of sensors will be required to detect and evaluate a wide range of possible damage types within a large and complex structure. Robustness, scalability, reliability and performance verification are also key SHM requirements. A future vision of self-monitoring robust aerospace vehicles includes both local and global SHM systems. The local actions are anticipated to identify, evaluate, and trigger repair for a wide range of damage or defect conditions in aerospace materials and structures. In parallel, global actions should enable dynamic evaluation of structural integrity across large and remote areas. This dual architecture, in turn, entails the need for dynamic and decentralised algorithms used in damage detection, evaluation, diagnostics, prognosis and repair. In order to address these requirements we have chosen to investigate the application of a complex multi-agent system approach to the architecture, and seek to develop methodologies that will enable the desired responses of the system (the remedial actions) to emerge as self-organised behaviors of the communicating system of sensing and acting agents.

CSIRO-NASA Ageless Aerospace Vehicle (AAV) project developed and examined several essential concepts for self-organising sensing and communication networks [1, 3, 8, 9]. Some of these concepts are being developed, implemented and tested in the AAV Concept Demonstrator (AAV-CD): a hardware multi-cellular sensing and communication network whose aim is to detect and react to impacts by projectiles that, for a

vehicle in space, might be micro-meteoroids or space debris. The CD consists of "cells" that not only form a physical shell, but also have sensors, logic, and communications. Currently, each cell contains a small number of passive piezoelectric polymer sensors bonded to an aluminium skin panel in order to detect the elastic waves generated in the structure by impacts. Each AAV cell contains two digital signal processors, one of which acquires data from the sensors, while the other runs the agent software and controls the communications with its neighboring cells. Importantly, a cell communicates only with four immediate neighbors. The CD does not employ centralised controllers or communication routers. A stand-alone Asynchronous Simulator capable of simulating the CD dealing with some environmental effects such as particle impacts of various energies has been developed and used in the reported experiments.

Single cells may detect impacts and triangulate their locations, while collections of cells may solve more complex tasks, for example, produce an impact boundary with desired characteristics or an impact network [3, 9] to pre-optimise secondary inspections and repairs. Some responses could be purely local, while some may require emergence of dynamic reconfigurable structures, with some cells taking the roles of "local hierarchs". In fact, most hierarchical clustering architectures for multi-agent networks are based on the concept of a cluster-head (a local hierarch). A cluster-head acts as a local coordinator of transmissions within the cluster. Often, a cluster-head is dynamically selected among the set of nodes. Moreover, clusters would form and re-form when new damage is detected on the basis of the local sensor signals. In the SHM context, an example of a coordinated task initiated by a cluster head is Active Damage Interrogation (ADI) with a piezoelectric transducer array, emitting moderate to high frequency energy from one or more transducers and using the other transducers as sensors to monitor the energy propagation through the structure. A meaningful ADI scenario may require an emergent formation of clusters of cells with similar damage levels. A cluster-head would then fit the data to a diagnostic model. The sensor-data clustering task has two primary challenges:

- Decentralised clustering: most existing algorithms for clustering focus on how to form clusters, given a file or database containing the items. Decentralization creates the additional complication that, even if a correct classification can be determined with the incomplete information available, the location of items belonging to a class also needs to be discovered [6];
- Dynamic (on-line) clustering: new events may require reconfiguration of clusters – thus, the resulting patterns or clusters have to be constantly refined.

This requires efficient algorithms for decentralised sensor-data clustering in a distributed multi-agent system. In Section 2 we describe an adaptive algorithm enabling self-organisation of stable but reconfigurable impact data clusters, connecting the cells which detected impacts with energies within a certain band (e.g., non-critical impacts). These clusters are expected to reconfigure in real-time if required. Importantly, the cluster algorithm should be robust in the face of changes caused by new damage, cells' failures and node insertion/ removal. Section 3 presents comparative analysis between decentralised and centralised versions of the developed algorithm, followed by a discussion of the obtained results and future work.

2 Adaptive Clustering Algorithms

In this section we describe the decentralised and centralised versions of the developed adaptive clustering algorithm. The input can be described as a series (a flux) of impact energies detected at different times and locations, while the output is a set of non-overlapping clusters, each with a dedicated cluster-head (an AAV cell) and a cluster map of its followers (AAV cells which detected the impacts) in terms of their sensor-data and relative coordinates.

Before presenting details of the algorithms, we would like to position our work in relation to the method of clustering within a fully decentralised multi-agent system, proposed recently by Ogston et al. [6]. The work of Ogston and her colleagues clearly points out the problems of centralised clustering when "data is widely distributed, data sets are volatile, or data items cannot be compactly represented". They present a method for grouping networked agents with similar objectives or data without collecting them into a centralised database, which shows very good scalability and speed in comparison with the k-means clustering algorithm. The method employs a heuristic for breaking large clusters when required, and a sophisticated technique dynamically matching agents objectives, represented as connections in the multi-agent network. The reported clustering results provide a considerable motivation for our effort. Nevertheless, we need to account for specifics of our application: a particular communication infrastructure where each cell is connected only to immediate neighbours in Von Neumann neighbourhood; constraints on the communication bandwidth; dynamic impact scenarios where density of impacts may vary in time and space; a decentralised architecture without absolute coordinates or id's of individual cells on a multi-cellular aerospace vehicle skin; etc. Therefore, our main goal is not a new clustering method *per se*, but rather an evaluation of a simple clustering technique in a dynamic and decentralised setting, exemplified by the CD sensor and communication network, in terms of scalability and convergence, under specific communication constraints. To this effect, we attempted to abstract away some sensor-data features. For example, instead of considering time-domain or frequency-domain impact data, detected and/or processed by cell sensors [8], we represent a cell sensory reading with a single aggregated value ("impact-energy"), define "differences" or "distances" between cells in terms of this value, and attempt to cluster cells while minimising these "distances". This approach can be relatively easily extended to cases where "distances" are defined in a multi-dimensional space. In short, our focus is on evaluating inter-agent communications required by a decentralised clustering algorithm, dynamically adapting to changes.

The algorithm involves a number of inter-agent messages notifying agents about their sensory data, and changes in their relationships and actions. For example, an agent may send a recruit message to another agent, delegate the role of cluster-head to another agent, or declare "independence" by initiating a new cluster. Most of these and similar decisions are based on the clustering heuristic described by Ogston et al. [7], and a dynamic offset range. This heuristic determines if a cluster should be split in two, and the location of this split.

2.1 Clustering Heuristic

Firstly, all n agents in a cluster are sorted in decreasing order according to their impact-energy value x. Then, a series of all possible divisions in the ordered set of agents is

generated. That is, the first ordering is a cluster with all agents in it; the second ordering has the agent with the largest value in the first cluster and all other agents in the second cluster; and so forth (the n-th division has only the last n-th agent in the second cluster). For each of these divisions, the quality of clustering is measured by the total square error:

$$E_j^2 = \sum_{i=1}^{k} \sum_{x \in C_{i,j}} \|x - m_{i,j}\|^2 \ ,$$

where k is a number of considered clusters ($k = 2$ when only one split is considered), $C_{i,j}$ are the clusters resulting from a particular division and $m_{i,j}$ is the mean value of the cluster $C_{i,j}$. We divide E^2 values by their maximum to get a series of normalised values. Then we approximate the second derivative of the normalised errors per division:

$$f''(E_j^2) = \frac{(E_{j+1}^2 + E_{j-1}^2 - 2E_j^2)}{h^2} \ ,$$

where $h = \frac{1}{n}$. If the peak of the second derivative is greater than some threshold T for a division j, we split the set accordingly; otherwise, the set will remain as one cluster.

2.2   Decentralised Adaptive Clustering (DAC)

Each agent is initially a follower to itself, and its followers' list will contain only itself. Each agent is also a cluster-head initially (a singleton cluster). The communication messages (shown in *italic*) can be "flooding" broadcasts or dead-reckoning packets using relative coordinates of their destination on the AAV grid. The algorithm involves the following steps carried out by each cell (agent) which detected an impact with the value x (henceforth, references like "larger" or "smaller" are relative to this value):

1. Keeps broadcasting its *recruit* message initially (*recruit* messages will always contain the followers' list of an agent). This broadcasting is done periodically, with a broadcasting-period P, affecting all agents with values within a particular offset of the value x of this agent, i.e., with values between $x - \varepsilon$ and $x + \varepsilon$. The offset ε is initially set to a proportion α of its agent value: $\varepsilon = \alpha x$.
2. If an agent in a singleton cluster receives a *recruit* message from a "smaller" agent, it ignores it.
3. If an agent p in a singleton cluster receives a *recruit* message from a "larger" agent q in a singleton cluster, it becomes its follower, stops broadcasting its own *recruit* messages and sends its information to its new cluster-head q: an *acceptance-message* with its relative coordinates and the agent-value x. It also stores details of the cluster-head q: the agent-value x_q and relative coordinates.
4. If an agent p in a singleton cluster receives a *recruit* message from a "larger" agent q which does have other followers, it ignores the message: simply because the "larger" agent q would also receive and handle a *recruit* message from p itself (see step 6).
5. If an agent receives an *acceptance-message* from some potential follower agent, it adds the agent involved in its followers' list.
6. If a member of a non-singleton cluster, either the head or a follower, receives a *recruit* message (either from a "larger", "smaller" or "equal" agent), it *forwards* it to its present cluster-head.

7. After *forwarding* a recruit message to its cluster-head, a follower ignores further *recruit* messages until the identity of its head has been re-asserted (as a result of the clustering heuristic being invoked somewhere).

8. The cluster-head waits for a certain period W, collecting all such *forward* messages (the period W, called heuristic-period, should be greater than $2P$). At the end of the heuristic-period, the clustering heuristic is invoked by the cluster-head on the union set of followers and all agents who *forwarded* the messages. The "largest" agent in any resulting cluster is appointed as its cluster-head.

9. The cluster-head which invoked the heuristic notifies new cluster-heads about their appointment, and sends their cluster maps to them: a *cluster-information* message.

10. A cluster-head stops sending its *recruit* messages P cycles before it invokes the clustering heuristic. If it is re-appointed as a cluster-head, it resumes sending *recruit* messages.

11. If an agent receives a *cluster-information* message it becomes a cluster-head. If it was already a cluster-head with a cluster map, it erases that cluster map and accepts the new cluster map. It also *notifies* all its new followers.

12. A follower will periodically get *recruit* messages from its cluster-head. If this does not happen for a while, then it means that this follower is no longer in the followers' list of its cluster-head. Then it will make itself a cluster-head and start sending its own *recruit* messages. The offset of these *recruit* messages will be determined by the offsets it had when it was a cluster-head the last time (not necessarily the same as ε).

Because of the unpredictable timing of the clustering heuristics being invoked in various agents, it is possible that a cluster-head keeps a particular agent as its follower even after its offset ε has changed and this particular agent is now out of range. To counter this, the cluster-head checks its followers' list periodically and removes agents with values out of range. It is also possible that a cell detects a new impact, possibly increasing the agent-value by a large amount. If this agent was a follower, it immediately becomes a cluster-head and *updates* its former cluster-head. The former cluster-head will delete it from its followers' list.

Depending on the nature of the set of agent values, the offset ε may be initially too small to reach any other agent. To counter this, an agent periodically (with a period Δ) increases its offsets exponentially until a certain limit: $\varepsilon_{k+1} = max(2\varepsilon_k, \beta x)$, where $\varepsilon_0 = \varepsilon$ initially, and β is the limit proportion (e.g., the initial ε_0 may be $0.01x$ and after 5 periods the offset would become $\varepsilon_5 = 0.32x$). Alternatively, the increase will stop when the offsets of an agent have been reset by the clustering heuristic. When the clustering heuristic is applied, it may produce either one or two clusters as a result. If there are two clusters, the offset of each new cluster-heads is modified. It is adjusted in such a way that the cluster-head of the "smaller" agents can now reach up to, but not including, the "smallest" agent in the cluster of "larger" agents. Similarly, the cluster-head of "larger" agents can now reach down to, but not including, the "largest" agent (the cluster-head) of the cluster of "smaller" agents. These adjusted offsets are sent to the new cluster-heads along with their cluster maps.

2.3   Fixed-Order Centralised Adaptive Clustering (FCAC)

In order to evaluate DAC, its centralised version was developed. To achieve a congruence between decentralised and centralised versions we define the notion of "reachability" imitating the *recruit* messages. That is, any agent that is within a particular offset of

the agent-value of a cluster-head, i.e., with values between $x - \varepsilon$ and $x + \varepsilon$, is said to be *reachable* from that cluster-head. This allows us to essentially replace broadcast messages with simple data-array searches. Direct cell-to-cell messages (e.g., *forwarding*) are replaced with simple data-array operations, such as inclusion, deletion, merge, split, etc. For example, an addition of a follower p is decided when an agent p is "reachable" from a cluster-head q, and is accomplished by inclusion of p into an ordered list, headed by q.

The Centralised Clustering Algorithm is an iterative process, involving two functions. The first function is a sequential processing of each cluster-head "reaching" to all other agents – analogous to *recruit* messages, while each "reaching" event may trigger an addition of a list of followers to an existing cluster – analogous to *forward* messages. The second function follows the first, and invokes the clusering heuristic by all current cluster-heads, potentially changing the clustering outcome. These functions are repeatedly invoked on the list of agents, and offsets are continuously modified (as described at the end of the previous subsection), until any agent is "reachable" only from a unique cluster-head. Once no cluster-head can "reach" agents that are not its followers, the centralised algorithm terminates.

The quality of clustering is measured by the weighted average cluster diameter [10]. The average pair-wise distance D for a cluster C with points $\{x_1, x_2, \ldots, x_m\}$ is given by

$$D = \frac{\sum_{i=1}^{m} \sum_{j=1}^{m} d(x_i, x_j)}{m(m-1)/2} \,,$$

where $d(x_i, x_j)$ is the Euclidean distance between points x_i and x_j. The weighted average cluster diameter for k clusters is given by:

$$\bar{D} = \sum_{i=1}^{k} m_i(m_i - 1)D_i \ / \ \sum_{i=1}^{k} m_i(m_i - 1) \,,$$

where m_i is the number of elements in the cluster C_i with the pair-wise distance D_i. This metric is known to scale well with the size of data points and number of clusters in a particular clustering. It does not, however, account for singleton clusters, while favouring small clusters.

We would like to point out, at this stage, that neither decentralised nor centralised algorithm guarantees a convergence minimising the weighted average cluster diameter $\bar{D}$. In fact, DAC may give different clusterings for the same set of agent values, depending on the physical locations of the impact points. The reason is a different communication flow affecting the adjustment of the offsets. Each time the clustering heuristic is executed in an agent, its offsets are either left alone or reduced (they are never increased). The scope of agents involved in the clustering heuristic depends on the order of message passing, which in turn depends on the physical locations of impacts. The adjusted offsets determine which agents can be reached by a cluster-head, and this will affect the result of clustering. Therefore, for any set of agent values, there are certain sequences of events which yield better clustering results than others.

The developed centralised algorithm does not simulate all possible sequences of events – hence, the name: Fixed-order Centralised Adaptive Clustering (FCAC). Agent

values are entered in a random order, which may not initiate the "best" ordering of events and yield the best clustering for a particular data set. In other words, centralisation of sensor-data is not a guarantee of a superior performance, and processing of all permutations is prohibitive even for a very small number of elements. On occasions, DAC may even outperform FCAC. Nevertheless, randomisation of data proved to be sufficient for the purposes of our comparison.

3   Experimental Results

We conducted extensive simulations to compare DAC and FCAC algorithms, with the intention of determining whether DAC algorithm is robust and scales well in terms of the quality of clustering and convergence. The quality of clustering is measured by the weighted average cluster diameter $\bar{D}$. The convergence is measured by the number of times (denoted H) the clustering heuristic was invoked before stability is achieved with each data set. This number of times is is chosen as a metric instead of the total time taken in order not to reward DAC for employing essentially parallel computation.

Since we wanted to compare the DAC and FCAC algorithms, we considered the ratio of the weighted average cluster diameters $\bar{D}$ given by FCAC and DAC: $\bar{D}_{FCAC}/\bar{D}_{DAC}$. Similarly, we traced the ratio of numbers of times the clustering heuristic was executed before DAC and FCAC algorithms stabilised with each data set: H_{FCAC}/H_{DAC}. Random values were used as agent values. Each resulting data series was approximated with polynomials, and the best fit was selected according to Mallows' criterion [5]:

$$C_p = RSS_p/s^2 - L + 2p \, ,$$

where p is the polynomial's order, L is the data sample size, RSS_p is the residual sum of squares, and s^2 is given by $RSS_M/(L - M)$, which is the estimate for the residual variance obtained from the complete model with all M regressors included. We used Mallows' criterion because predictive and descriptive ability of the model dealing with noisy data is important.

The scalability analysis considered two scenarios. The first scenario kept the AAV grid array size constant, while increasing the number of impacts detected within it. The second scenario, on the contrary, fixed the number of impacts, while increasing the grid size. In other words, the density of impacts was increasing in the first case, and decreasing in the second.

Effect of Increasing Density. Figure 1[top-left] shows the weighted average cluster diameter ratio $\bar{D}_{FCAC}/\bar{D}_{DAC}$ against the increasing number of impacts. It illustrates that, while the relative performance of DAC decreases with the number of impacts, it scales well and decreases "gracefully". The relative decrease in performance is linear, and DAC gives comparable performance with FCAC for large data sets. Figure 1[top-right] shows that, as the number of impacts increases, DAC needs relatively more and more executions of clustering heuristics to stabilise than the FCAC algorithm. This is expected, as the clustering heuristic needs to be invoked in many different agents with limited information in the DAC algorithm. As the number of impacts increases, DAC takes relatively more computation and time before stabilising. However, the best fit (a
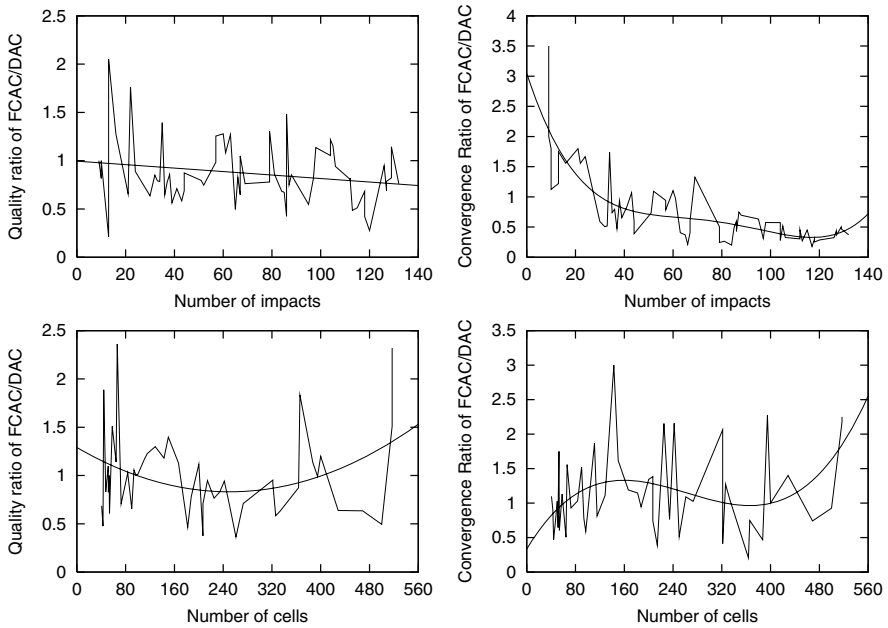
Fig. 1. Top-left: the quality ratio $\bar{D}_{FCAC}/\bar{D}_{DAC}$ for increasing impact density. Top-right: the convergence ratio H_{FCAC}/H_{DAC}; increasing density. Bottom-left: the quality ratio $\bar{D}_{FCAC}/\bar{D}_{DAC}$; decreasing density. Bottom-right: the convergence ratio H_{FCAC}/H_{DAC}; decreasing density.

4-th order polynomial) indicates that DAC may in fact stop losing ground with respect to FCAC for even larger numbers of impacts. The reason is that larger data samples have more possible orderings and FCAC has a lesser chance to process the best one. At this stage it is unclear whether DAC will begin to outperform FCAC, but at least it does not perform worse than 4 times compared to the latter.

Effect of Decreasing Density. Figure 1[bottom-left] shows that an increase in the grid size has an interesting effect on the relative performance of the DAC algorithm. Not only the DAC algorithm scales very well with respect to the network array sizes, but it begins to outperform the FCAC algorithm when the array becomes larger. This is indicated by both the best fit (a 2-nd order polynomial) and the next best (linear) fit. Figure 1[bottom-right] shows that the relative convergence rate of DAC algorithm can be best approximated by a 3-rd order polynomial (again, the next best fit was linear). Both approximations indicate that, as the array becomes larger, the decentralised algorithm begin to outperform the centralised version.

The described experiments simulated impacts detected at the same time – to evaluate scalability with respect to array sizes and the number of impacts. Another factor is dynamics of the impact flux. To analyse robustness of the DAC algorithm in the face of a spatiotemporal impact flux, we developed scenarios where impacts appear periodically, with varying periods. For the FCAC algorithm, still all data were given at once since,

in this case, periodic data insertion is not relevant. Again the ratio $\bar{D}_{FCAC}/\bar{D}_{DAC}$ of weighted average cluster diameters was taken. It was observed that DAC is robust against the timing of impacts: the impact period does not affect the relative performance of the DAC algorithm. In other words, DAC is as good with dynamic data insertion as FCAC is with static (or dynamic) data insertion.

4 Conclusions and Future Work

We presented Decentralised Adaptive Clustering (DAC) and Fixed-order Centralised Adaptive Clustering (FCAC) algorithms for self-monitoring impact sensing networks. The experiments indicate that DAC algorithm can be used to cluster sensor-data, achieving a high quality in a dynamic impact sensing network. The DAC algorithm scales reasonably well with respect to array sizes and the number of impacts, and is robust in the face of a spatiotemporal impact flux. This provides a very good support for deploying other, more sophisticated algorithms in the sensing networks. The density-based algorithms may particularly be relevant in our application: e.g., DBSCAN algorithm would allow us to discover clusters with arbitrary shape [2]. Another avenue is, of course, deployment of hierarchical clustering algorithms. However, in this case, rather than parallelising clustering by partitioning the data set over the AAV network and maintaining a central point as done in P-CLUSTER algorithm [4], we are investigating dynamic hierarchies emerging in response to the constraint on the number of clusters, and the communication protocols required for that. In addition, such a dynamic hierarchy in a self-monitoring impact sensing network may be capable of diagnostic actions and localised SHM responses to global patterns detected by the network.

References

1. Abbott, D., B. Doyle, J. Dunlop, T. Farmer, M. Hedley, J. Herrmann, G. James, M. Johnson, B. Joshi, G. Poulton, D. Price, M. Prokopenko, T. Reda, D. Rees, A. Scott, P. Valencia, D. Ward, and J. Winter. Development and Evaluation of Sensor Concepts for Ageless Aerospace Vehicles. Development of Concepts for an Intelligent Sensing System. NASA technical report NASA/CR-2002-211773, Langley Research Center, Hampton, Virginia, 2002.
2. Ester, M., Kriegel, H., Sander, J., and Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Second International Conference on Knowledge Discovery and Data Mining, 226–231, 1996.
3. Foreman, M., M. Prokopenko, P. Wang. Phase Transitions in Self-organising Sensor Networks. In Banzhaf, W., Christaller, T., Dittrich, P., Kim, J.T. Ziegler, J. (Eds.) *Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life*, 781–791, LNAI 2801, Springer, 2003.
4. Judd, D., McKinley, P., and Jain A. Large-Scale Parallel Data Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* Vol 20, 8, 871–876, 1998.
5. C. L. Mallows. Some comments on C_p, *Technometrics*, vol. 15, 661–675, 1973.
6. Ogston E., Overeinder, B., Van Steen, M., and Brazier, F. A Method for Decentralized Clustering in Large Multi-Agent Systems. Proceedings of the Second International Joint Conference on Autonomous Agent and Multi Agent Systems, 798–796, 2003.

7. Ogston E., Overeinder, B., Van Steen, M., and Brazier, F. Group Formation Among Peer-to-Peer Agents: Learning Group Characteristics. Second International Workshop on Agents and Peer-to-Peer Computing, Lecture Notes in Computer Science series vol. no. 2872, 59–70, Springer, 2004.
8. Price, D., Scott, A., Edwards, G., Batten, A., Farmer, A., Hedley, M., Johnson, M., Lewis, C., Poulton, G., Prokopenko, M., Valencia, P., Wang, P. An Integrated Health Monitoring System for an Ageless Aerospace Vehicle. Proceedings of 4th International Workshop on Structural Health Monitoring, Stanford, September 2003.
9. Wang, P., P. Valencia, M. Prokopenko, D. Price, and G. Poulton. Self-reconfigurable sensor networks in ageless aerospace vehicles. Proceedings of the Eleventh International Conference on Advanced Robotics, 1098–1103, Coimbra, 2003.
10. Zhang, T., Ramakrishnan, R., Livny, M. BIRCH:A New Data Clustering Algorithm and Its Applications. *Data Mining and Knowledge Discovery*, 1(2), 141–182, 1997.

Shared Learning Vector Quantization in a New Agent Architecture for Intelligent Deliberation

Prasanna Lokuge and Damminda Alahakoon

School of Business Systems, Monash University, Australia
{Prasanna.lokuge,damminda.alahakoon}@infotech.monash.edu.au

Abstract. The basic belief-desire-intention (BDI) agent model appears to be inappropriate for building complex system that that must learn and adapt their behaviour dynamically. The contribution of the paper is the introduction of a new "intelligent-Deliberation" process in the hybrid BDI (h-BD[I]) architecture that enables an improved decision making features in a dynamic, and complex environment. Shared learning vector quantization (SLVQ) based neural network is proposed for the intelligent deliberation of the agent model. Paper discusses the benefits of incorporating knowledge based techniques in the deliberation process of the extended h-BD[I] model.

1 Introduction

Intelligent agents have been used for modeling rational behaviors in a wide rage of distributed application systems. Intelligent agent receives various, if not contradictory, definitions; by general consensus, they must show some degree of autonomy, social ability and combine pro-active and reactive behaviors [1]. One solution in particular, that is currently the subject of much ongoing research, is the belief-desire-intention (BDI) approach [2].In particular, the intelligent knowledge improves the quality of the decisions made by the basic BDI model for building complex systems in complex dynamic environments. One of the two key issues in the design of BDI agents are the intention reconsideration policy [8] and the deliberation process in a complex dynamic application. Our previous papers addressed the explicit behaviour in learning and adaptability and their contribution towards improving the agent intention reconsideration process, especially in complex and dynamic real time application systems [3, 4]. The paper describes a novel agent architecture which adopts the shared learning vector quantization neural network in the deliberation for selecting the optimal option in the hybrid BDI agents. The use of SLVQ enhances the autonomous and adaptive behaviour of the hybrid BDI agents in complex environments.

The research described in this article is motivated by a berth assignment problem faced by terminal operators in large container hub ports and investigates the possibility of using intelligent agents for efficient management of vessel berthing operations. Numerous studies have been conducted in vessel and port related operations in the past. Most of the research focused on static vessel berthing system, where the main issue is to find a good plan to assign vessels at the out harbor. Brown [5] used integer programming for berthing vessels, Luca and Andrea [6] used genetic algorithms for the optimization of the resource allocation in a container terminal. However, it is not addressed the incorporation of knowledge based deliberation process in container vessel berthing in the literature. Ideally, the intelligent deliberation to choose most desirable state of affairs or the intention can minimize the unnecessary computations.

The deliberation typically begins by trying to understand what the options available to carry out [7, 10]. For example, in a typical vessel berthing application, when the expected arrival time of a vessel is received, many berth options are available for the vessel to berth. The deliberation process can be decomposed into two distinct functional components: *"options generation"* and *"filtering"* [7]. The few key issues need attentions are: how to predict the available options?, what is the strategy adopted in the filtering process to identify the optimal option?, Is it possible to suggest some alternative options which may use if the optimal option is no longer valid in the present environment? We strongly belief the intelligent knowledge and the domain experience can provide reliable solutions to the above key issues in the deliberation process.

The research is carried out at the School of Business Systems, Monash University, Australia, in collaboration with the Jaya Container Terminal at the port of Colombo, Sri Lanka and Patrick Terminals in Melbourne, Australia. The rest of the paper is organized as follows: Section 2 outlines a vessel berthing system, Section 3 the intelligent deliberation process and the filtering of alternative options. The experiment results are described in section 4 and conclusion is in section 5.

2 A Vessel Berthing System

Shipping lines inform the port the expected time of arrival with other vessel details. The arrival declaration notice sent by shipping lines generally contains the date of arrival, the expected Time of Arrival, the number of containers to be discharged, the number of containers to be loaded any remarks such as Cargo type, berthing and sailing draft requirements etc. The terminal managers investigate the availability and the efficiency of the cranes, trucks, and labour before finding the optimal berth for calling vessels. The port of Colombo has been used as the test bed for our experiments, which handled approximately 1.8 million container boxes annually. The main container terminal is called the "Jaya container terminal" (JCT) which has four main berths called jct1, jct2, jct3 and jct4.

3 The Intelligent Deliberation and the Filtering Strategies

In this section, we describe the intelligent deliberation process which is used to select the appropriate options in a dynamic environment. The two for the selection of we describe the *options selection* and *filtering* strategies adopted in the intelligent deliberation process in the h-BD[I] agents. Our previous papers describe the proposed *h*-BD[I] architecture (Figure 1) which is designed to overcome the limitations in the generic BDI model in dynamic complex environments [4]. We use of SLVQ, reinforcement learning and Adaptive Neuro Fuzzy Inference System (ANFIS) techniques in the *h*-BD[I]model. The letter "*h*" denotes the composition of intelligent tools and "*[I]*" denotes that the extended BDI architecture is capable of handling many alternative options in additions to the intention committed. The whole architecture of the *h*-BD[I] agent is not addressed in the paper and our aim is to present only the Intelligent Deliberation process. The option selection in the intelligent-deliberation process is described in the next sub section.
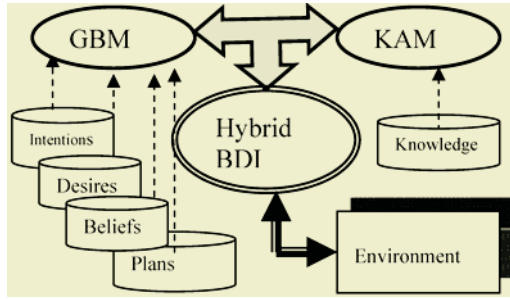
Fig. 1. Main modules in the hybrid Architecture

3.1 The Option Selection in the Intelligent Deliberation

The architecture of the shared LVQ neural net is shown in Figure 2 and is similar to the Kohonen self-organizing map. Each output unit has known class that it represents. Input training vector is belonged to one class or category in the original LVQ algorithm. But is our proposed architecture, we allow to define many output units that an input unit may belongs.
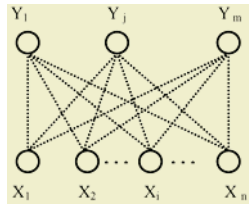
Fig. 2. Shared learning vector quantization neural network

Set of attributes or beliefs in the world, $X \equiv \{x_1\, x_2 \cdots x_n\}$ is the training vector used in the SLVQ. Each attribute, $x_i \,|\, \forall i \in [1, \cdots n]$, where n is the number of attributes used to describe the world. W_j is the weight vector for the jth output unit, where $W_j \equiv (w_{1j}, w_{2j}, \cdots w_{ij}, \ldots w_{nj})$. The closet cluster or the category to the training vector, $t_i \,|\, \forall i \in [1, \cdots noc]$, where noc indicates the number of different classes or clusters in the net. For example, input vector X may belong to the both classes $t_i\, and\, t_j$, but with different distances. Mapping the input vectors to various classes varies from application to application. Original LVQ algorithm introduced does not encourage to assign many classes into a single input vector in the net. But in the real time applications, it is required to define the input vector not only to one class or category but also to some other classes if required. Let $C_j \,|\, \forall j \in [1, \cdots noc]$ is the category of class represented by j^{th} output unit in the net. The outline of the shared LVQ algorithm is describes below.

In the Initialization Phase, define the number of classes or categories belong in the application and initialize the weight vectors. The first m training vectors and their weights are used as weights of the reference vector. The Reference Ratio Matrix (RRM) is defined to represent the degree of how much an input vector is belonged to the output units relative to the closet vector. The closest vector has the lowest value which is equal to zero. Reference vector defined in the original LVQ algorithm represents only one output unit that an input is belonged. In our case, we can define output units that an input vector represents and their degree of the closure. Reference ratio matrix is shown below. Human feed back is very important for the optimization. The introduction of RRM have encouraged the human interaction for finding the optimal option in the proposed agent model.

$$ RRM = \begin{bmatrix} \varepsilon_{1,1} & \varepsilon_{1,2} & \cdots & \varepsilon_{1,noc} \\ \cdots\cdots \\ \varepsilon_{noi,1} & \varepsilon_{noi,2} & \cdots \varepsilon_{noi,noc} \end{bmatrix} \tag{1} $$

Where $\varepsilon_{i,j}$ ($0 \le \varepsilon_{i,j} \le 1$) is called as the relative degree of closeness of the ith input vector for the cluster j. noi and noc are used to indicate the number of input vectors and the number of clusters or classes used in the network.

In the training phase, provide the inputs to the network and compute the weight vector that is closest to the input vector. Euclidean distance of the (similar Self Organizing Maps) input vector i and (weight vector for) j^{th} output unit is shown in equation 2 below,

$$ D_j = \sum_i \left(w_{i,j} - x_i \right)^2 \tag{2} $$

Then find j so that $\left\| X - w_j \right\|$ is the minimum. i.e. D_j has the minimum distance between the cluster j and the input vector i. The two main methods used to update the weights of the output vector are: Case 1: If the input vector belongs to the same cluster which is found as the minimum:

$$ If \left(X_i \in t_j \right) \wedge \left(\left\| X - w_j \right\| \right) \text{ is the minimum then} $$
$$ w_j^{new} = w_j^{old} + \alpha \left(x - w_j^{old} \right) \tag{3} $$

For all the other clusters in the net (which are not closer as the jth output in the net), find the euclidean distances between them and the input vector x. That is,

$$ D_p = \sum_i \left(w_{i,p} - x_i \right)^2 , \text{ where } (\forall j, p : j \ne p). \tag{4} $$

The weight adjustments of other output units are given in equation 5 and 6.

$$ If \left| D_p - D_j \right| < \varepsilon_p \text{ then, } w_p^{new} = w_p^{old} - \alpha \left(x - w_p^{old} \right). \tag{5} $$

If the distance ratio is grater than the reference ratio for output unit p, then we decrease the distance between the output vector and the input vector as shown below:

$$If \left| D_p - D_j \right| > \varepsilon_p \ \text{then} , \ w_p^{new} = w_p^{old} + \alpha \left(x - w_p^{old} \right). \tag{6}$$

In the second case (II), if the input vector does not belonged to the same cluster which is found as the minimum, then the weights of the cluster is updated as follows:

$$If \left(X_i \notin t_j \right) \wedge \left(\| X - w_j \| \right) \ \text{is the minimum then}$$

$$w_j^{new} = w_j^{old} - \alpha \left(x - w_j^{old} \right) \tag{7}$$

Find the preferred class or cluster that the input vector belongs, i.e.

$$If \left(X_i \notin t_p \right) \text{then} \ w_p^{new} = w_p^{old} + \alpha \left(x - w_p^{old} \right) \tag{8}$$

The least distance output unit in the network is the most suitable option for the agent to commit in achieving the desire or goal, which is known as the '*Intention*'. Finally, this SLVQ method produces the other alternative options in addition to the optimal option and can be controlled externally.

3.2 The Filtering Strategy

The introduction of the filtering process enables agent to filter the undesirable options in the environment. The unnecessary means end reasoning or the execution of plans over undesirable options is therefore removed. A threshold called, Option Control Factor, θ^{opt} (OCF) is defined to limit the number of alternative options that an agent can remember during the execution of plans for a committed intention. The positive values in θ^{opt} allow agent to produce many alternative options. The negative values in θ^{opt} produces less number of alternative options, but most of them would be appropriate options to achieve the agent desire. The mean value of the distances between output units μ is calculated as given in equation 9.

$$\mu = \frac{\sum_{i=2}^{noc} \left(D_i - D_{\min i} \right)}{noc - 1} \tag{9}$$

The selection of jth alternative option in the filtering strategy is therefore given as

$$\left(D_j - D_{\min i} \right) \leq \mu \pm \theta^{opt} \tag{10}$$

4 Experiment Results

A vessel berthing data of the main container terminal (*JCT*) in the Port Colombo has been used in our experiments. The SLVQ neural network is trained with using three months real data which contains 32 attributes. Few attributes used in the training are: the vessel berthing and sailing drafts, the vessel crane outreach requirements, the expected arrival time of the vessel, the number of container boxes in the vessel, the length of the vessel, the expected time of completion of berths, the berth drafts, the equipments availability of berths, the berth productivities etc. Four output vectors are considered to represent the individual berths in the terminal.
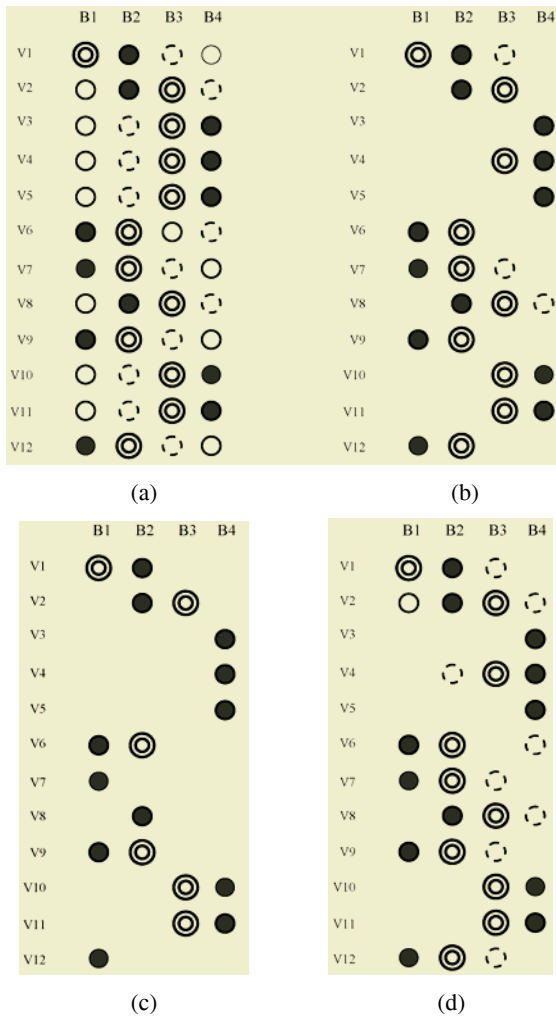
Fig. 3. (a) Berth Options selected (b) Berth options after filtering (c) Berth options when (d) Berth Options when OCF= +0.1

Figure 3(a) shows the berth options selected for a window of 12 incoming vessel at the *JCT*. Figure 3(b) indicates the berths selected after applying the filtering process described earlier. Figures 3(c) and 3(d) show the berth selection when θ^{opt} (OCF) is -0.3 and +0.1. Figure 4(a) and 4(b) indicate the new berths selected (before and after the filtering process) according to the environmental changes observed. B1, B2, B3 and B4 in the following Figures denote the berths *JCT1, JCT2, JCT3* and *JCT4*.

The solid circle indicates the optimal berth or the intention, circle with the double line indicates the second best berth option, circle with the dotted line shows the third alternative option identified and the circle with the single line is the least possible berth option identified by the intelligent deliberator for the incoming vessels.
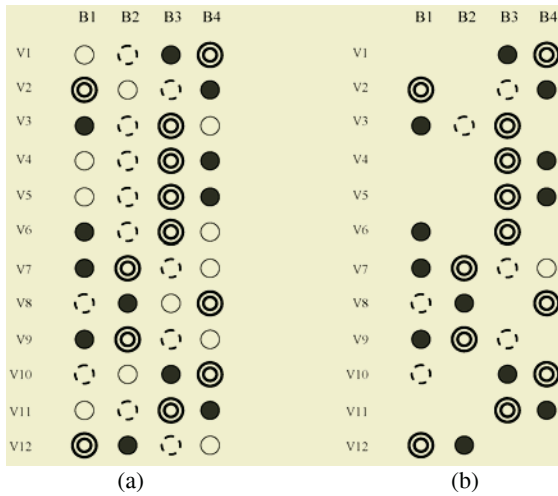
Fig. 4. (a) Berth options selected (b) Berth options after filtering is applied

For example, first row of the Figure 3(c) has shown that the Berth B2 has been se-lected as the optimal berth for the incoming vessel v1 and berth B1 as the second best option. But due the changes observed in the vessel berth draft requirement, it is noted that the current berths B2 or B1 are now not in a position to accommodate the vessel v1. Therefore, the intelligent deliberator has produced the new optimal berth for the incoming vessel v1 as berth B3 and berth B4 as the second option. The investigation we conducted showed interesting results. Primarily, the introduction of an agent's intelligent knowledge in the deliberation process assure in producing the optimal solution and other alternative options in a dynamic environment.

It is observed that the intelligent-deliberator produces less number of alternative options but most desirable ones when negative values are applied to the OCF. Figure 3(b) shows that vessel *v1* has been originally assigned to berth *B2 (JCT2)*. It is noted that the events received: "*vessel draft*" and "*crane change*" caused the intelligent deliberator to reconsider its original intention of assigning the vessel to *JCT2*. Figure 4(b) shows that berth *B3 (JCT3)* has been chosen as the new optimal berth due to the changes observed in the environment. The filtering process applied has correctly cho-sen the optimal berth and other alternative options for the incoming vessels as shown in Figure 3(b) and 4(b). The intelligent deliberator uses the past experience and the intelligent knowledge in the option selection and filtering assuring the optimal results in the execution of plans.

5   Conclusion

The deliberation process of the generic BDI agent is improved with the introduction of SLVQ neural network. The use of SLVQ based neural network is able to predict not only the optimal option but also the other desirable alternative berthing options for incoming vessels. The introduction of θ^{opt} enables users to control the number of alternative options produced in addition to the committed intention. Further this

minimizes the unnecessary deliberations at every time. The autonomous adaptive behavior to environmental changes is therefore assured in the *h*-BD[I] agent which enhanced the autonomous and adaptive behavior in the generic BDI model.

References

1. Wooldridge M. and Jennings N.R: Intelligent agents: Theory to practice, the knowledge engineering review. V. 10, no 12, p. 115-152, 1995.
2. Georgeff, M, B. Pell, M. Pollack, M. Tambe, and M. Wooldridge: The Belief-Desire-Intention Model of Agency, Springer Publishers, 1998.
3. Lokuge D.P.S and Alahakoon L.D: Hybrid BDI Agents with Improved learning capabilities for Adaptive Planning in a Container Terminal Application. In proceeding of IEEE/WIC/ACM International conference on Intelligent Agent Technology (IAT), pp. 120-127, IEEE computer society, China, 2004..
4. Lokuge D.P.S. and Alahakoon L.D: Motivation based behavior in Hybrid Intelligent Agents for Intention Reconsideration Process in Vessel Berthing applications. In proceeding of the 4th International conference on Hybrid Intelligent Systems (HIS04), IEEE computer Society, Japan, 2004.
5. Brown, G. G., Lawphongpanich, S. and Thurman, K.P, Optimizing Ship Berthing: Naval research logistics, 41, 1-15.1994.
6. Gambardeella L. M and Rizzoli A.E and Zaffalon M: Simulation and planning of an intermodal container terminal. Special Issue simulation on Harbour and Maritime Simulation, 1998.
7. Wooldridge, M. 2000: Reasoning about rational agents. The MIT Press, Cambridge, MA.
8. Kinny D and Georgeff M: Commitment and effectiveness of situated agents. In Proceeding of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI-91), pp. 82-88, Sydney, Australia, 1991.
9. Bloodsworth P, Greenwood S and Nealon J: A Generic Model for Distributed Real-time Scheduling Based on Dynamic Heterogeneous Data. Intelligent and Multi-agent Systems, Springer-verlag, Eds Lee J, et.al, 2004.
10. Wooldridge, M: An Introduction to Multi-Agent Systems. John Wiley & Sons Ltd, 2002.

Patterns for Agent Oriented e-Bidding Practices

Ivan Jureta[1], Manuel Kolp[1], Stéphane Faulkner[2], and T. Tung Do[1]

[1] Information Systems Research Unit, University of Louvain, Belgium
kolp@isys.ucl.ac.be
[2] Department of Management Sciences, University of Namur, Belgium
stephane.faulkner@fundp.ac.be

Abstract. Today high volume of goods and services is being traded using on-line auction systems. The growth in size and complexity of architectures to support online auctions requires the use of distributed and cooperative software techniques. In this context, the agent software development paradigm seems appropriate both for their modelling, development and implementation. This paper proposes an agent-oriented patterns analysis of best practices for online auction. The patterns are intended to help both IT managers and software engineers during the requirement specification of an on-line auction system while integrating benefits of agent software engineering.

1 Introduction

The emergence and growing popularity of Internet-based electronic commerce has raised the challenge to explore scalable global electronic market information systems, involving both human and automated traders [1].

Online auctions are a particular type of Internet-based markets, i.e., worldwide-open markets in which participants buy and sell goods and services in exchange for money. Most online auctions rely on classical auction economics (see e.g., [2]). "An auction is an economic mechanism for determining the price of an item. It requires a pre-announced methodology, one or more bidders who want the item, and an item for sale" [3]. The item is usually sold to the highest bidder. An online auction can be defined as an auction which is organized using a software system, and accessible to participants exclusively through a website.

This paper proposes an agent-oriented patterns analysis of best practices for online auction. Providing agent-oriented patterns for such systems can reduce their development cost and time, while integrating benefits of agent-orientation in software development. Agent-oriented development is a modern software engineering paradigm for analyzing and designing distributed and dynamic systems such as online auctions. An agent is an autonomous software entity that is responsive to its environment, proactive (in that it exhibits goal-oriented behavior), and social (in that it can interact with other agents to complete goals). Multi- agent systems involve the interaction of multiple agents, both software and human, so that they may achieve common or individual goals through cooperative or competitive behavior.

Patterns of best practices in the online auction domain will facilitate the development of new auction systems, by clearly showing the functional and non functional aspects that are particularly valued by auction participants. Patterns – which are reusable solutions to recurring development problems – for online auction have already been proposed in the literature (see e.g., [4]). However, these patterns have been specified using object-oriented concepts. Consequently, they do not show agents as intentional, autonomous, and social entities. In addition, the patterns usually do not integrate best practices identified in current operating auction systems on the Internet. [5] only provides a global architecture of a basic online auction system in the context of object-oriented software development. GEM [1] provides system architecture for developing large distributed electronic markets but it only addresses the system's basic functionalities required to organize trading among agents. It provides patterns without treating intentional aspects, and uses agents at implementation level.

The rest of this text is organized as follows. Section 2 discusses the *i** agent framework we have used to represent the patterns. Section 3 describes the best practices agent patterns we have analyzed in the domain of online auctions. Due to lack of space, we only present 4 patterns. For a more complete catalogue, we refer the reader to [6]. Section 4 concludes the text and discusses some further work of our research.

2   The i* Framework

In the following, we analyse each pattern using the *i** framework [7]. *i** is an agent-oriented modelling framework used to support the early phase of requirements engineering, during which the analyst represents and understands the wider context in which the system will be used. The framework focuses on intentional dependencies that exist among actors, and provides two types of models to represent them: a strategic dependency (SD) model used for describing processes as networks of strategic dependencies among actors, and a strategic rationale (SR) model used to describe each actor's reasoning in the process, as well as to explore alternative process structures.

The main modelling constructs of the *i** framework are Actors, Roles, Goals, Softgoals, Resources, and Tasks (See Fig. 1). Both the SD and SR models can represent dependencies among Actors or Roles. A dependency describes an "agreement" (called *dependum*) between two actors: the *depender* and the *dependee*. The *depender* is the depending actor, and the *dependee,* the actor who is depended upon. The type of the dependency describes the nature of the agreement. *Goal* dependencies represent delegation of responsibility for fulfilling a goal; *softgoal* dependencies are similar to goal dependencies, but their fulfilment cannot be defined precisely; *task* dependencies are used in situations where the dependee is required.

Actors are represented as circles; dependums – goals, softgoals, tasks and resources – are respectively represented as ovals, clouds, hexagons and rectangles; dependencies have the form *depender → dependum → dependee.*

In *i**, software agents are represented as Actors. Actors can play Roles. A Role is an abstract characterization of the common behaviour of an Actor in some specific context (e.g., a consumer, a salesman, a buyer, a seller, etc.).

3   Best Practices Patterns

Online auctions are highly dynamic processes which involve numerous participants. Their structure changes rapidly to reflect the entry and exit of bidders as well as the impact of their behaviour on the price of the item being auctioned. The most common mechanism for on-line sales are the "English", "Vickrey", "Dutch", and "first-price sealed bid" auctions [3]. We briefly describe them below.

English Auction. Each bidder sees the highest current bid, can place a bid and update it many times. The winner of the auction is the highest bidder who pays the price bid, i.e. the final auction bid that this bidder placed.

First-Price Sealed Bid Auction: Each bidder makes a single secret bid; the winner is the highest bidder, and the price paid is the highest bid.

Vickrey Auction: Each bidder makes a single secret bid; the winner is the highest bidder. However, the price paid is the amount of the second highest bid.

Dutch Auction: The seller steadily lowers the price of the item over time. The bidders can see the current price and must decide if they wish to purchase at that price or wait until it drops further. The winner is the first bidder to pay the current price.

Today's online auction offer features more complex to those that automate the traditional auction mechanisms (e.g. user authentication, auction setup, auction item search, bidding, ... [4]). In addition to enhancing the user experience, these additional features are essential to the commercial success of an online auction. This paper focuses on best practices to better understand and build these features. The analysis is applicable on any type of auction as far as the participant type is concerned: both the seller and buyer may be either customer and/or business. It is independent of the auction mechanism (english, vickrey, dutch, ...), as long as it involves a single seller and many buyers.

Some of the features can be introduced in the system in several ways, requiring comparison and evaluation. To select the most adequate alternative, we represent relevant system qualities (e.g., security, privacy, usability, etc.) as softgoals and use contribution links to show how these softgoals are affected by each alternative, as in the Non-Functional Requirements framework [8].

Due to lack of space, we only present 4 patterns. For a more complete catalogue, we refer the reader to [6]

Proxy Bidding. Online auctions can last for several days, making it impossible for human buyers to follow the auction in its integrity, as is the case in traditional ones. Proxy bidding allows buyers to specify their maximum willingness to pay.

A procedure is then used to automatically increase their bid until the specified maximum is reached, or the auction is closed [9]. This enables human buyers to be represented in the auction, without requiring their physical presence in order to interact with their *Buyer* agent. Proxy bidding is applicable only when English Auction rules are enforced in the auction.
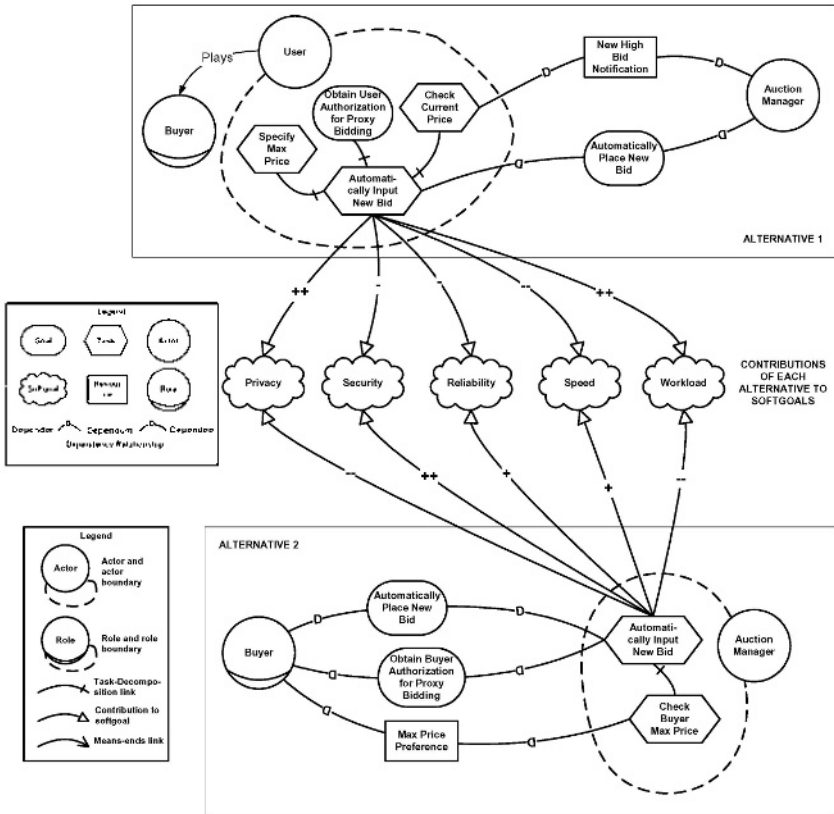
Fig. 1. Two alternative responsibility assignments of the Proxy Bidding. Positive (favorable) (+) and negative (not favorable) (-) contributions of each alternative structure aid in selection

Proxy bidding can be introduced in the basic online auction in several ways in terms of responsibility assignment. Two alternatives are shown in Fig. 1. Each one is represented as a simple Strategic Rationale model. A series of softgoals have been selected as criteria for alternative comparison – *Privacy*, *Security*, *Reliability*, *Speed*, and *Workload*. These are non-functional requirements [8] for the information system and have been selected according to issues often raised in e-commerce system design, online auction design, etc.

Dispute Resolution (Fig. 2). The trade settlement that follows the closure of the auction may not be successful for many reasons (e.g., late deliveries, late payment, no payment at all, etc.). It then results in dispute that can require mediation by a third party in order to be resolved. The third party (here, a *Negotiation Assistant*) can be either a software agent that manages an automated dispute resolution process, or a human mediator.

The *Negotiation Assistant* collects *Buyer* and *Seller Arguments*, and makes them available to both parties. On the basis of these *Arguments* and its *Solution Knowledge*

Base, the agent *Selects Solution* – both the *Buyer* and the *Seller* depend on the agent to *Suggest Solution* to their dispute.
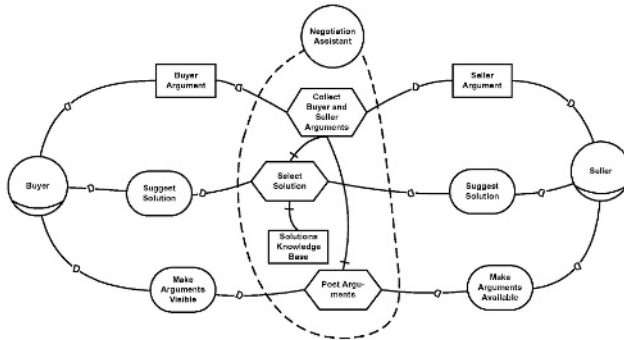
Fig. 2. Strategic Rationale model of the Dispute Resolution pattern with focus on *Negotiation Assistant* agent rationale

Payment. Payment can be accomplished in numerous ways in the context of an online auction. They can be either managed (in part) through the online auction – e.g., credit card based transactions –, or outside the scope of the online auction information system (OAIS) – e.g., cash, checks, etc. The payment choice of auction participants is not repetitive and differs according to the payment cost, convenience, and protection [10]. Consequently, it is important to take these criteria into account when structuring an online auction.
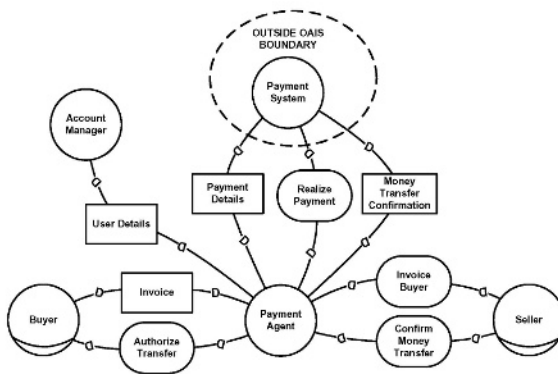
Fig. 3. Payment pattern

In the Payment pattern, the *Payment Agent* (specialization of the *Negotiating and Contracting Agent* [11]) mediates the payment interaction between the *Seller* and the *Buyer*. This agent depends on the *Account Manager* for data on *Users*, which is then used in providing *Payment Details* to the *Payment System*. In addition to user identification, *Payment Details* should also contain transaction-related data. The *Payment Agent* depends on the *Payment System* to *Realize Payment* and to provide *Money Transfer Confirmation*, which is used to *Confirm Money Transfer* to the *Seller*. The

Payment System is outside the boundary of the online auction. Upon closure of the auction, the *Seller* depends on the *Payment Agent* to *Invoice Buyer*. The *Buyer* depends on the *Payment System* to provide *Invoice* and in return, the *Buyer* is expected to *Authorize Transfer*.

The pattern structure in Fig. 3 is adapted to PayPal [5] and all common credit card based payment systems. Any of these payment systems intervenes in the pattern as the *Payment System*, which is specialized in money transfers.

Fraud Detection. Fraud is common in online auctions. In 2002, more than 33.000 fraud complaints were filled [12]. Fraud issues are strongly related to trust and reputation, and should be accounted for an online auction system, in terms of specific parts of the system that are specialized in fraud detection activities. These may be human and/or automated agents, provided that the latter dispose of high performance automated methods for fraud detection.

The fraud detection pattern (Fig. 4) is based on two main agents: *The Fraud Complaint Centre* and the *Fraud Detector*. The *Fraud Complaint Centre* gathers all the *Fraud Complaints* posted by *Sellers* and/or *Buyers*. The Internet Fraud Complaint Centre (IFCC) typically plays this role. Besides, *Users* also expect active fraud detection. They depend on a *Fraud Detector* to *Detect Fraud* in a secure and reliable way.
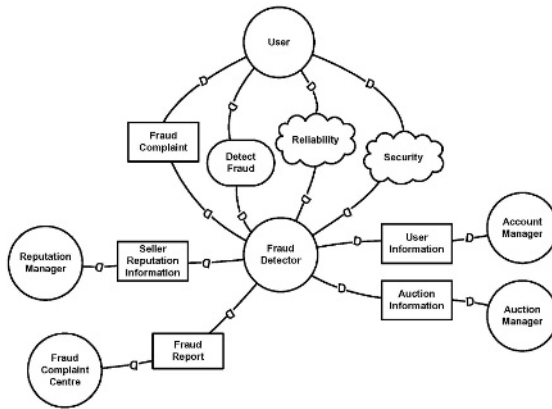
Fig. 4. Fraud Detection pattern

The *Fraud Detector* requires specific *User* information and therefore depends on the *Reputation Manager* for *Seller Reputation Information*, the *Account Manager* for *User Information*, and on the *Auction Manager* for *Auction Information*.

The fraud detection methods proposed in [13] are based on statistical methods and association analysis. This is particularly helpful to detect *shilling*, one of the most common fraud practice in online auction. The Seller tries to hike up the prices in auction by placing buy bids under distinct aliases or through associates.

4   Conclusion

Online auctions have become increasingly popular in e-business transactions [12]. Companies require such systems to be developed on tight budgets and in short time. Patterns of best practices of online auctions can provide significant aid in the development process of such systems.

This paper explores such patterns, by analyzing some advanced online auctions functionalities through the lens of the agent paradigm. Compared to the literature, our approach is innovative in several respects: we consider that multi-agent systems are particularly adapted to modelling and implementing online auction systems; we provided the i^* agent-oriented modelling perspective of each of the patterns we consider and we focused on specifying best practices in current online auction systems.

There are limitations to our work. We have not provided other dimensions than the i^* (social and intentional) ones for the patterns. This is well beyond the scope of this paper as it requires much more time and space. As future work, the patterns will be modelled using UML-based notations and formally specified with the Z language.

References

[1] Rachlevsky-Reich B., et al.: "GEM: A Global Electronic Market System". In *Information Systems* 24, pp. 495-518, 1999.

[2] Bikhchandani S., de Vries S., Schummer S., Vohra R.V.: "Linear Programming and Vickrey Auctions". In Dietrich B., Vohra R.: *Mathematics of the Internet: E-auction and Markets*. Springer, 2001.

[3] Beam C., Segev A.: *Auctions on the Internet: A Field Study*. Working Paper 98-WP-1032, Haas School of Business, University of California, Berkeley, USA, 1998.

[4] Re R., Braga R.T.V., Masiero P.C.: "A Pattern Language for Online Auctions Management". *Proc. of the 8th Conf. on Pattern Languages of Programs* (PLoP'2001), 2001.

[5] Kumar M., Feldman S.I.: "Internet Auctions" *Proceedings of the 3rd USENIX Workshop on Electronic Commerce*, Boston, September, 1998.

[6] Jureta I., Kolp M., Faulkner S. And Do T., *On-line Bidding Patterns*, IAG working paper 132/05, University of Louvain, Belgium, 2005.

[7] Yu E.: *Modelling Strategic Relationships for Process Reengineering*. Ph.D. thesis, Dept. of Computer Science, University of Toronto, 1994.

[8] Chung L.K., Nixon B.A., Yu E., Mylopoulos J.: *Non-Functional Requirements in Software Engineering*. Kluwer, 2000.

[9] Kurbel K., Loutchko I.: "A Framework for Multi-agent Electronic Marketplaces: Analysis and Classification of Existing Systems". *Proceedings of International Congress on Information Science Innovations* (ISI'2001), Dubai, U.A.E., 2001.

[10] Li H., Ward R., Zhang H.: *Risk, Convenience, Cost and Online Payment Choice: A Study of eBay Transactions*. Georgia Institute of Technology Working Paper, 2003.

[11] Papazoglou M.P.: "Agent-Oriented Technology in Support of E-Business". *Communications of the ACM* 44-4, pp. 71-77, 2001.

[12] Wood C. "Current and Future Insights From Online Auctions. In M. Shaw, R. Blanning, T. Strader & A. Whinston (eds.) *Handbook on Electronic Commerce*, Springer, 2004.

[13] Shah H.S., Joshi N.J., Wurman P.R., *Mining For Bidding Strategies on eBay*, Springer, 2002.

Innovations in Intelligent Agents

Jeff Tweedale[1] and Nikhil Ichalkaranje[2]

[1] Air Operations Division, Defence Science and Technology Organisation
Edinburgh SA 5011, Australia
Jeffrey.Tweedale@dsto.defence.gov.au
[2] School of Electrical and Information Engineering,
Knowledge Based Intelligent Engineering Systems Centre
University of South Australia, Mawson Lakes SA 5095, Australia
Nikhil.Ichalkaranje@unisa.edu.au
http://www.kes.unisa.edu.au

Abstract. This paper introduces the invited session of "Innovations in Intelligent Agents" being presented at the ninth International Conference on Knowledge-Based & Intelligent Information & Engineering Systems. This session concentrates on the use of the Belief-Desire-Intention (BDI) architecture to produce Multi-Agent Systems (MAS) that enable "human-centric" teaming. A brief summary of the innovations that led to the evolution of research into this topic is also provided with the papers that detail the concepts raised in this session.

1   Introduction

KES 2004 was one the highest profiling conferences of its type in the world and the presenters are regarded as experts in their field. Many of these presenters were eminent scientists and engineers who shared their knowledge on current and innovative research. Most of this research was described by attendees as being leading edge research that had been inspired through innovation and an inescapable quest to improve their field of endeavour. Last year our conference was fully subscribed with only the highest quality papers across a range of prescribed, invited and special sessions. Attendance was the best in its history and the organisers suggested stricter guidelines for future conferences to enable attendees more freedom to attend parallel sessions [1].

Many fields of computing are realising their potential through advances in technology. Observations within the DNA computing domain show how it has been used to nurture evolvable hardware that now boosts the ability to adapt and self heal [2]. Given this continued improvement in "silicon to biological" interfaces [3], our ability to control a sequence of predictable, repeatable processes could soon be feasible [4]. Combining this with advances in nano-technology we could witness a paradigm shift from radical surgery to less invasive DNA correction [5, 6]. Researchers have already determined how the human brain works and compared it to advanced machines. Many of these researchers agreed that we require a current microprocessor to run at 25GHz to emulate the brain [7]. The speed is slowly being delivered, however the Von Neumann architecture does not have the right structure to support human thought, because symbolic association and pattern recognition both require parallel processing [8]. This issue should be addressed with the next generation of microprocessors from Intel, AMD, IBM and HP. For instance the launch of the Intel's latest Pentium 4TM (The Smithfield PD 840, a dual core micro-processor) was intended to arrest the heat/speed

problem by introducing parallel processing. Intel's Pat Gelsinger announced the Averill microprocessor, to be launched in 2006, would provide ten times the current Pentium 4TM performance [9]. This evolution should pave the way in providing truly smart intelligent agent systems.

Each leap in technology tends to ignite new research into Artificial Intelligence because the preceding domains become bogged down when attempting to communicate with humans [10, 11]. This disconnect is generally cognitive in nature and is being addressed using agents capable of "intelligence through interaction". Research into Distributed Artificial Intelligence (DAI) and intelligent agents have benefited from agent-oriented technology [12]. This session has been structured to introduce a number of innovations that show how the Belief-Desire-Intention (BDI) architecture has the potential of becoming the method of choice for complex reactive systems. It also describes how future trends in agent technology could be categorized using "human-centric" teaming that enables operators to actively engage or be in the next generation of intelligent agent systems. Such systems should reflect heavily on "intelligence with interaction", autonomous target-tracking, exhibit powerful reasoning and learning, with cognitive models of trust and reasoning models based on time in BDI architectures.

2   Session Papers

This session includes five papers. The first paper by Ichalkaranje and Tweedale is on the trends and directions of future agent design. This paper illustrates how the agent technologies have developed along with increasingly sophisticated techniques to adapt to solution of problem rather than adhering to core principles of its design. The paper traces the origin of agent technology and its principles especially those that reflect heavily on "intelligence with interaction" with an emphasis as an important component of modern Distributed Artificial Intelligence field. Furthermore this paper points out that human centric interaction in agents will be the future direction by using the evolutionary term "next generation of agents". This human centric interaction can be achieved by adding one or more ideal attributes such as learning, coordination and autonomy, resulting in a "truly smart agent" [13].

The second paper by Park, Fullam, Han and Barber illustrates agent technologies deployed in the Unmanned Aerial Vehicle (UAV) target-tracking domain. This is the classic example of implementation of agents reasoning and decision making characteristics, in dynamic, unpredictable and coordination intense environments. This paper describes how agents capabilities have enabled: (1) coordination of the tracking of multiple targets among a set of UAVs, (2) identification of the best subset of assigned UAVs from which to collect location information, and (3) evaluation of location information accuracy. These capabilities aid the efficient and effective collection and verification of target location information [14].

The third paper by Sioutis, Ichalkaranje presents a novel architecture for implementing powerful reasoning and learning systems based on Intelligent Agents. It is particularly suited for agent systems working in complex and cooperative dynamic environments. This architecture came into existence to overcome the limitations that were confronted while researching the use of agents in such environments. It focuses on one of the ideal agent attributes "learning" by taking advantages of three primary

human decision making models such as Bratman's BDI, Rasmussen's decision ladder and Boyd's OODA loop. This paper describes a brief background of the underlying conceptual reasoning model, followed by details of the implemented hybrid architecture and concludes with some final remarks and future work [15].

The fourth paper by Jarvis, Corbett and Jain is on a Belief-Desire-Intention model of confidence in an agent's intention. A cognitive model of trust is presented. Trust plays a fundamental role in multi-agent systems in which tasks are delegated or agents must rely on others to perform actions that they themselves cannot do. The concept of trust may be generalised and considered as a level of confidence in one's predictions of another agent's future behaviour. This has applicability beyond that normally ascribed to trust: for instance, one may be confident that a particular agent's intentions are hostile, and that this will be borne out by particular behaviours. In this paper authors present a cognitive model of trust in which the central component is a Belief-Desire-Intention model or 'theory of mind' of a person or agent that evolves over time [16].

The fifth paper by Jarvis, Corbett and Jain is on reasoning about time in a Belief-Desire-Intention architecture. It outlines a possible strategy for implementing the representation of, and the capacity to reason about actual time. The philosophical roots of the Belief-Desire-Intention model lie in Bratman's formulation of an intention theory of planning, in which he sought to make sense of the notion of future-directed intention. The principal implementations of BDI follow the original Procedural Reasoning System model, which has a sound logical basis, exemplified by the Logic Of Rational Agents (LORA). While the LORA formulation has a temporal logic component, however, this does not translate into any ability for the agent to reason about actual time. Being able to reason about actual time would bring significant benefits for BDI agents. Among these benefits are richer descriptions of the relationships between beliefs, desires and intentions; and the ability for agents to communicate deadlines and to plan and schedule activities in a cooperating group. Given a suitable representation of temporal knowledge, an agent could learn about the temporal aspects of its own actions and processes, and this knowledge could be used as input to the planning process [17].

3   Summary

The research papers presented in this session contribute to many new ideas including new direction in multi-agents systems, agent technology for coordinating UAV target tracking, a novel architecture for implementing reasoning and learning systems, a novel cognitive model of trust and a strategy for implementing the representation of and the capacity to reason about actual time. This collection of papers documents key aspects of a maturing technology that is emerging as the paradigm of choice for MAS being developed for complex reactive systems in a hostile environment.

Acknowledgements

References

1. M. G. Negiota, L. Jain, and B. Howlett: "Kes'2004 - conference programme", presented at The 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems, 2004, pp. 1 - 40.
2. T. Higuchi: "Evolvable hardware (EHW) and its industrial applications," presented at The 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems, 2004, pp. 6 - 7.
3. S. Vedantam: "Brain cells, silicon chips are linked electronically," Last visited 12 Oct 2004, http://wwww.mercola.com/2001/sep/12/silicon chips.htm
4. A. Stoica: "Advanced evolvable hardware for autonomous systems," presented at The 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems-Tutorial CD, 2004.
5. G. Bouriano: "Nanoenergetics, nanomaterials, nanodevices, nano-computing: Putting together the pieces," European Materials Research Society, pp. 1 - 35, 2004.
6. S. Chou: "65-nanometer process technology extends benefit of Moore's law," Last visited 11 Oct 2004, http://www.intel.com/labs/features/si08042.htm?iid=search&
7. J. Henning: "Spec CPU2000: measuring CPU performance in the new millennium," in IEEE Computer, 2000, vol. 33, Issue 7, pp. 28-35.
8. H. C. Anderson: "Why artificial intelligence isn't (yet)," in AI Expert Magazine, July 1987, pp. 1-9.
9. S. M. Fulton: "Duelling Multi-Cores: Intel and AMD fight for the future, "Tom's Hardware Guide, TPG Publishing AG, Last visited 29 Apr 2005, http://www20. tomshardware.com /business/20050330/multicores-01.html
10. J. Rasmussen: Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering: Elsevier Science Inc, 1986.
11. L. A. Suchman: Plans and Situated Actions: The Problem of Human-machine Communication: Cambridge, University Press, 1987.
12. S. J. Russel and P. Norvig: "Artificial Intelligence: A Modern Approach," 2nd ed.: Prentice-Hall, New Jersey, 2003.
13. N. Ichalkaranje, J. Tweedale,: "Multi-Agents Systems: New Directions," presented at 9th Knowledge-Based Intelligent Information & Engineering Systems (KES) Conference, Melbourne, 2005.
14. J. Park, K. K. Fullam, D. C. Han, and K. S. Barber: "Agent Technology for Coordinating UAV Target Tracking," presented at 9th Knowledge-Based Intelligent Information & Engineering Systems (KES) Conference, Melbourne, 2005.
15. C. Sioutis, N. Ichalkaranje: "Cognitive Hybrid Reasoning Intelligent Agent System," presented at 9th Knowledge-Based Intelligent Information & Engineering Systems (KES) Conference, Melbourne, 2005.
16. B. Jarvis, D. Corbett, and L. Jain: "Beyond Trust: A Belief-Desire-Intention Model of Confidence in an Agent's Intentions," presented at 9th Knowledge-Based Intelligent Information & Engineering Systems (KES) Conference, Melbourne, 2005.
17. B. Jarvis, D. Corbett, and L. Jain: "Reasoning about Time in a BDI Architecture," presented at 9th Knowledge-Based Intelligent Information & Engineering Systems (KES) Conference, Melbourne, 2005.

Multi-agent Systems: New Directions

Nikhil Ichalkaranje[1] and Jeff Tweedale[2]

[1] School of Electrical and Information Eng., University of South Australia
`Nikhil.Ichalkaranje@unisa.edu.au`
[2] Airborne Mission Systems, Defence Science and Technology Organisation
`Jeffrey.Tweedale@dsto.defence.gov.au`

Abstract. This paper presents the trends and directions of future agent design. These trends focus on advances in Agent-Oriented programming techniques and describe how agent technologies have developed with increasingly sophisticated techniques. The paper traces the origins of agent technology and describes related principles, especially those that reflect heavily on "Intelligence with Interaction". This interaction, especially with operator-in-the-loop scenarios, is demonstrated as a critical component of modern Distributed Artificial intelligence. Unfortunately many applications fail to address this fact, creating an interruption to ongoing research. Therefore we believe the next generation of agent systems should focus on human centric interaction to achieve a level of shared- intelligence made possible with modern computational techniques.

1 Introduction

Leading researchers believe that agent technology is a result of the convergence of many related technologies within the computer science domain [1, 2]. The key domains include: object-oriented programming, distributed computing and Artificial Intelligence (AI). The important aspect contributing to the evolution of agents is their ability to offer intelligence with interaction. This ability facilitated the merger of techniques traditionally reserved for AI and Human Computer Interaction (HCI) research. The traditional AI focus was dominated by research into pure intelligence. Unfortunately when it came to interacting externally with agent-to-agent or human-to-agent, bottlenecks occurred, especially when attempting to troubleshoot faults within such systems [3]. These restrictions interrupted ongoing development of AI systems during the late 1990's [2]. A renewed surge in AI research occurred after the Distributed Artificial Intelligence (DAI) community acknowledged the need for "Intelligence with Interaction". This new domain incorporates agent technology in order to form the interaction. Thus agents provided the means to bridge the gap between humans and machines facilitated by creating intelligence with interaction between agents and other external contacts.

Although the scope of DAI technology relies heavily on intelligence with interaction, most of the current systems or applications lack the vision of utilizing human computer interaction. Recently there is a shift towards human-agent centric interaction with many researchers contributing to the field [4]. We discuss the current status of agent technology and the authors' point of view of future agent technology and the desirable drive towards a human-agent centric system.

1.1   Agent Classification

The question of how to classify agents has been debated since the first scripts originated. Some researchers preferred functionality, while others used utility or topology. Agreement on categories has not been rationalised, however three classes dominate. These classifications include: mobility, reasoning (such as reactive-deliberative), attributed models (such as those used in autonomy), planning, learning, and cooperative/collaborative/communicative agents [3]. Nwana chooses to classify agents using mobility, reasoning, learning, information and data fusion. Noting these differences, pressure is also emanating from within the DIA community to include interaction within the beliefs, desires and intentions (BDI) agent paradigm with interaction into this definition. BDI can be used to achieve human like intelligence, however such research cannot be fitted into the above definition of truly "smart agents", as most applications still lack the main ideal characteristics of "coordination and learning". To be acknowledged as the next interruption along the AI paradigm continuum, intelligence with interaction must be demonstrated as the major impediment of next generation AI applications by incorporating agents capable of learning and self managed group coordination (teaming).

1.2   Agents: Where They Are Going?

The next major step towards a new generation of agents is to incorporate "human-centric" functionality. The current evolution of agent systems has concentrated solely on agent-to-agent interaction and fails the above because they lack "social ability" [4]. Wooldridge describes social ability as "… the ability to interact with other agents and possibly humans via some communication language [5]". Based on this statement, we would like to suggest that 'interaction' with humans is not limited to a communication language but should be extended to include other means of interaction, such as observation and adaptation. We would also like to suggest that smart agents could work in conjunction with humans if they adapt or translate their skills (communication, learning and coordination) rather than the functions they conduct. These concepts help focus our development of agents adopting a human centric nature by merging one or more of the ideal attributes used in coordination, learning and autonomy.

1.3   Building Blocks: MAS Theory

We have chosen to use the JACK BDI agent architecture and intend to demonstrate that it is the paradigm of choice for complex reactive systems [6]. We believe the architecture needs some refinement and the development environment extended in order to suit most systems where human involvement is the most crucial factor. Examples of such systems could readily be used in the cockpit of an airplane and in complex reactive systems involving critical human decision making in real time, such as surveillance and control operators building situation awareness. MAS capable of coordination and learning could conceivably be incorporated to automate or assist operators in conducting these roles.

1.4 Human Centric Agents: An Answer to Early Automation Pitfalls

One of the major issues in early human-machine automation was the lack of focus on the human and the human cognitive process [7, 8]. This was due to aggressive introduction of automation based on perceived needs and the tools available. Agent research has matured and applications using agent-oriented techniques based on BDI reasoning are becoming common. Early Agent models based on theories such as BDI appeared as an attractive solution during human based decision support due to their human like intelligence and reasoning behavior [8]. Existing agent models can act as independent substitutes for humans and their cognitive decision support behavior. If this human like substitute should fail at a critical point, the humans/agent ability to regain control would be impaired, creating a situation where the man/machine also fails to achieve the goal. The answer to this deficiency is provided by developing machine-assistants conducting advisory and decision support roles to human-operators experiencing critical or high cognitive workloads. Intelligent agent technology has matured and it is attractive enough to implement such machine-assistants in the role of more independent cooperative and associate assistants [9].

2 New Directions for Agent Development

Intelligent Agent (IA) technology is gaining commercial strength, with applications being developed in domains as diverse as meteorology; manufacturing, war gaming, capability assessment and UAV mission management. Furthermore, commercially supported development environments are increasingly being provided complete with design methodologies and reference architectures. In our opinion some early standards are beginning to emerge with key researchers collaborating using agreed methodologies (i.e. LORA [10], JACK [11] and SOAR [12]). BDI MAS applications have also emerged as the paradigm of choice for the development of complex distributed systems. This combination of indicators confirms a maturing of agent-orient technology. The commercial transition, however, will only be acknowledged after a number of real-world problems are solved. To achieve this transition, increased collaboration between organisations such as academia, industry and defence is required. Although we consider the transition has begun, the next step is to address the availability of skills within the community to support agent-oriented programming techniques.

From a software engineering perspective, one would expect to gain major benefits from intelligent agent technology through its deployment in complex distributed applications such as virtual enterprise management and the management of sensor networks. Given the introduction of standards and an improved programming support base, we believe the underlying theories of cognition will continue to prove adequate for large-scale software developments. The key theories enveloping system development and BDI MAS have long pedigrees in terms of their use and they are now gaining commercial support in the form of tools and maturing development environments. This foundation and new cognitive concepts will be used to help to transition agent research into commercial-strength applications.

2.1   Alternative Reasoning Models

We perceive a need to develop alternative reasoning models to accommodate cognitive concepts. Experience gained by using the JACK implementation of BDI enabled us to implement a reasoning MAS. This model employs the logic that shapes the program flow, not the underlying cognitive theory. To go further we needed to implement "higher level" reasoning by incorporating an Observe Orient Decide and Act (OODA) loop. This application is still being developed as we are also attempting to incorporate a learning facility. When complete, this model is intended to provide the cognitive environment required to implement human-agent interaction.

The development of applications using the current generation of intelligent agent is not routine. The distinguishing feature of the agent-oriented paradigm is that an agent can have autonomy over its execution as opposed to the object-oriented paradigm, where there is no notion of autonomy because objects invoke the services that they require from other objects. Being able to provide a model capable of more intuitive reasoning that will support MAS cognitive frameworks will help. Behaviour acquisition is a major impediment to the widespread application of human-agent interaction in the IA paradigm. However acquiring behaviour is necessary to achieve the degree of autonomy required. This operation can be a major undertaking and one in which there is little support. The problem can be likened to the knowledge acquisition bottleneck that beset expert systems of the 1980's [7]. There is a need for principled approaches to behaviour acquisition, particularly when agents are to be deployed in behaviour rich applications such as enterprise management. Cognitive Work Analysis has shown promise in this regard, but further studies are required, especially if the knowledge is being acquired through self-learning.

2.2   Human Interaction

The requirements for autonomous operation are weakened when the need for human interaction is introduced, often increasing the risk of failure. Rather than having purely agent-based applications, we need cooperative applications comprised of teams of agents and humans that are interchangeable. Agent-based advisory systems can be seen as a special case of cooperative applications, but we see the interaction operating in both directions – the agent advises the human, while the human also directs and influences the reasoning processes of the agent. Existing architectures provide little in the way of support for this two-way interaction. It should also be noted that in many applications, such as the cockpit of an airplane or surveillance and control operators building situation awareness, full autonomy is not desirable due to the potential lethality of the decision. In these instances an agent should provide advice that is acted upon by a human. For this reason the interface between MAS and operators must be improved, with either agents and/or humans being able to adopt a variety of tasks and responsibilities as the situation or environment changes. Learning has an important role to play in both cooperative and autonomous systems. The alternative is to provide the agent with predefined behaviours (or context) based on a priori knowledge of the system that can be modified manually based on experience gained with the system and/or its operators. This has worked well in practice and we expect that it will remain the status quo for the immediate future.

3   Conclusion

In summary, we expect that intelligent agents will retain the architectural foundations discussed. They are expected to evolve with reasoning models and methodologies to become standards within the MAS community. Furthermore, better support for human/agent teams will continue to mature and form the basis of a new class of intelligent decision support applications.

Intelligent agents have been used to develop many applications over the past decade, especially in operations analysis where the modelling of "human-like" intelligence and decision-making components are required. We have witnessed the transition from object-oriented to agent-oriented software development and MAS are maturing to a point where standards are being discussed. The future design of agent-based systems will rely on the collaboration of academia, industry and Defence with a focus on solving a significant real-world problem. Our initial efforts are focused on demonstrating a multi-agent system capable of learning, teaming and coordination.

Acknowledgements

References

1. Decker, K.: A Vision for Multi-agent Systems Programming. In: Dastani, M., Dix, J. and El Fallah Seghrouchni, A., (eds.): Programming Multi-agent Systems, First International Workshop, PROMAS 2003. Lecture Notes in Artificial Intelligence, Vol. 3067. Springer, (2004) pp. 1-17.
2. Wooldridge, M.: An Introduction to MultiAgent Systems. Wiley, (2002)
3. Nwana, H. S.: Knowledge Engineering Review, Cambridge University Press, (1996), Vol. 11, No 3, pp. 1-40, Sept 1996.
4. Heinze, C., Goss, S., Josefsonn, T., Bennett, K., Waugh, S., Lloyd, I., Murray, G. and Oldfield, J.: Interchanging Agents and Humans in Military Simulation. AI Magazine, 23(2), (2002), pp. 37-48.
5. Wooldridge, M., Jennings, N. R.: Theories, Architectures, and Languages: A Survey, Intelligent Agents: ECAI-94 Proceedings of the Workshop on Agent Theories, Architectures, and Languages, eds Wooldridge, M., Jennings, N. R., Lecture Notes in Artificial Intelligence, vol.890, Springer-Verlag, Heidelberg, (1995), pp. 1-39.
6. Kinny, D., Georgeff, M., & Rao, A.: A Methodology and Modelling Techniques for Systems of BDI Agents, Van de Velde, W. & Perram, J. W., (Eds.), Agents Breaking Away: Proceedings of 7th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW '96, LNAI 1038, Springer-Verlag, (1996), pp. 56-71.

7. Russel, S. J., and Norvig, P. (Eds): Artificial Intelligence: A Modern Approach, 2nd Ed., Prentice-Hall, (2003), New Jersey.
8. Bratman, M, E.: Intention, Plans and Practical Reasoning. Harvard University Press, (1987)
9. Urlings, P.: "Human Machine Teaming", PhD thesis, University of South Australia, (2004)
10. Wooldridge, M., and Jennings, N.R.: "Intelligent Agents: Theory and Practice", Knowledge Engineering Review, Vol. 10, No. 2, June 1995
11. Unknown: Jack Intelligent Agents Manual, Agent Oriented Software, Melbourne, (2002)
12. Laird, J. E., Newell, A., and Rosenbloom, P. S.: Soar: An architecture for general intelligence, Artificial Intelligence, (1987) 33(1), pp. 1-64.

Agent Technology
for Coordinating UAV Target Tracking

Jisun Park, Karen K. Fullam, David C. Han, and K. Suzanne Barber

The Laboratory for Intelligent Processes and Systems
Electrical and Computer Engineering Department
The University of Texas at Austin
{jisun,kfullam,dhan,barber}@lips.utexas.edu

Abstract. This paper illustrates three agent technologies deployed in the Un-manned Aerial Vehicle (UAV) target tracking domain. These capabilities enable: (1) coordination of the tracking of multiple targets among a set of UAVs, (2) identification of the best subset of assigned UAVs from which to collect location information, and (3) evaluation of location information accuracy. These capabilities aid the efficient and effective collection and verification of target location information.

1 Introduction

Dynamic and unexpected events are the defining characteristics of numerous application domains, requiring decision-makers to solve problems with insufficient resources and time. In the Unmanned Aerial Vehicle (UAV) target tracking domain, UAVs continually collect and deliver information about moving enemy targets to a tactical operation center (Central Command). UAVs must coordinate with each other to divide tracking responsibilities; uncoordinated execution of tracking activity among UAVs may result in inefficient information collection and resource usage. Second, the Central Command must decide which UAVs are the most appropriate information providers of the target locations. Finally, target location information may be inaccurate, perhaps due to sensor failure, noise caused by target movement, or communication delays. Therefore, the operations center must gauge the accuracy of incoming information. These challenges provide an appropriate environment for multi-agent system deployment. Agent technologies proposed in this research provide novel capabilities for (1) coordinating the tracking of multiple targets among a set of UAVs, (2) identifying the best subset of assigned UAVs from which to collect location information, and (3) evaluating the accuracy of location information

2 Overview of Technical Details

This section explains the technologies deployed in the UAV target tracking domain. Fig 1 depicts the overall picture of target tracking activity by UAVs and the Central Command. Three major technologies deployed in this target tracking are explained in the following sections: autonomous action selection by agents controlling the UAVs and information provider selection and information valuation by an agent in the Central Command.
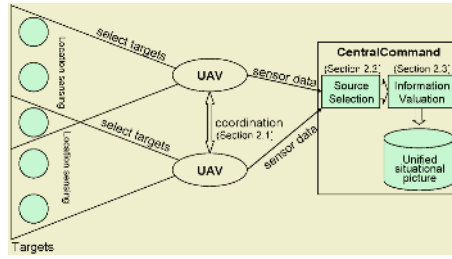
Fig. 1. UAV Target Tracking Activity Overview

2.1   Autonomous Action Selection

An agent must decide what actions to perform (e.g., selecting which targets to track in Fig 1) to affect the environment according to the desires the agent holds. When faced with complex desires (i.e., multiple goals), an important characteristic of an autonomous agent is the ability to decide which goals to pursue and in what order. Markov Decision Processes (MDP) are used to calculate the value of an action as the expected value to be received immediately and in the future through continued rational action. The basic reward structure for MDP models is limiting since, if the agent has multiple goals to achieve, their achievement state must be represented as part of the decision problem. This research separates the concepts of desires (achievement states) and domain states to enable reasoning at various abstraction levels. For example, the domain states are defined as a product of state variables, $S_{domain} = V_1 \times V_2 \times \ldots \times V_L$. If an agent desires to sequentially visit multiple goal states in the domain, the actions that the agent selects will be different depending on which of the goal states the agent has already visited. Desire states of the agent can be defined as a product of the goal variables (boolean values indicating whether each goal has been achieved), $S_{desire} = G_1 \times G_2 \times \ldots \times G_K$. The states represented in an MDP must be able to differentiate between the same domain states when the agent has different desire states, hence the total decision space is the cross product of the domain space with the desire space, $S = V_1 \times V_2 \times \ldots \times V_L \times G_1 \times G_2 \times \ldots \times G_K$.

Macro actions are combinations of primitive actions whose execution will lead to specific domain states of interest [5]. In this domain, macros can be designed through domain analysis as a sequence of heading changes that will reach a given location. Factoring using macros separates analysis of the domain characteristics (S_{domain}) from the desires states (S_{desire}), allowing reuse of domain analyses when desires change. The estimated costs of macro actions are used to impose structure over S_{desire}, relating the various achievement states to each other in terms of costs and rewards. The motivations for reasoning in the desire space include: (1) the desire space is much smaller than the complete decision space, and (2) the structure of the desire space can be exploited algorithmically by the agent during computation of macro action values. Full details on this method can be found in [2].

When multiple agents interact, either through design or chance encounter, the idea of task allocation comes into play. Each agent is assigned goals by its respective commander (i.e., the person who deployed that autonomous agent). Interaction with other agents changes the goal evaluation model by either changing the allocation of goals among the agents or by changing the expected values associated with the con-

stituent goals. Agents can perform "what-if" evaluations for possible new goal allocations through modifications of the desire-space model.

Once the UAV agents decide which targets to track based on the allocated individual goals and send the sensed location information of the corresponding targets to the Central Command, Central Command must determine which UAVs are the "best" target information providers for generating the unified situational picture, as demonstrated in Section 2.2.

2.2   Selecting Target Information Providers

When multiple UAVs can provide location information about the same target, selecting the appropriate information providing UAVs (information sources) requires the Central Command to decide "from whom to request what", and involves two tasks which are search space construction and search. A search space is constructed to represent the potential information providers accompanied with the evaluation of the information provider set. The challenge comes from the fact that the number of possible information sources increases exponentially as the diversity of information requirements and the number of sources meeting those requirements increase. A proposed heuristic search called *hill-climbing with mutation operation* explores the search space efficiently and finds the best solution while minimizing the possibility of local optima [3].

The search space consists of a set of nodes representing potential information source combinations. Each instance of an information source combination constitutes a node in the search space [4]. The valuation of each node represents how good an information source combination is. Three factors are considered to valuate each node – trustworthiness, coverage, and cost [4]. Trustworthiness of an information source is represented by the probability distribution of the error, from the estimate true value, of the provided information [1]. The coverage measure represents the amount of contribution the information sources make to an agent's goal achievement. Information cost is derived from the message-passing and computational burden required to communicate information. The nodes are connected with each other where one node can be mapped into another node by adding or removing a single source of a single requirement, so that the adjacent nodes are minimally different. The currently selected information sources are intended to be the best sources combination at the current time, and the information sources are replaced by new sources when the new sources are proved to (1) include better partners, (2) exclude bad partners from the current partners or (3) be an almost or completely new set of information sources having a higher evaluation value than the current ones do. In the case of (1) and (2), it is likely that most of the current partners are still included. Therefore, it is desirable to keep the search path near the current node because it is more likely that a better node exists nearby. By harnessing this locality, search can be efficiently performed. However, depending only on the locality can reach a local optimum or can slow down the search in the case of (3), so we need to expand the exploration of the search space to reduce the possibility of the local optimum. Expanding the exploration in turn requires more computation and may waste resources if the exploration does not return a better result.

The *hill-climbing with mutation operation* concurrently makes use of the locality by adopting the hill climbing search method while expanding the exploration by

adopting a mutation operation from genetic algorithms [3]. Hill-climbing is a heuristic search to head towards a neighboring state which is better than the current state, guaranteeing a result which is equal or better than the current state. The mutation operation reduces the possibility of the local optimum problem. Mutation generates a new node to be inspected by applying a simple modification rule to the current node, and it enables random walks in the search to escape from the local optimum.

Once Central Command has decided from which UAVs to request which information, it must valuate the incoming information to build an accurate estimate of target location, as demonstrated in Section 2.3.

2.3   Valuation of Target Location Information

After Central Command has identified potential tracking UAVs for a given target, as explained in Section 2.2, it must select the best location information from those UAVs; some information may not be reliable due to sensor inaccuracies or noise. By adopting a set of policies for choosing information, Central Command can attempt to minimize the risks associated with dependencies on information-providing UAVs.

By using a set of information valuation policies to the select information Central Command uses to build its situational picture, target location estimate accuracy can be ensured. Since Central Command does not know true target locations, it has no basis to judge the quality of information based on a UAV's communicated values alone. However, it can follow general policies [1] for identifying the most accurate information. An algorithm employing a compromise of the following policies can identify robustly the most valuable information for good decision-making.

1. **Priority of Maximum Information:** incorporate information from as many sources as possible. Given information from a greater number of sources, the calculated location estimate should be closer to the trust target location if UAVs are statistically independent in their reporting.
2. **Priority to Corroborated Information:** give priority to information that can be corroborated.
3. **Priority for Source Certainty:** give priority to information from UAVs conveying high certainty on that information if the UAV is proficient at conveying a quality certainty assessment.
4. **Priority to Reliable Sources:** give priority to information from UAVs estimated to be most reliable, based on past experience or recommendations.
5. **Priority to Recent Information:** give priority to data recently received.

These policies ensure secure decision-making despite uncertainty by identifying three types of error in reported information: 1) error due to information age (the value being estimated may have changed since the information was received), 2) error due to UAV unreliability (perhaps due to failed sensors), and 3) error due to a UAV's uncertainty in its information (a UAV may be uncertain about the quality of its sensors).

3   Application in UAV Target Tracking

The technologies described in Section 2 have been implemented in a simulation of UAV target tracking for experimentation and demonstration. In this simulation, tar-

gets are located at points in a 2-dimensional surveillance field and agents, controlling UAVs, track target locations. Fig 2 shows the user interface to demonstrate action selection. As a simplification, each UAV is represented by a dot surrounded by a circle indicating the UAV's sensor radius. Targets are depicted by small dots surrounded by shaded circles, whose size and darkness represent the time elapsed since the target was within range of, and thus sensed by, any UAV. The scenario assumes a Central Command receives target location information from UAV agents and forms estimates about target locations. UAV agents are modeled by Central Command to have varying reliability depending on sensor quality (the UAVs are not necessarily homogenous). UAV agents obtain target location information from their sensors at varying intervals depending on their flight paths.
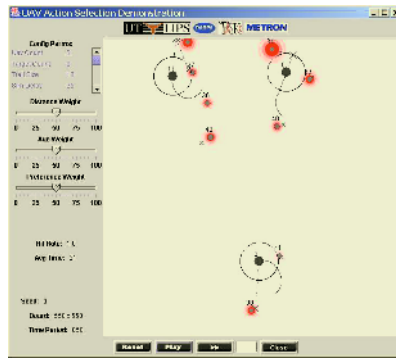
Fig. 2. Simulation Demonstrating Autonomous UAV Target Selection

Lines out of the UAV indicate the path taken and intended future destinations. An agent has control over the heading and speed of a single UAV. Each agent is charged with selecting which target their UAV will service next with the system goal of keeping models of target positions updated. Thus, the agents' decision-making is guided by two principles: (1) the desire to service a target increases as time passes (affecting the reward values), and (2) the desire to service a target decreases as distance to that target increases (modeled in the cost function). The first principle attempts to ensure fairness, so that all targets will eventually be serviced. The second principle increases the efficiency of servicing, encouraging agents to service nearby targets first, reducing travel, and leading to a partitioning of the targets amongst UAVs.

Uncertainty is introduced by target movement, forcing the UAVs to operate with possibly outdated location information. To ameliorate the effects of uncertainty, trust and information sharing create an up to date, unified situational picture. UAV movement also adds uncertainty to modeling the expected reward value received from visiting a target since the agents are, in a sense, competing for targets. Coordination among agents reduces uncertainty and improves system performance.

Fig 3 is the Topology Viewer showing which information sources are selected by the Central Command. The Central Command is depicted by a dot surrounded by a circle, connected by lines to selected information sources. Among the three evaluation criteria (trustworthiness, coverage, cost), the trustworthiness evaluation value is represented by different colors. This viewer demonstrates the dynamic information source selection of the Central Command while UAVs and targets are moving around.
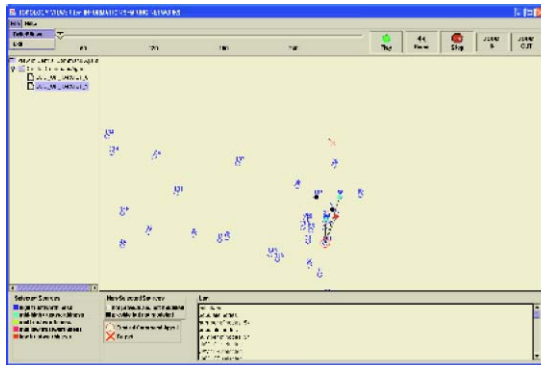
Fig. 3. Topology Viewer demonstrating UAV selection as location providers
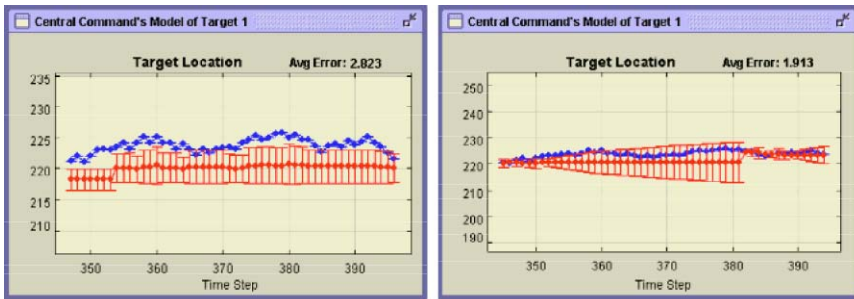
Fig. 4. Target location estimation using (a) Naïve Algorithm and (b) All-Policy Algorithm

Fig 4(a) and (b) demonstrate the improvement in location estimate accuracy with information valuation policies. An algorithm based on the five policies of Section 2.3 (All-Policy Algorithm) is compared against an algorithm that simply averages a target's location information from all UAV agents (Naïve Algorithm). Fig 4 shows Central Command's target location estimates as computed using the Naïve Algorithm, averaging location information from all UAV agents (Fig 4(a)) and using the All-Policy algorithm (Fig 4(b)). Dots with error bars represent Central Command's estimate of target location and certainty. For comparison purposes, lone dots represent the target's actual location. First, note that the All-Policy algorithm achieves lower error in its target location estimates, shown both on the chart and by the computed average error. Second, when Central Command is receiving no UAV Agent reports (when the location estimate remains static), the All-Policy algorithm adjusts its error bars to accommodate the known decrease in target location certainty.

4   Conclusion

This paper illustrates three agent technologies deployed in the UAV target tracking domain for (1) coordinating target tracking among multiple UAV agents, (2) identifying the best UAVs to be used for location information collection, and (3) evaluating the accuracy of location information. Using available information, a UAV agent tries to make the best (rational) decisions it can to affect the domain according to its de-

sires. Using macro actions to factor the state space and estimate the cost of pursuing each goal, reasoning is performed in the "desire space," which describes the expected value of pursuing a goal in the context of how selected actions facilitate the pursuit of other goals in the future. To facilitate the target tracking activity, reliable information channels are constructed by finding the most appropriate UAVs from which to acquire necessary target location information. Finally, a policy-based scheme for assessing information accuracy overcomes the potential risks of uncertain location information.

Acknowledgment

References

1. Fullam, K. and Barber, K. S.: Using Policies for Information Valuation to Justify Beliefs. In Proceedings of the Third International Conference on Autonomous and Multi Agent Systems (AAMAS 2004), New York City, New York (2004) 404-411.
2. Han, D. C. and Barber, K. S.: Desire-space Analysis and Action Selection for Multiple, Dynamic Goals. In Proceedings of the 5th Workshop on Computational Logic in Multi-Agent Systems (CLIMA V), Lisbon, Portugal (2004)
3. Park, J. and Barber, K. S. Building Information Sharing Networks with Local Partnership Decision and Cooperation Motivation, Technical Report. TR2004-UT-LIPS-034, The University of Texas at Austin, Austin, TX, (2005).
4. Park, J. and Barber, K. S.: Information Quality Assurance by Lazy Exploration of Information Sources Combination Space in Open Multi-Agent Systems. Journal of Universal Computer Science, 11 (2005) 193-209.
5. Sutton, R. S., Precup, D., and Singh, S. P.: Between MDPs and Semi-MDPs: A Framework for Temporal Abstraction in Reinforcement Learning. Artificial Intelligence, 112, 1-2 (1999) 181-211.

Cognitive Hybrid Reasoning
Intelligent Agent System

Christos Sioutis and Nikhil Ichalkaranje

School of Electrical and Information Engineering
Knowledge Based Intelligent Engineering Systems Centre
University of South Australia, Mawson Lakes SA 5095, Australia
{Christos.Sioutis,Nikhil.Ichalkaranje}@unisa.edu.au
http://www.kes.unisa.edu.au

Abstract. This paper presents a novel approach for implementing reasoning and learning systems based on Intelligent Agents. It is particularly suited for agent systems working in complex and cooperative dynamic environments. This approach was developed to overcome limitations confronted while researching the use of agents in such environments. A brief background of the underlying conceptual reasoning model is described, followed by details of the implemented framework and concluded with some final remarks and future work.

1 Introduction

Recent research directions [7, 9] for enhancing the performance of agent based systems have lead to the investigation of human reasoning models as a main paradigm for implementing intelligent agents. The widely used Beliefs Desires and Intentions (BDI) model was developed by Bratman [2] to describe how humans perform practical reasoning. Practical reasoning is defined as considering actions that need to be performed given a particular situation [10]. An important assumption about the BDI model is that a BDI agent's actions must be logically consistent with the combination of its beliefs and goals. The BDI reasoning process is believed to consist of at least two distinct activities, the first activity being deliberation, which involves deciding what goals to achieve [10], and the second activity, means-ends reasoning, involves deciding how to achieve this goal [10].

JACK is a commercial platform for developing intelligent agents based on the BDI model that uses events and plans to define different behaviours that an agent can exhibit [1]. A behaviour is executed by posting an event which is handled by a plan. The plan contains any relevant code that causes the agent to exhibit the desired behaviour. A plan is also able to post other events and the JACK language provides a number of ways in which events can be posted and handled. This makes it a powerful platform for building complex agent behaviours. However, one limitation that was observed is that all plans must be pre-defined to be used. There is no way in the current distribution (version 5.0) to modify a behaviour, short of modifying the plan code (or the way the plan is

selected) and re-compiling it off-line. This limitation is specifically addressed by this research.

Section 2 presents a brief overview of a conceptual reasoning and learning model previously developed and ties our Cognitive Hybrid Reasoning Intelligent Agent System. Our case study related to a learning agent is presented in section 3. The final section presents concluding remarks and future directions.

2 Cognitive Hybrid Reasoning Intelligent Agent System

Rasmussen's decision ladder [5], Boyd's OODA loop [4] and Bratman's BDI are three popular models for describing human decision making. Recent research has also revealed that these models are complementary and can in-fact be fused together to yield a new, hybrid and more detailed conceptual reasoning model that supports learning [6]. The Cognitive Hybrid Reasoning Intelligent Agent System (*CHRIS*) is an implementation of this model that was developed as an extension to JACK. It provides a complete framework for designing agents based on the reasoning model. The framework however is not directly executable. A specific application must be designed such that it uses the framework and implements any required additional components. The only reason for this, is that any non-implemented components are inherently application specific, in-fact these components 'hook' the framework code with code from the specific application into one complete working system.

The framework divides the Agent's reasoning process into six stages. Each of the stages have been implemented in different Java packages which are required to be imported for successful compilation. Two additional packages are also required as listed below.

1. Agent: Provides an extended JACK learning agent, the CHRIS capability and associated Java interface classes.
2. Observation: Receives sensations and converts raw data into useable information.
3. Orientation: Uses new information to update the Agent's beliefs accordingly.
4. Decision: Manages goals that the agent is trying to achieve.
5. Action: Executes tasks that cause the agent take actions in the environment.
6. Learning: Modifies the actions taken by the agent according to specified learning goals, active and passive learning methods are available.
7. Logging: Controls the logging functionality built into the framework.
8. Util: Utility classes that were created while developing the framework.

Active learning is realised by traditional Reinforcement Learning (RL) algorithms. It involves handling a event that has an attached learning goal. The agent then activates the learning algorithm specified in the learning goal and executes different actions while observing the results of the actions. The learning is achieved by the algorithm finding an optimal state-action value function based on rewards generated by its learning goal. That is, a function that translates a

particular state to a specific action that is to be taken in that state. The optimal value function causes the agent to make choices that give it maximum rewards.

Passive learning is a feature that has been implemented specifically in the *CHRIS* framework but was inspired by the 'exploration control module' presented by Dixon [3]. It allows an RL action-selection policy to be replaced by a JACK plan written in a specific way. This allows an agent to learn the behaviour of a pre-written JACK plan, and evaluate the JACK plan with respect to the given learning goal. Skill control is another feature that allows the agent to switch from passive learning to active learning. When enabled, skill control uses specified thresholds for switching between the two learning methods.

The interested readers are invited to contact the first author for a detailed description of *CHRIS*.

3   Example Learning Agent

A simple example system was implemented in order to test the features provided by the framework, it involves an agent that learns how to play the game of Tic Tac Toe. The actions available to the agent are variable and are based upon the moves available to the agent for a particular state. The Sarsa [8] RL algorithm and the EGreedy [8] policy were selected. The learning goal's reward function is based on the action that the agent chooses, it returns a -20 if the action results in the agent losing the game, a +20 if the action results in the agent winning the game and -1 if the action results in a stalemate or if the game is not finished. This instructs the agent to avoid losing and try to win in the quickest possible way. The required components for both active and passive learning were implemented and four runs of 2500 plays were made for each type of learning. The learned data was preserved after each run for that particular learning type to demonstrate that the agent is able to save, load and improve upon the value function over multiple executions.

Figure 1(a) shows the performance of the agent for active learning. The agent uses an exploration rate of 50% for the first run of 2500 plays. This is indicated by the very low performance of the agent during the first 2500 plays where it loses approximately 75%, draws approximately 25% and wins less than 1% of the games. The exploration rate is then set to 10% for the subsequent runs and the performance suddenly changes to approx 70% drawing, 15% winning and 15% losing. The small spikes in the lines every 2500 correspond to starting a new run.

Figure 1(b) shows the performance of the agent for passive learning. During the first 2500 runs the agent's skill control is disabled. Which means that it learns from its actions but it acts based on a regular JACK plan that contains a behaviour for playing Tic Tac Toe, the learning package has no effect on the actions taken by the agent which performs at 65% draws, 35% losses and 5% wins. For subsequent runs the skill control is enabled which causes the agent to switch back and forth between active and passive learning depending on its performance. Again a substantial change in performance is achieved, draws remain
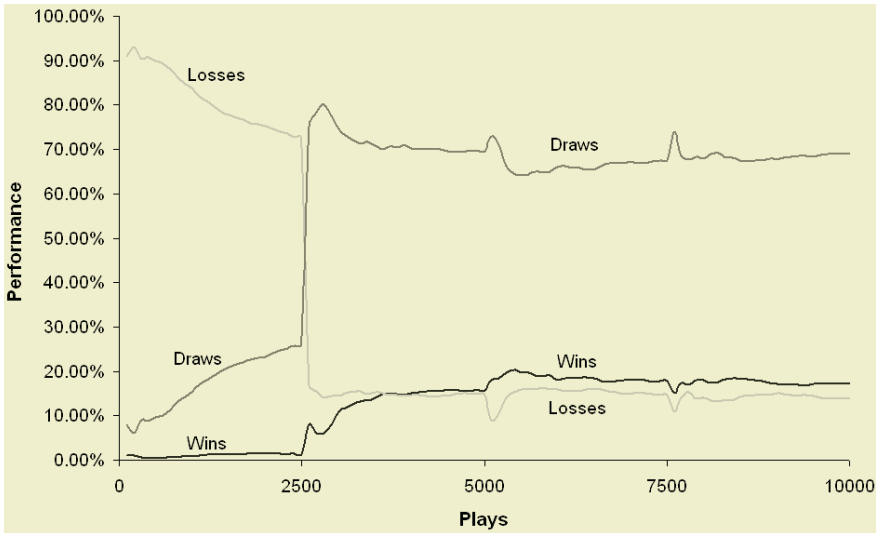
(a) Active Learning
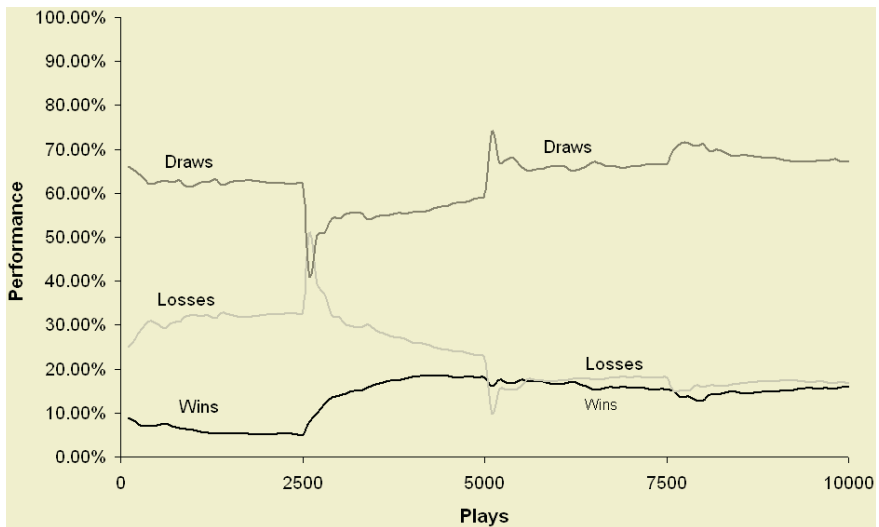
(b) Passive Learning

Fig. 1. Learning Agent performance when playing Tic Tac Toe

at approximately 65% while both wins and losses stabilise at approximately 18%. This example demonstrates an advantage of passive learning. It performs well at the start when learning is inexperienced, but it also improves its performance later on when learning is so experienced that it provides a better behaviour than the pre-written JACK plan. The behaviour of the agent can be even further improved through tweaking its reward function or improving a state generalisation function defined within its learning goal.

4   Conclusions and Future Work

This paper describes a framework that is based upon a hybrid, conceptual reasoning model. The framework allows the creation of JACK agents with the ability to learn from their experiences. Learning is achieved through RL algorithms that execute different actions while observing the results of the actions until an optimal state-action value function is reached for a given learning goal. The optimal value function causes the agent to make choices that give it maximum rewards.

Future work for *CHRIS* includes implementing different types of learning agents and assessing their performance. Some additional RL algorithms (eg. Eligibility Traces) may also be implemented in order to improve learning performance. A Planning capability based on [8] is also intended to be integrated into *CHRIS*. The planning capability would allow the agent to learn a Model of the environment. The Model would be used to 'mimic' the environment in a limited way allowing the agent to do off-line learning using simulated experiences against the model.

Acknowledgements

References

1. Agent Oriented Software, *Jack intelligent agents agent manual*, [Online, accessed April 2005], (2005), URL:http://www.agent-software.com.au/shared/resources/index.html.
2. Bratman ME, *Intention, Plans, and Practical Reason*, Center for the Study of Language and Information, (1999).
3. Dixon KR, Malak RJ, and Khosla PK, *Incorporating Prior Knowledge and Previously Learned Information into Reinforcement Learning Agents*, Technical Report, Institute for Complex Engineered Systems, Carnegie Mellon University, (2000) pp13.
4. Hammond GT, *The Mind of War: John Boyd and American Security*, Smithsonian Institution Press, Washington, USA, (2004).
5. Rasmussen J, Pejtersen AM, Goodstein LP, *Cognitive Systems Engineering*, Wiley & Sons, New York, NY, (1994).
6. Sioutis C, Ichalkaranje N, Jain LC, Urlings P, and Tweedale J, *A conceptual reasoning and learning model for intelligent agents*, In proceedings of the 2nd International Conference on Artificial Intelligence in Science and Technology (AISAT 2004), University of Tasmania, Hobart, Australia, November, (2004) 301-307.

7. Sioutis C, Tweedale J, Urlings P, Ichalkaranje N and Jain LC 2004, *Teaming Humans and Agents in a Simulated World*, In proceedings of the 8th International Conference on Kowledge Based Intelligent Information and Engineering Systems (KES 2004), Springer Verlag, Berlin, pp. 80-86.
8. Sutton RS and Barto AG, *Reinforcement Learning An Introduction*, The MIT Press, London (2000).
9. Urlings P., *Teaming Human and Machine*, PhD Thesis, University of South Australia (2004).
10. Wooldridge M, *Reasoning About Rational Agents*. The MIT Press, London, (2000).

Beyond Trust: A Belief-Desire-Intention Model of Confidence in an Agent's Intentions

Bevan Jarvis[1], Dan Corbett[2], and Lakhmi C. Jain[1]

[1] School of Electrical and Information Engineering, University of South Australia
Bevan.Jarvis@postgrads.unisa.edu, Lakhmi.Jain@unisa.edu.au
[2] Science Applications International Corporation
corbettd@saic.com

Abstract. Trust plays a fundamental role in multi-agent systems in which tasks are delegated or agents must rely on others to perform actions that they themselves cannot do. The concept of trust may be generalised and considered as a level of confidence in one's predictions of another agent's future behaviour. This has applicability beyond that normally ascribed to trust: for instance, one may be confident that a particular agent's intentions are hostile, and that this will be borne out by particular behaviours. In this paper we present a cognitive model of trust in which the central component is a Belief-Desire-Intention model or 'theory of mind' of a person or agent that evolves over time.

1 Introduction

In this paper we propose, as a concept for discussion, a generalisation of trust that appears to follow naturally from the cognitive approach. We regard trust as an agent's confidence in its belief in the 'theory of mind' that it has of another. We call this theory a 'model of trust'. In consequence an agent can be said to be better trusted if it behaves more predictably, whether or not the behaviour is beneficial to the trusting agent. An untrustworthy agent is one that is not predictable.

Delegation can be regarded as the incorporation of an action of another agent into a plan. In the generalised model delegation becomes an acceptance that the second agent will act in a certain way. In the case that the expected action is not considered beneficial, it would serve to constrain planning, by filtering out plans that are inconsistent with the expected action. In our conception, trust and acceptance can still serve to reduce the amount of time spent deliberating over which plan to adopt.

The application of interest is that of human-agent teams. Trust is of concern in the formation and maintenance of any team, but with human-agent teams it is feasible to analyse the intentions of other agents, either human or artificial. The primary concern then becomes developing a representation of our model of trust that is understandable both to the human and software agents in the team.

Our interest is in individual-level trust. However, the generalisation we propose is relevant also with regard to system-level trust, or more precisely, an individual's trust in the system. The system in which an agent operates is not usually overtly helpful to the agent, and can be expected to prevent it from doing certain things. The agent is entitled to expect, however, that the system be consistent, and to some extent predictable. The agent should be able to make plans under the assumption that the system will behave as specified.

We now discuss research context. Ramchurn, Huynh and Jennings [10] survey the state of the art of the subject of trust in multi-agent systems. They distinguish between system-level trust and individual-level trust. The former they classify into trustworthy interaction mechanisms (protocols), reputation mechanisms and distributed security mechanisms. Individual-level trust they divide into socio-cognitive models, reputation models, and evolutionary and learning models.

Castelfranchi and Falcone [5, 6] put the case for the cognitive model of trust. They assert that: (1) only a cognitive agent (one with goals and beliefs) can trust another agent, (2) trust is a complex mental attitude of one agent about another with regard to behaviours or actions relevant to particular goals, and (3) trust is the mental counterpart to delegation.

Marsella, Pynadath and Read [9] describe PsychSim, a multi-agent system for simulating social interactions that incorporates theory of mind. Its basis lies in decision theory, rather than in the Belief-Desire-Intention model [3].

The layout of the remainder of this paper is as follows. In section 2 we present motivating examples illustrating different aspects of the problem. Section 3 develops the model of trust in more detail and emphasises the role of commitment. Section 4 gives a categorisation of strategies of commitment. In section 5 we discuss a possible way of implementing our model as an extension of JACK Teams [2]. Finally we summarise and discuss future research.

2   Motivating Examples

2.1   Scenario 1

John, who works in a bar, approaches his employer, Sarah, asking that he be considered for the position of bar manager; this would mean more money but requires greater responsibility. Sarah explains that in his performance to date he has not yet demonstrated all the characteristics she requires of someone in that position. She explains her requirements: she must be confident he has a high level of knowledge and skills; that his goals are consistent with the success of the bar; that he will act appropriately in any given situation; and that he has a high level of commitment to the success of the bar and an appropriate level of commitment to any task that he undertakes. Her intention is that whoever takes on the position will fulfil it in the way she would want it to be done, but without her having to closely supervise the work. Together, they devise a plan by which John can convince Sarah that she can trust him in the role of bar manager. John's commitment to this plan is part of the test.

2.2   Scenario 2

Corporal Jackson is monitoring airborne surveillance of shipping movements in the area north-west of Australia. He is watching a boat that is outside of normal shipping lanes. From its speed he knows it is powered, but apparently not fast, and he deduces from its course that the captain is unaware of a cyclone 100 km to the east currently moving towards the craft. Over time, he becomes more certain that it is not a fishing boat and that the captain's intention is to land on the Australian coast. Based on his confidence in that intention, he requests a vessel to intercept the suspect craft, provid-

ing a predicted course, which includes direction changes Corporal Jackson expects the captain to make when he becomes aware of the approaching cyclone.

2.3   Scenario 3

Alicia is a highly skilled tennis player. When she plays Mairi for the first time, she will yet have some expectations. Primary among these will be that Mairi will attempt to win, and will, to some extent, be committed to that attempt. Over the course of the match, Alicia will develop a notion of Mairi's knowledge of tennis, her ability, the tactics she uses, and her level of commitment to the game. This estimation will evolve over time, as the two meet in future tournaments. (Commitment – more precisely, courage and resolution, or 'heart' – is ascribed a particular importance in sport, and is often cited as what enables a champion player to come from behind and win.)

2.4   Scenario 4

Parker is an experienced criminal planning a heist. After much thought, and some telephone calls, he has gathered together a small group to execute the plan he has conceived. Parker has worked with some of the team before, and knows how they might act. Others he knows only by reputation, and one or two are unknown to him (but are necessary because they know the 'score'). In common, they have (Parker believes) an intention to carry out his plan, and some level of commitment to that intention. As part of his deliberation, Parker has assigned a role for each member of the team, and to some extent the roles are designed to suit the person – their knowledge, abilities and desires. Parker himself has a role, suited to his own beliefs about himself. A nagging doubt persists for him about the intentions of certain members of the group, and their commitment to his plan. This doubt leads him to remain watchful and observant of their behaviour. Once he understands their motivations and intentions, he can make any necessary contingency plans.

2.5   Discussion

The common theme of these scenarios is the prediction of the future behaviour of another person, to the extent of believing that it will stay within certain bounds. This implies a theory of mind: a model of their knowledge, abilities and motivations.

Of interest in a computing context is the Belief-Desire-Intention model [3], the inspiration for implemented BDI agent architectures. This model uses the notions of Belief, Desire, Intention and Plan to describe rational behaviour of humans. Intention is defined as a desire to which a person has commitment. In implemented BDI systems, this means an objective that an agent is part-way through executing a plan to achieve.

Desires may not be known, but can sometimes be assumed or inferred. Thus, Corporal Jackson believes the captain has a general desire to avoid cyclones. Desires are usually only hypothetical indicators of future action, however, and are not as useful to know about as intentions.

In these scenarios, commitment plays a central role. A high level of commitment enhances predictability of future action, and therefore trust. Different levels and

strategies of commitment apply to different tasks and goals. John's intention to put out the rubbish bins on Tuesday night does not acquire the same kind of commitment as his intention to keep the bar clean. Commitment can usefully attach to things that are not actions – for example, norms and obligations: telling the truth or not stealing from the till. This differs from Bratman's notion of commitment as an aspect of intention, which in turn implies performing a plan. We regard commitment as a separate mental attitude, additional to desire and intention – as it is regarded, for example, in [4].

Plans are not so important in the scenarios. Generally speaking, it is not necessary to know or to specify how a person will do something: it often suffices simply to know they are able to do it. Rather than plans, we speak mainly about abilities.

3   Towards a BDI Model of Trust

3.1   The Model of Trust

We consider trust as a set of beliefs about another person or agent. This means beliefs about knowledge (beliefs), abilities, desires and commitments. We disregard intention in favour of commitment, which (according to Bratman) converts desire into intention. Other modalities, such as norms and obligations, can also be modified by commitment. We regard commitment as being more fundamental.

The purpose of the model of trust is to predict behaviour to the extent of determining whether it will stay within certain bounds.

3.2   Building a Model of Trust

Building a model of trust can be a joint exercise. The proactive version of the process is described in Scenario 1. Sarah lets John know what is required, and John demonstrates behaviour that leads her to believe that he has the characteristics she requires. Each has some level of commitment to the process of building trust, and each may reconsider their intention to continue. In an important sense, however Sarah is in control, and might be expected to withhold certain aspects of her model.

The process can also be one-sided, as in scenario 2, although it can be undermined. The captain may be deliberately travelling slowly to disguise the vessel's actual speed, and may in fact know about the cyclone. The captain may be said to be participating in the building of trust, but in a way that is at cross-purposes with the corporal.

The process is non-monotonic – the model of trust should be reconsidered with new information. This may be subject to non-reconsideration. If John uncharacteristically arrives late one day without a good excuse, Sarah may yet decide to overlook the incident, perhaps simply because she is in a good mood.

If the model for a person is negatively altered then any delegations of authority should in principle be reconsidered. There is room here too for non-reconsideration. Having gone through a long process of building trust and finally giving John his promotion, Sarah might be expected to have some commitment to keeping him on.

There is also the possibility that a person's level of trust is such that they distrust the 'facts' rather than the person. The possibility of such mistrust, and the possibility of non-reconsideration, leads to 'stickiness' of the model of trust.

3.3 Fitting the Model: Roles and Stereotypes

In the positive-trust situation, the goal is to delegate an action, task or role. It makes sense to specify a model of trust desired for the delegatee. This model is essentially a job specification, and should be compared to the model developed for the people being considered. We refer to the specification as a 'role definition'.

More generally, we refer instead to 'stereotypes'. Corporal Jackson has (among others) stereotypical models for 'fisherman' and for 'people smuggler'. He compares his deduced model of trust of the captain to each of these stereotypes. Over time, he becomes more certain about which of the stereotypes best fits the situation.

4 Strategies of Commitment

The following describes a characterisation of strategies of commitment that is detailed in [8]. This description builds upon work by Rao and Georgeff and by Wooldridge, but ultimately relies on Bratman's original conception.

1. Null commitment means either single-shot commitment or delegated commitment. John may make a single attempt to clean a stain off a glass; if he does not succeed he throws the glass away. Sarah asks John to put the rubbish out, and then forgets about the request, trusting him to do as he was asked.
2. Blind commitment means the intention is maintained until the agent believes it has been achieved. John sets to the task of washing glasses, and continues until all the glasses that have been stacked up are washed.
3. Single-minded commitment means the intention is maintained until it is believed either that it has been achieved or that it is impossible. Mairi continues to try to win against Alicia until either she has succeeded or Alicia has won.
4. Open-minded commitment means the intention is maintained until it is believed impossible. Corporal Jackson maintains his duties of surveillance so long as information continues to arrive.

Each of these strategies can be subject to reconsideration. John may discontinue washing the glasses to help Kylie with a difficult customer.

Finally, there is the issue of non-reconsideration: an intention due for reconsideration may not be reconsidered. A typical reason for this is lack of time.

5 Implementing the Model of Trust: Extending JACK Teams

JACK [1] is an implementation of BDI. JACK Teams [2, 7] is an extension of JACK, and of the BDI model, that provides a team-oriented modelling framework. A team in JACK Teams is also an agent, and inherits all the characteristics of an agent in JACK.

The purpose of JACK Teams is to model the process of delegation. It achieves this through what are called roles. Teams require roles, which relevant team members (role fillers) must fulfil. A role consists of a set of tasks, for each of which a role filler must have a relevant plan, as well as 'teamdata' constructs, which allow beliefs (data) to be exchanged between team and team member. The details of the plan need not be specified – only the fact of its existence is required. The concept of role thus appears to be quite similar to that of ability, as we have defined it.

We consider extending the concept of role in JACK Teams to incorporate our model of trust.

5.1   Implementing Positive Trust

First we consider the 'positive' situation, where trust leads to delegation. Ability is already incorporated in JACK Teams, matching the concept of role. As with ability, required knowledge is not specified in detail. The role filler is expected to have, or to be able to obtain, the knowledge it requires. In effect, knowledge is subsumed into ability. With teamdata it is possible to keep track of what a team member believes (or what it says it believes). It is thus possible to define a requirement for knowledge in terms of the ability to respond to a query.

Of more interest is introducing the notion of commitment, either to the role or to specific events that the role handles. The team cannot simply require the team member to state its level of commitment and leave it at that. It must have a means of checking each team member's level of commitment, or else the specification itself would become pointless.

The latter principle also applies to ability and to knowledge, but we emphasise the importance of commitment. Lack of knowledge, ability or commitment all produce the same effect: consistent failure at a particular task. Conversely, it is hard to think of a situation where too much knowledge or ability would be a problem. Too much commitment to a task, however, can easily be seen to cause problems, even assuming the agent is not malevolent. Over-commitment to a task can cause other tasks to fail.

5.2   Implementing Generalised Trust

In the general situation, the requirement is to maintain a model of trust for an agent, and occasionally to compare this model to various stereotypes. We propose a strategy for implementing the model.

The model of trust should be maintained as a team. This means using the team construct as a storage mechanism for information about the observed agent. The model-of-trust team should incorporate the surmised commitments the observed agent has to its intentions (roles).

Secondly, the stereotypes should be specified as roles. As discussed above, these roles should include specifications of level of commitment.

Finally, the model should be compared to the stereotype. The aim is that the model-of-trust team should 'tender' for the stereotype role. A successful tender means that the model fits the stereotype.

5.3   Completing the Model

The generalised situation should encompass the positive. A team should have a model-of-trust team for any agent (sub-team) in which it is interested. The model-of-trust team is subject to reconfiguration whenever more is learned about the sub-team, either through supervision or more indirect observation.

6   Summary and Future Research

In this paper we have generalised the cognitive model of trust, regarding it as confidence in a theory of mind of another agent, which we call a model of trust. We have

emphasised the idea of commitment as a contributor to this trust. Commitment we have defined to incorporate intention, norms and obligations. We have briefly explored how to implement this in JACK Teams.

The model we have presented falls squarely into the classification of individual-level trust given in [10]. While it is a cognitive model, it is also capable of evolving ('learning') over time. It is also possible to use the model to provide reputation evaluations, and indeed to evaluate other agents in their 'role' of reputation provider.

Future research lies in looking more closely at the implementation in JACK Teams, leading to modelling the mechanism described. Initially, the focus will be on implementing the positive model of trust, as this is closest to the purpose and mechanism of JACK Teams itself. Achieving this may give a clearer insight into the means of implementing generalised trust.

References

1. AOS: JACK Intelligent Agents: JACK Manual, Release 4.1, Agent Oriented Software, Pty Ltd, 2004.
2. AOS: JACK Intelligent Agents: JACK Teams Manual, Release 4.1, Agent Oriented Software, Pty Ltd, 2004.
3. Bratman, M. E.: Intention, Plans, and Practical Reason, Harvard University Press, Cambridge, MA, 1987.
4. Brazier, F. M. T., Dunin-Keplicz, B., Treur, J. & Verbrugge, R.: 'Modelling Internal Dynamic Behaviour of BDI Agents', in Formal Models of Agents: ESPRIT Project ModelAge Final Workshop: Selected Papers, LNCS 1760, Springer-Verlag, 1999, pp. 36-56.
5. Castelfranchi, C. & Falcone, R.: 'Trust is Much More than Subjective Probability: Mental Components and Sources of Trust', in Proceedings of the 33rd Hawaii International Conference on System Sciences, IEEE, 2000.
6. Castelfranchi, C. & Falcone, R.: 'Social Trust: A Cognitive Approach', in Trust and Deception in Virtual Societies, Kluwer Academic Publications, 2001, pp. 55-90.
7. Evertsz, R., Fletcher, M., Jones, R., Jarvis, J., Brusey, J. & Dance, S.: 'Implementing Industrial Multi-agent Systems Using JACK', in First International Workshop on Programming Multiagent Systems (PROMAS 2003), vol. LNAI 3067, Springer-Verlag, Melbourne, Australia, 2004.
8. Jarvis, B.: 'Expressing Commitment in a BDI Architecture', in review.
9. Marsella, S. C., Pynadath, D. V. & Read, S. J.: 'PsychSim: Agent-based modeling of social interactions and influence', paper presented at the Sixth International Conference on Cognitive Modeling, 2004.
10. Ramchurn, S., Huynh, D. & Jennings, N. R.: 'Trust in Multi-Agent Systems', Knowledge Engineering Review, 2004.

Reasoning About Time in a BDI Architecture

Bevan Jarvis[1], Dan Corbett[2], and Lakhmi C. Jain[1]

[1] School of Electrical and Information Engineering, University of South Australia
`Bevan.Jarvis@postgrads.unisa.edu, Lakhmi.Jain@unisa.edu.au`
[2] Science Applications International Corporation
`corbettd@saic.com`

Abstract. The philosophical roots of the Belief-Desire-Intention model lie in Bratman's formulation of an intention theory of planning, in which he sought to make sense of the notion of future-directed intention. Implementations of BDI mainly follow the original Procedural Reasoning System model. BDI has a sound logical basis, exemplified by the Logic Of Rational Agents. While the LORA formulation has a temporal logic component, however, this does not translate into any ability for the agent to reason about actual time. Being able to reason about actual time would bring significant benefits for BDI agents, such as the ability for agents to communicate deadlines and to plan and schedule activities in a cooperating group. Given a suitable representation of temporal knowledge, an agent could learn about the temporal aspects of its own actions and processes, and this knowledge could be used as input to the planning process. This paper outlines a possible implementation strategy for the representation of, and the capacity to reason about, actual time.

1 Introduction

The ability to reason about actual time means firstly that an agent should be able to predict how long its own actions will take, and secondly that it should be able to plan and schedule activities, for itself or for a team. The first requirement entails some self-knowledge of the agent's own programming; the latter is the realm of the well-researched problem of planning. In this paper we propose, as a concept for discussion, a scheme for implementing the ability to reason about time in a BDI agent.

BDI is an architecture for building intelligent agents that balance the time spent on deliberation and planning, exhibiting both reactive and goal-directed behaviour in real time [13].

LORA represents a formal model of a single BDI agent, and provides the capacity to reason about inter-agent communication and cooperation [24]. Communication is achieved through speech acts. The internal workings of an agent are invisible to other agents. In particular the internal clock of a BDI agent has no global relevance.

Two kinds of time are at issue. The internal time within LORA is modelled time. The nodes in LORA's branching-time temporal logic represent agent states, not times. The links between nodes are labelled with actions, and the labelled tree becomes, essentially, a dynamic logic program specification.

Actual time is neither a part of the BDI architecture nor of the LORA model. This is properly so: actual time may be viewed as a perceived aspect of the environment.

In order to facilitate reasoning about actual time, and make full use of this function, agents should be given three distinct capabilities:

1. Knowledge about the time they take to execute their own plans;
2. The ability to compute estimated or minimum and maximum plan execution times;
3. The ability to plan and schedule activities (either for a single agent or a team).

The first of these might be subject to evolution over time, as plans are executed and execution times observed. An agent might thus be said to learn about the temporal nature of its plans.

The research that has led to this paper, towards extending BDI with temporal logics, is being conducted under the auspices of the KES group of the University of South Australia. The research interests of the group include BDI and in particular a commercial implementation of BDI known as JACK Intelligent Agents, or JACK [14, 3]. In this context, for this paper we use JACK as an example implementation of BDI.

In other research, Prouskas and Pitt [20] describe a layered time-aware agent architecture that enables real-time operation and allows humans and agents to interact in different time frames. The principal feature is the specification of three layers: a real-time layer, a time-aware layer and an application layer.

Jonker et al. [18] investigate dynamic BDI profiles related to route locations. They assume that the location of an agent effects its beliefs, desires and intentions. The history of agent's mobility and observed states in different locations are used to predict future states. They describe an information agent architecture for reasoning about the intentions of the customers of a mobile location-based service.

The outline of the remainder of this paper is as follows. In section 2 we give an overview of the BDI model. Section 3 briefly discusses time and temporal logics. In section 4 we introduce our proposal for incorporating temporal reasoning into BDI. Finally, we summarise and discuss future research.

2   BDI: An Overview

The Belief-Desire-Intention model, as implemented in PRS, is a mature architecture dating from the mid 1980s. PRS is the most durable agent architecture developed to date [24]. BDI combines three important elements: it has a sound philosophical basis [5]; it has been rigorously formalised in a family of BDI logics [21, 24]; and it has been implemented several times and used in a number of complex applications.

The philosophy behind BDI is Bratman's theory of human practical reasoning [5]. Bratman's work relies ultimately on folk psychology: the concept that our mental models of the world are, in essence, theories. He shows the importance of intention, a distinct mental attitude different in character from both desire and belief, in understanding how limited agents plan ahead and coordinate their actions.

Rao and Georgeff developed a formal BDI theory that connects possible worlds by modal operators representing beliefs, desires and intentions [21]. Wooldridge has extended this work to define the Logic Of Rational Agents [24].

Implementations of BDI have mainly followed the PRS model (dMARS [8], JACK [6]), with some hybrid architectures and extensions of PRS, such as JAM, which

draws from PRS (particularly UM-PRS), Structured Circuit Semantics and Act plan interlingua [15]; GRATE* [17]; and JACK Teams [4].

2.1 Implementing BDI: Hierarchies of Intention

In this paper we refer to JACK as an example BDI implementation. Programming in JACK means defining certain entities: Agent, Capability, Event, Plan and BeliefSet. Events are the driving force in the JACK model: without events, agents would lie idle. Events may be triggered by plan statements, by changes in a beliefset, or by time trigger; events may also be received from other agents. Broadly speaking, an agent handles an event by selecting a plan instance from among those that are declared to handle the event. Beliefsets encapsulate the agent's beliefs about the current state of the world. Capabilities are functional groupings within an agent [6].

The handling of an event develops over time into a 'hierarchy of plans and events' [16]. At the top is an 'event-as-intention' representing the agent agreeing to handle an external event. This links to the plan selected to handle the event, which links to the steps in that plan. A plan step may instigate an event: if so, and if the event is handled by the agent, it links to a new event-as-intention.

When an event is instigated by a plan, the handling of the new plan may be either synchronous or asynchronous with the original plan. The hierarchy of plans and events can be split into separate hierarchies, by stipulating a new hierarchy when the posted event is to be handled asynchronously. We call these separated-out hierarchies the 'hierarchies of intention' [16].

When a plan fails, either the event it is handling fails or another plan is selected to run in its place. Thus, the structure of the hierarchy of intention, and therefore the plans executed and the order of their execution, depends upon circumstance.

An agent's processes can be considered as a set of asynchronous tasks, each of which is a hierarchy of intention.

2.2 Logic of Rational Agents

LORA [24] provides a rigorous basis for BDI agent theory. It combines and extends (notably) seminal work by Rao and Georgeff on the theoretical basis of BDI [21] with the work of Cohen and Levesque on agent communication and cooperation, and in particular the dynamic logic incorporated into LORA [7].

LORA consists of four components: first-order logic; a BDI logic; a temporal logic (CTL*); and a dynamic logic.

The first component is classical first-order logic with the standard quantifiers $\forall$ and $\exists$. LORA is therefore an extension of classical logic.

The temporal component is CTL*, which combines a linear-time past with a branching-time future to represent a model of the agent's world. Since this is a model of a world, rather than a description of the real world, the nodes in the CTL* tree should properly be regarded as agent states, rather than as time points.

The BDI component introduces modal operators representing beliefs, desires and intentions, which connect possible worlds (each represented by a CTL* tree). Wooldridge systematises the basic relationships between these operators, including

those that involve the temporal operators A ("on all future time lines") and E ("on at least one future time line"), and considers each in turn to determine their usefulness.

The action component consists in labelling the state transitions in the CTL* tree with actions [24, p62]. The labelled tree is essentially a specification of a program to be followed until the requirements for the next perception / belief revision cycle are fulfilled. At that point the BDI relationships and the CTL* tree may be renewed.

Again we emphasise that the internal temporal logic (or 'clock') of an agent has no direct relevance to actual, external time. However, the form, or structure, of the logic is relevant to the problem of calculating how long a plan will take to complete.

3   Temporal Logic

In this section we discuss some aspects of time and temporal logic. We look at two approaches: first-order temporal logic, which attaches times to logical propositions, and tense logic, which is an application of modal logic. Through the modal logic approach it is possible to express certain theories about the structure of time. We present some theories as alternative answers to certain questions, and consider those questions also with regard to the computer representation of time. Finally, we identify two structures which together should enable the twin aims of reasoning about how long a plan will take to complete, and reasoning about planning and scheduling.

3.1   The Logic of Time: Two Approaches

Modern philosophers have been interested in logic mainly as it relates to language. The first modern approach to the logic of time was that of Russell and Quine. They apply the mathematical paradigm of first-order terms with associated times. Russell carefully speaks of propositions about an entity 'and' a time and refers to propositions such as $f(x, t)$. Quine speaks of properties of temporal stages of entities, substituting propositions of the form $f(x@t)$. The latter is to be understood as a proposition about a time slice of an entity that exists in the four dimensions of time and space [11].

Tense Logic grew from Prior's attempt to use modal logic principles to describe time. It is generally regarded as a multiply modal logic, with operators for the past and for the future. Prior's style of analysis is known as the modal approach to time, distinguishing it from Russell and Quine's first-order approach [11].

3.2   The Topology of Time

One benefit of tense logic is the ability to express certain theses about the structure of time [11]. Galton nominates three questions: (1) Is time discrete or continuous? (2) Is time bounded or unbounded? (3) Is time linear, parallel or branching? These questions are discussed at length from a philosophical point of view by Newton-Smith [19].

Newton-Smith argues a prima facie case against discrete time. Time is at least dense, although there is no empirical justification for stating it to be continuous. His argument against discrete time applies, he claims, even assuming there is a smallest measurable time. He argues that intervals of time are more basic than instants, and demonstrates that instants can be defined in terms of intervals.

Question 2 does not assume a defined metric, and so is not the same as the question whether time is finite or infinite in extent [11]. Rather, it asks whether there is a first or last moment in time. Newton-Smith notes that the former violates the causal principle, while the latter breaks the law of conservation of mass-energy.

For Question 3, Newton-Smith gives no prima facie case or definitive answer.

Newton-Smith also considers circular time, shows that truly circular time (linear time looping back on itself) is indistinguishable from cyclical time, where world states periodically recur. It seems likely that a similar argument should apply to a branching-time version of circular time.

Parallel time corresponds to the situation of agents that do not interact and therefore have no knowledge of each other's time lines. It is difficult to see any use for a logic to define this situation.

3.3 The Simulation of Time

Galton's questions can also be asked of computer representations of time. Dates and times are measured by clocks and recorded as integers. Measurement and recording both introduce limits to accuracy. The representation is finite and discrete; a recorded time represents an interval of actual time. The syntax of a discrete number corresponds to the semantics of an interval.

This correspondence causes a typing problem for a representation of bounded time. A start or end time is an instant. Representing either in the simulation means using the time variable sometimes as an instant and sometimes as an interval. This would serve to complicate the associated logic.

3.4 Two Uses of Temporal Logic

For the purposes of this paper, just two uses of temporal logic need be highlighted.

Firstly, branching-time logics are used in the logical specification of programs. For an example of this we need look no further than the CTL* component of LORA.

The second application is planning and scheduling. In particular, Allen's linear, interval-based logic has been the basis for much work in this area [1, 2].

4 BDI + Temporal Logic

The problem at hand is that of reasoning about actual time. We are now ready to specify the components to achieve that task:

1. Timed Tasks. Tasks are the internal representations of hierarchies of intention. The task manager should record the time whenever a task is swapped in or out.
2. Timed Plans. Plans should record the time when they start and finish and whether the plan succeeded or failed. Of interest also is the calculation of time actually spent executing each plan, discounting plans lower in the intention hierarchy and plans executed asynchronously.
3. A model for each plan, including the expected execution time and any plans immediately lower in the intention hierarchy or that are instigated asynchronously. Logical conditions or probabilities would add structure to the model.

4. A branching-time temporal logic to convert the model into a prediction of future behaviour. We want to predict how much time an agent might spend executing the plan and any processes generated by the plan.
5. An implementation of Allen's logic to support planning and scheduling based on the predicted times.

4.1 Implementing Temporal Logics: Temporal Prolog

Gabbay's Temporal Prolog is a language to express temporal logics [10]. It extends Prolog with the modal operators 'past', 'future' and 'always'. Gabbay states three ways to describe a temporal logic, of which just one is relevant to Temporal Prolog and to this paper: presentation in predicate logic form (particularly in Horn clause logic). In the PRS model the beliefs of an agent are represented as Prolog-like facts [23].

An adaptation of Temporal Prolog may provide a basis for defining the temporal logics. Benefits would derive from modularity in code and in testing. Sánchez and Augusto describe an approach to testing an extension of Temporal Prolog in [22].

5 Conclusion and Further Research

The BDI architecture does not provide the capability for agents to reason about actual time. This paper has described a strategy aimed at remedying this by extending BDI to incorporate two temporal logics and the concepts of timed tasks and timed plans.

We have focused on enabling reasoning about planning. However, the temporal model of a plan, together with a branching-time logic to reason about it, may serve as a form of self-knowledge about the agent's own processes. The strategy outlined allows the model to some extent to evolve over time, so enabling the agent to 'learn'.

We have not discussed implementation. It is implied that this would be done using JACK. JACK is a commercial product, however, and the source code is not available. Thus, the constructs of timed tasks and timed plans would need to be simulated.

References

1. Allen, J. F.: 'Towards a general theory of action and time', Artificial Intelligence 23(2): 123-154, 1984.
2. Allen, J. F.: 'Planning as Temporal Reasoning', in Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91), J. F. Allen, R. Fikes and E. Sandewall, eds, Cambridge, MA, Morgan Kaufman, 1991.
3. AOS: JACK Intelligent Agents: JACK Manual, Release 4.1, Agent Oriented Software, Pty Ltd, 2004.
4. AOS: JACK Intelligent Agents: JACK Teams Manual, Release 4.1, Agent Oriented Software, Pty Ltd, 2004.
5. Bratman, M. E.: Intention, Plans, and Practical Reason. Cambridge, MA, Harvard University Press, 1987.
6. Busetta, P., Howden, N., et al.: 'Structuring BDI Agents in Functional Clusters', in Intelligent Agents VI: Proceedings of the 6th International Workshop, ATAL 1999, Berlin, Springer-Verlag, LNAI 1757, 2000.
7. Cohen, P. R. and Levesque, H. J., 'Intention is choice with commitment', in Artificial Intelligence 42:213-256, 1990.

8. d'Inverno, M., Kinny, D., et al.: 'A Formal Specification of dMARS', in Intelligent Agents IV: Proceedings of the International Workshop, ATAL 1997, Berlin, Springer-Verlag, LNAI 1365, 1998.

9. Emerson, E. A. and Halpern, J. Y.: '"Sometimes" and "not ever" revisited: on branching time versus linear time temporal logic', in Journal of the ACM 33(1): 151-178, 1986.

10. Gabbay, D. M.: 'Modal and Temporal Logic Programming', in Temporal Logics and their Applications, Academic Press: 197-237, 1987.

11. Galton, A.: 'Temporal Logic and Computer Science: An Overview', in Temporal Logics and their Applications, A. Galton, ed, Academic Press: 1-52, 1987.

12. Georgeff, M. P. and Lansky, A. L.: 'Procedural Knowledge', in Proceedings of the IEEE 74: 1383-1398, 1986.

13. Georgeff, M. P. and Pollack, M. E.: 'Rational Agency Project', in CSLI Monthly 2(3), 1986.

14. Howden, N., Rönnquist, R., Hodgson, A. and Lucas, A.: 'JACK Intelligent Agents - Summary of an Agent Infrastructure', paper presented at the 5th International Conference on Autonomous Agents, Montreal, Canada, 2001.

15. Huber, M.: 'JAM: A BDI-theoretic mobile agent architecture', in Proceedings of the Third International Conference on Autonomous Agents (Agents'99). New York, ACM Press, 236-243, 1999.

16. Jarvis, B.: 'Expressing Commitment in a BDI Architecture', in review.

17. Jennings, N. R.: 'Specification and implementation of a belief-desire-joint-intention architecture for collaborative problem solving', in Journal of Intelligent and Cooperative Information Systems 2(3): 289-318, 1993.

18. Jonker, C. M., Terziyan, V., and Treur, J.: 'Temporal and Spatial Analysis to Personalise an Agent's Dynamic Belief, Desire, and Intention Profiles', in Cooperative Information Agents VII: Proceedings of the Seventh International Workshop on Cooperative Information Agents, CIA 2003, Springer Verlag, LNAI. 2782: 298-315, 2003.

19. Newton-Smith, W. H.: The Structure of Time. London, Routledge & Kegan Paul, 1980.

20. Prouskas, K.-V. and Pitt, J. V.: 'A Real-Time Architecture for Time-Aware Agents', in IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics 34(3), 2004.

21. Rao, A. S. and Georgeff, M. P.: 'Decision procedures for BDI logics', in Journal of Logic and Computation 8(3): 293-344, 1998.

22. Sánchez, M. A. and Augusto, J. C.: 'Testing an Implementation of a Temporal Language', paper presented at the 20th International Conference of the Chilean Computer Science Society, 2000.

23. van der Hoek, W. and Wooldridge, M.: 'Towards a Logic of Rational Agency', in Logic Journal of the IGPL 11(2): 133-157, 2003.

24. Wooldridge, M.: Reasoning About Rational Agents, The MIT Press, 2000.

A Delegation Model
for Designing Collaborative Multi-agent Systems

Stéphane Faulkner and Stéphane Dehousse

Information Management Unit, University of Namur
Rempart de la Vierge 8, 5000 Namur, Belgium
{stephane.faulkner,stephane.dehousse}@fundp.ac.be

Abstract. Multi-agent systems are now being considered a promising architectural approach for designing collaborative information systems. In such a perspective, the concept of delegation has often been considered as a key concept for modeling cooperative behavior in MAS. However, despite considerable work on delegation mechanisms in MAS, few research efforts have aimed at truly defining a delegation model for designing MAS. This paper deals with this issue in defining the foundations for a delegation model which is aimed to help the developers during the phase of design of collaborative MAS.

1 Introduction

The collaborative information system paradigm has been growing and gaining substance in technological infrastructure (e.g., middleware and Web technologies) and application areas (e.g., Business Process Management, e-Commerce, e-Government, and virtual enterprises). They involves large networks of information systems that manage large amounts of information and computing services and cooperate as-needed to fulfill their mission.

In the last few years, one promising source of ideas for designing collaborative information systems is the area of multi-agent systems (MAS) architectures. They appear to be more flexible, modular, and robust than traditional including object-oriented ones. Research in this area has notably emphasized that MAS is conceived as a society of autonomous, collaborative, and goal-driven software components (agents).

In such a perspective, the concept of delegation has often been considered as a key concept for modeling cooperative behavior in MAS [3, 12]. A delegation is defined as the action by which an agent delegates to another agent the execution of an action or the fulfilment of a goal. However, even if the concept of delegation has received increasing attention, the majority of researches are focused either on requirement analysis [7] or on the definition of delegation models through communication acts [14] or socio-cognitive theories [4].

Few research efforts have aimed at truly defining delegation models for designing MAS. This paper deals with this issue in defining a "core" set of concepts, including relationships and constraint, that are fundamental to propose a delegation model. This model is aimed to help the developers during the phase of design of collaborative MAS.

The paper is structured as follows. Section 2 provides an overview of the delegation model and details some concepts using the Z specification language. Section 3 discusses some of the related research in delegation and concludes the paper.

2   The Delegation Model

Figure 1 depicts the Delegation Model using a UML type class diagram. An Agent is an intentional entity, which has some Beliefs and Goals that guide its action. A Belief defines current states about the MAS, while a Goal defines desired states that the Agent wants to bring about. Each Agent occupies one or more Roles that are a characterization of the expected behavior of an Agent in the MAS. A Role requires a set of Plans to fulfill the Goals for which it is responsible. A Plan defines a sequence of actions.
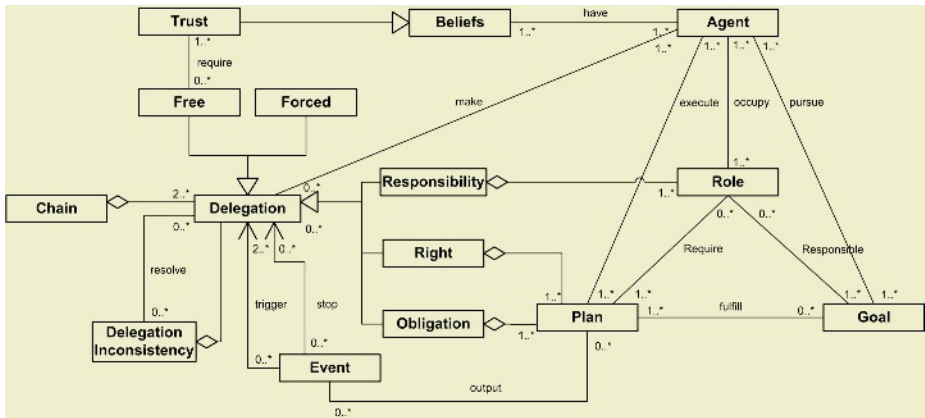
Fig. 1. The delegation Model

In order to easier or better achieve its Goals, an Agent can delegate to another Agent, called the delegatee, the execution of a Plan or the responsibility of a Role (i.e., Responsibility). The Delegation of a plan can be classified as a Right or an Obligation. A Delegation of Right defines one or more Plans that the delegatee is allowed (or not) to execute, while a Delegation of Obligation defines Plans that the delegatee must execute. The Delegation can also be self-determined (i.e., Free) or imposed by other Agents (i.e., Forced). In the case of Free Delegation, Agent requires Trust in the delegatee in order to effectively make the Delegation. Trust is a Belief an Agent has on the ability and the dependability of another agent to execute a Plan or to achieve a Goal. An agent can also be allowed to re-delegate a given delegation to another agent. This kind of situation leads to a Chain of delegation.

A Delegation is trigged or stopped by an Event. An Event is an instantaneous state of the system. It is either output of a Plan, or exogenous to the system.

Because Agents can collaborate autonomously, Delegation Inconsistencies may exist. A Delegation Inconsistency concerns two or more Delegations that cannot be assigned to the same Agent in the system. The resolution to a Delegation Inconsistency takes the form of new Delegations.

Figure 1 shows only concepts and relationships. In the next sections, we specify attributes and constraints of some concepts using the Z state-based specification language [16]. Following UML to Z translation rules from Shroff and France [15], attributes are specified as Z state variables and constraints as Z predicates. However, due a lack of space we only specify in details the concept of Delegation and its specializations (i.e., Responsibility, Right, Obligation, Free and Forced).

2.1 Delegation and Chain of Delegation

Figure 2 shows the Z formal specification of the Delegation concept. The first part of the specification represents the definition of types. A given type defines a finite set of items. The Delegation specification first defines the type Name (which represents the Name attribute) by writing [Name]. Such a declaration introduces the set of all names, without making assumptions about the type (i.e., whether the name is a string of characters and numbers, or only characters, etc.).

[Name]
[Value]

___Delegation___

delegation_name : *Name*
delegator : *Agent*
delegatee : $\mathbb{P}$ *Agent*
trigger: *Event*
expiration *:* $\mathbb{P}$ *Event*
depth : *Value*
linked_ delegation : *Boolean*

name $\neq \varnothing$ $\wedge$ delegator $\neq \varnothing$ $\wedge$ delegatee $\neq \varnothing$
$\wedge$ trigger $\neq \varnothing$

dom depth $=$ N

disjoint@Responsibility, Right, Obligationfl$\rangle$

disjoint@Free, Forcedfl$\rangle$
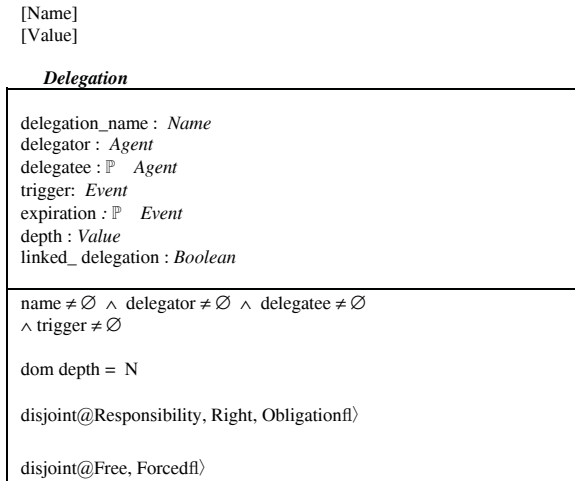
Fig. 2. Formal specification of the Delegation concept

More complex and structured types are defined with schemas. A schema groups a collection of related declarations and predicates into a separate namespace or scope. The schema in Figure 2 is entitled Delegation and is partitioned by a horizontal line into two sections: the declaration section above and the predicate section below the line. The declaration section introduces a set of variable declarations, while the predicate section provides predicates that constrain values of the variables.

A delegation defines an action by which an agent delegates to another agent the responsibility of a role, or the right or the obligation to execute a plan. It is specified with the following variables:

- delegator: identifies the agents initiating the delegation;
- delegatee: identifies the agents to which the delegation has been assigned;
- trigger: defines when the delegation starts;

- expiration: defines when the delegation stops;
- depth: defines restrictions on re-delegations. A depth set to 0 implies that no re-delegation is allowed, while a depth set to a value greater than 0 allows a chain of delegation composed of a number of re-delegations equivalent to the depth's value. The depth attribute is particularly interesting in a chain of delegation [2, 7, 9, 10] to control delegation propagation and avoid erratic re-delegation that could destabilize the system;
- linked_delegation: specifies if the delegation stops its effects when the delegator leaves the system. This variable is useful to design open MAS architectures [8] in which agents can constantly integrate or leave the system;

A Chain of delegation is an aggregation of Delegations. The predicate section of the Chain schema specifies that all delegations which belong to the same chain have an identical value for their respective linked_delegation variable.

Chain

delegations: $\mathbb{P}$ *Delegation*

#delegations $\geq$ 2

$\triangleq d_1$, d_2 : Delegation, c : Chain

$d_1 \odot c \: f \: d_2 \odot c$
  $\Rightarrow d_1.linked\_delegation = d_2.linked\_delegation$

Fig. 3. Formal specification of the Chain concept

2.2 Delegation of Responsibility

A delegation of responsibility is specified as follows:

Responsibility

Delegation
roles : $\mathbb{P}$ *Role*

---

role s $\neq \varnothing$

Fig. 4. Formal specification of the Responsibility Delegation concept

A delegation of responsibility defines a set of roles that the delegatee has to occupy. Roles provide the building blocks for agent social systems and the requirements by which agents interact [6]. The concept of role is important to abstractly model the agents in multi-agent systems and helpful to manage its complexity without considering the concrete details of agents (e.g., implementation architectures and technologies) [18]. They enable separation between different functionalities of software agents (e.g., mobility from collaboration), or between different phases of the development process (e.g., functions in the design from methods in the implementation).

A delegation of responsibility involves that the delegatee which has to occupy the roles is constrained to fulfill the goals for which the roles are responsible. However,

the delegation of responsibility is the least restrictive kind of delegation. Indeed, it only constrains the delegatee to fulfill goals without constraining the selection or the execution of plans which will make possible to fulfill them.

2.3 Delegation of Right

A delegation of right defines a set of plans that the delegatee is allowed to execute. It is specified in Figure 5. This definition allows the representation of the concept of authorization or permission [4, 7, 9, 10] that are often described in the literature. An authorization is defined as the right to grant to an agent an access to a resource in the system. In our case, this kind of authorization can be modeled by a delegation of right on the plan with which the resource is associated. For instance, the authorization on a personal data resource can be modeled as a delegation of right on the plan which accesses to the personal data.

Right
_____

Delegation
plans : $\mathbb{P}$ *Plan*
_____

plans $\neq \varnothing$

Fig. 5. Formal specification of the Right Delegation concept

2.4 Delegation of Obligation

A delegation of obligation defines a set of plans that the delegatee must (or not) execute. It is specified in Figure 6. The polarity variable describes whether the obligation is positive or negative. A negative delegation of obligation enables an agent to forbid another agent to execute a plan. It corresponds to the notion of prohibition [7].

By combining delegation of responsibility and delegation of obligation, an agent can delegate a role and force the execution of one or more plans in order to fulfill a goal for which this role is responsible. Such a combination is useful to constraint some agent's behaviors in cooperative MAS. For instance, an agent may delegate to another agent a role with the goal *get personal data* and a delegation of obligation on the plan, *encrypt the data*. For the delegator, this ensures that whatever the plans used to get the personal data, they will be encrypted.

[Obligation_polarity] := Positive | Negative

Obligation
_____

Delegation
polarity: *Obligation_Polarity*
plans : $\mathbb{P}$ *Plan*
_____

polarity $\neq \varnothing$ $\wedge$ plan $\neq \varnothing$

Fig. 6. Formal specification of the Obligation Delegation concept

2.5 Forced Delegation

Delegation is generally strictly based on trust [4]. However, [4, 7] has mentioned rarer cases of delegation where trust is not required. This kind of delegation is called a forced delegation. It occurs when an agent is in a situation of blind or coercive delegation [4]. A blind delegation occurs when the delegator does not have sufficient information to form a trust opinion on the delegatee. A coercive delegation implies one or more agents, called the requirers, that force the delegator to delegate the responsibility of a role or the execution of a plan.

Forced

Delegation
requirers : $\mathbb{P}$ *Agent*
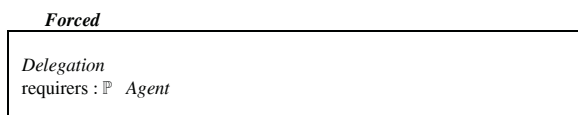
Fig. 7. Formal specification of the Forced Delegation concept

2.6 Free Delegation

The free delegation is specified as follows:

Free

Delegation
require : *Trust*
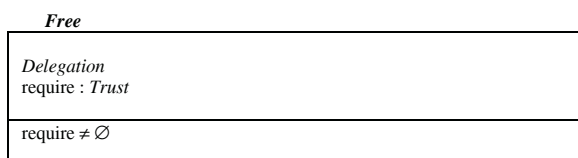
require $\neq \varnothing$

Fig. 8. Formal specification of the Free Delegation concept

Contrary to the forced delegation, a free delegation requires trust in the delegatee. Indeed, a delegation implies that the delegator exposes himself through the behavior of a delegatee. A free delegation is defined as an intentional delegation (i.e., the delegator can freely decide to delegate or not). Consequently, the delegator decide to delegate only if it trusts the delegatee.

Trust is a belief an agent has on the ability and the dependability of another agent to execute a plan or to achieve a goal. The concept of trust is essentially human mental state and therefore difficult to transpose in the agent paradigm. Authors have, many times [3, 4, 11, 13], tried to formalize this concept through, for example, computation of different components without being able to achieve the definition of a universally admitted model. However, due to a lack of space, we do not address in this work the issues related to trust models.

3 Related Work and Conclusions

The work by Giorgini et al. [7] uses a depth attribute to limit the possibility to re-delegate. This attribute is only available to delegation to grant. Our model brings the power of the depth attribute to any kind of delegation by adding it to the concept of

delegation. As a consequence, agents become able to control delegation chain either in case of delegation of right, obligation or responsibility.

In addition to the use of this attribute, they introduce the concept of prohibition which forbids an agent to use a service (task, resource or goal). The prohibition can be described in our model by a delegation of obligation with a negative polarity.

Norman and Reed [12] insist on the importance for a theory of delegation to encompass distinction between delegation on an action to be done and on a goal to be achieved. If the first form of delegation is treated in the agent literature, the second form is not present. To this end, we propose to use a delegation of responsibility on a role which is responsible of the goal to be achieved. Indeed, a delegation of responsibility involves that the delegatee which has to occupy the roles is constrained to fulfill the goals for which the roles are responsible.

In their attempt to define principles of trust for MAS, Castelfranchi and Falcone mention [4] that in exceptional cases (coercive or blind delegation) the delegator is not free to delegate. From this statement, we have defined the concept of forced delegation.

This paper has attempted to gather the key points of the different perspectives on delegation and has defined a conceptual model to help developers to design delegation mechanisms in collaborative MAS. The main contribution of this work is that our approach aims at modeling delegation to use it at the design level, while others approaches focus on requirement analysis or on the definition of delegation models through communication acts or socio-cognitive theories.

References

1. K. S. Barber and J. Kim, "Soft Security: Isolating Unreliable Agents", Proceedings of the AAMAS 2002 Workshop on Deception, Fraud and Trust in Agent Societies, Bologna, Italy, July 2002.
2. M. Blaze, J. Feigenbaum, J.Ioannidis, and K. Keromytis, "The Role of Trust Management in Distributed System Security", Secure Internet Programming, J. Vitec, and C. Jensen (Eds)., 1999.
3. J. Carter, E. Bitting and A. A. Ghorbani: Reputation Formalization Within Information Sharing Multiagent Architectures, Computational Intelligence, Vol 18, No. 4, pp. 45-64, 2002.
4. C. Castelfranchi and R. Falcone. Principles of trust for MAS: cognitive anatomy, social importance, and quantification. In Proceedings of the International Conference of Multi-Agent Systems (ICMAS'98), pages 72–79, 1998.
5. S. Faulkner, M. Kolp and P. Y Shobbens.: An Architectural Description Language for BDI Multi-Agent System,Submitted to Journal of Autonomous Agents and Multi-Agent Systems, 2005
6. J. Ferber, Multi-agent System: An Introduction to Distributed Artificial Intelligence. Harlow, Addison-Wesley Longman, 1999.
7. P. Giorgini, F. Masscci, J. Mylopoulos, and N. Zannone: Filling the Gap Between Requirements Engineering and Public Key/Trust Management Infrastructures, Proc. of the 2nd Int. Conf. on Trust Management iTrust 2004.
8. D. Huynh, N. R. Jennings, and N. R. Shadbolt. FIRE: An integrated trust and reputation model for open multi-agent systems. In Proceedings of the 16th European Conference on Artificial Intelligence (ECAI), 2004.

9. L. Kagal, T. Finin, and Y. Peng. A Framework for Distributed Trust Management. In Proceedings of IJCAI-01, Workshop on Autonomy, Delegation and Control, 2001.
10. N. Li, B. N. Grosof, and J. Feigenbaum, "A Logic-based Knowledge Representation for Authorization with Delegation", IBM Research Report, May 1999
11. L. Mui, A. Halberstadt & M. Mohtashemi, "Evaluating Reputation in Multi-agents Systems", Trust, Reputation, and Security: Theories and Practices, Springer-Verlag, Berlin, 2003.
12. T. J. Norman and C. Reed: A Model of Delegation for Multi-Agent Systems, Computer Science, Vol. 2403, pp. 185-204, 2002.
13. S. D. Ramchurn, D. Huynh, and N. R. Jennings. Trust in multi-agent systems. The Knowledge Engineering Review, 2004.
14. J. Sabater. Trust and Reputation for Agent Societies. Phd thesis, Universitat Autnoma de Barcelona, 2003.
15. M. Shroff and R. B. France, Towards a formalization of UML class structures in Z, Proceedings of the 21st International Computer Software and Applications Conference, IEEE Computer Society, 1997
16. J. M. Spivey, The Z notation: a reference manual, Prentice Hall International (UK) Ltd., Hertfordshire, UK, 1992
17. H.C. Wong, and K. Sycara: Adding Security and Trust to Multi-Agent Systems, Proceedings of Autonomous Agents'99 (Workshop on Deception, Fraud and Trust in Agent Societies), 1999
18. Qi Yan, Xin-Jun Mao and Zhi-Chang Qi, Modeling role-based organization of agent system, UKMAS'02

An Interactive Visual Model for Web Ontologies

Yuxin Mao, Zhaohui Wu, Huajun Chen, and Xiaoqing Zheng

College of Computer Science, Zhejiang University, Hangzhou 310027, China
{maoyx,wzh,huajunsir}@zju.edu.cn

Abstract. Web ontologies as the foundation of the Semantic Web were proposed to integrate heterogeneous information resources in the Web; howbeit several intrinsic limitations of Web ontologies represented in current Web ontology languages make them unsuitable for interactivities. Concept maps that provide a visual language for organizing and representing knowledge are widely used in many disciplines and application domains. This paper mainly describes an interactive visual model Web Ontology Map (WOMap) that extends Web ontologies with concept maps to sharing and exploring large-scale knowledge in an attractive and efficient way towards the Web.

1 Introduction

Web ontologies[1] as the foundation of the Semantic Web [10] describe heterogeneous information resources in the existing Web. Utilizing Web ontologies to integrate information resources is no longer a novel technique and the emerging Semantic Web has brought us a lot of ontological applications; howbeit several intrinsic limitations of Web ontologies make them unsuitable for interactivities.

Web information will be ultimately utilized by human beings, but Web ontology aims at machine-processable representation and the interactivities with users may not be under advisement in many Semantic Web applications. Alternatively, concept map [1] that has been used as a powerful assistant tool for human beings to learn and create knowledge meaningfully may complement Web ontologies on the interactive aspect.

The expression in Web ontologies can be viewed as a directed labelled graph. However, each concept and each relation in Web ontologies is a first class object [11]. When we visualize Web ontologies as RDF graphs [5], everything must be a node including properties and it will be not intuitive to grasp the relations between two classes (three coequal nodes). Most ontology visualization tools (e.g. OntoRama[2]) may not take this issue into deep consideration. In contrast, relation is indicated by a line between two nodes in concept map.

The hierarchical structure in Web ontologies is organized with the inheritance of superclass and subclass. Multidimensional relationships except for inheritance cannot be represented as hierarchy trees evidently in Web ontologies. A remarkable characteristic of concept maps is concepts are represented in a hierarchical fashion with the

---

[1] Currently, OWL (http://www.w3.org/2001/sw/WebOnt/) is the *de facto* standard of the Web Ontology, so we mainly refer Web ontologies in OWL

[2] http://www.ontorama.com/

most inclusive, most general concepts on top and the more specific, less general concepts arranged below.

Concept maps are suggested to take advantage of the visual information representation and embody rich semantics, while other diagram based methods like UML may not own the attribute to extend Web ontologies without losing semantics. We can use concept map to extend the existent Web ontology model for sharing Web knowledge. There are also many related works. Olivier in [9] proposed a conceptual graph based Ontolingua, and Corbett in [3] described the semantic interoperability using conceptual graph. However, we enhanced the model of Web ontology in several aspects mentioned above and our model is expected to give a new vision for Web ontology. The paper is organized as follows. We first give an overview about concept map in Section 2. Then, we describe an interactive visual model for Web ontologies and the main algorithms in Section 3. Section 4 gives evaluations and a conclusion.

2   Concept Map

The concept mapping technique was developed by Novak and Gowin [6] [7]. Concept maps have been defined in a variety of ways using a wide range of names. In the narrow sense, a concept map is a special form of network diagram for exploring knowledge and gathering information, which consists of nodes that contain a concept, item or question and links. The labeled links explain the relationship between the nodes and the arrow describes the direction of the relationship (see figure 1).

Concept maps are used most frequently in education as tools facilitating meaningful learning; however, the characteristics of concept map make it suitable for Web applications. Web Brain[3], offers a graphical search engine and converts the users' searching behaviors into exploring a concept map. KMap [2] is an open architecture concept-mapping tool allowing different forms of concept map to be defined and supporting integration with other applications. Gram and Muldoon in [8] compare a concept map-based interface to a web page-based interface in terms of accurately finding relevant information and the results indicated that the concept map-based interface resulted in better search performance for learners.

3   Interactive Visual Model for Web Ontologies

In many cases end-users need to interact with Web ontologies, but it's not easy for common users to understand or accept Web ontologies immediately without hindrance. Concept mapping is suggested to take advantage of the remarkable capabilities of the human visual perception system and the benefits of visual information representation and concept map itself embodies semantics of knowledge, so we can construct an interactive interface between end-users and Web ontologies based on concept maps.

3.1   WOMap Model

We can describe the interactive visual model itself as a concept map (see figure 1), which is so called Web Ontology Map (WOMap) Model. If we want to represent

---

[3]   http://www.webbrain.com/

concept maps based on Web ontologies, what we should do at first is to establish a bridge from ontologies to maps.
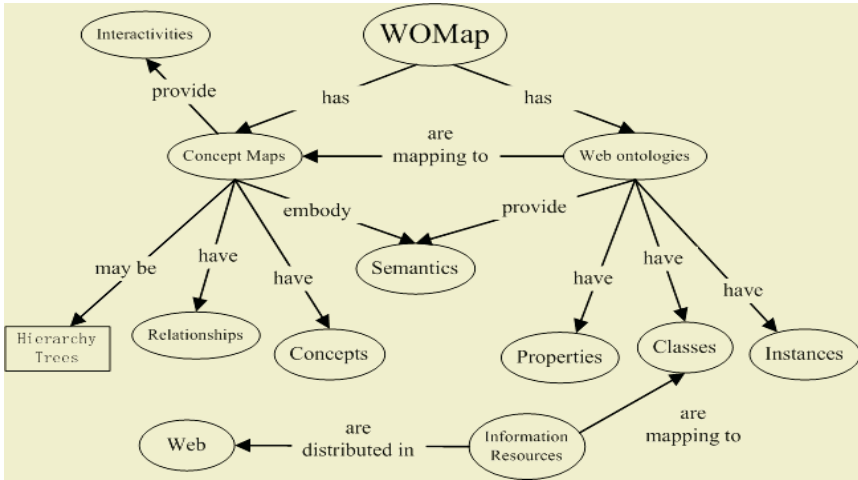
Fig. 1. The concept map of the interactive visual model for Web ontologies

Table 1. Mapping between Web ontology elements and concept map units

Web ontology elements	Concept map units	Graphical Types
Class	Concept	Circle or box
Property	Relationship	Line
Instance	Specific example	Circle or box
Statement	Proposition	Circle or box

This is, however, a basic mapping from Web ontology to concept map.

3.2 Links Construction

In concept maps, a concept is always linked to several other concepts by lines speci-fying the relationship between the two concepts. The more concepts to which a con-cept is linked, the better defined that concept is [4]. We illustrate the algorithm to find the linked classes of a focused class in Web ontologies as follow.

Algorithm of finding classes to be linked with the class user focuses on

```
SET Index to 0
EXTRACT all properties of FocusedClass into Iterator
FOR each Item in Iterator
     IF Item is not a property
         RAISE OntologyType exception
     END IF
     EXTRACT all property values of Item at "rdfs:range"
     into SubIterator
     FOR each SubItem in SubIterator
         IF SubItem is-a class or literal THEN
             STORE SubItem in LinkedClasses[Index]
```

```
                INCREMENT Index
        ELSE
             RAISE OntologyType exception
        END IF
     END LOOP
  END LOOP
```

Through this algorithm, three first class objects (the focused class, the linking property and the linked class) as well as the relations between them are reduced to a directed line with two endvertices in concept maps.

3.3 Rank Order

An important step in constructing a good concept map is to identify the key concepts that apply to the application domain and list them in a rank order. Let c be a class of Web ontology. The *semantic out-degree* of c in ontology graph is the number of properties that take c as domain, denoted by $d^+(c)$. Similarly, we can define the *Semantic In-degree* of c in ontology graph is the number of properties that take c as range, denoted by $d^-(c)$. The semantic out-degree (semantic in-degree) at a specific property p for c is represented as $d^+(c/p)$ ($d^-(c/p)$). Let p_s be the ontology property "*rdfs:subClassOf*". Then we can compute the semantic weight coefficient of c by the following equation:

$$sw(c) = d^+(c)/d^-(c) + \alpha\left(d^+(c/p_s) - \beta * d^-(c/p_s)\right). \qquad (1)$$

The semantic weight coefficient $sw(c)$ determines the relative order of c among a collection of classes (concepts) in the concept map: the larger $sw(c)$, the higher the order of c. And $sw(c)$ is mainly affected by the semantic out-degree and the semantic in-degree of c. The semantic out-degree and the semantic in-degree at the property "*rdfs:subClassOf*" also contribute to the semantic weight coefficient. The term $d^+(c/p_s)$ represents that how many direct subclasses c has and $d^-(c/p_s)$ is either 0 or 1 as most Web ontology languages disallow multi-inheritance of classes. The terms α and β are adjusted coefficients and for the convenience of computation we can assign integer value 1 to both of them.

3.4 Interactivities

Each node or line in Web Ontology Map is bound to objects in Web ontologies and thus embodies more explicit semantics than traditional concept maps. To enhance interactivities with end-users, WOMap Model extends concept maps with interactive operations:

Trace. Users can trace a map unit in concept maps of WOMap to view its complete semantic definition in Web ontologies, either in a new concept map or a list of se-

mantic information. For example, a kind of Chinese medical formula is composed of several single flavor drugs but only the knowledge about which diseases the formula can treat is visualized in a concept map, so tracing the formula concept will show the detail compounding structure.

Map. The mapping from Web ontology elements to concept map units is not injective; that is to say, several classes can be mapped to the same concept in a concept map or the concept is abstracted from a collection of synonymous or analogous classes.

Specify. Sometimes concepts are too abstract for unsophisticated users to understand. Then expert users can attach one or more relative or opposite ontology instances to a concept as its specific examples of events or objects that help others to clarify the meaning of the given concept.

Extend. Domain specific problem solving in the Web environment usually requires large-scale information other than single domain ontology. WOMap is an open structure model for Web information sharing and users can dynamically register heterogeneous information resources at Web ontology classes through visual elements in concept maps to extend the knowledge scope of the model.

Although there have been a number of ontological tools (e.g. Protégé[4] and OntoRama) that can process Web ontologies for users, those tools mainly focus on graphical control of ontologies rather semantics itself during interaction. In contrast, our model provides users with visual interactivities based on semantics of ontologies.

3.5 Hierarchy Trees

During the process of knowledge exploration a concept hierarchy will facilitate users to locate and discover knowledge efficiently. The concept hierarchical structure in Web ontologies is directly organized based on the subclass-superclass inheritance. For various Web information resources there are multidimensional relationships and the relations between ontology classes can be divided into direct relations that directly relate two classes in Web ontologies (inheritance, disjoints, equivalents et al) and indirect relations. Unlike inheritance, indirect class relations can't be viewed as concept hierarchy trees evidently in Web ontologies.

A remarkable characteristic of concept maps is that concepts are represented in a hierarchical fashion with the most inclusive, most general concepts at the top of the map and the more specific, less general concepts arranged hierarchically below and hence we can represent different kinds of relationships in a hierarchical fashion for special exploration preference. We propose an algorithm to construct a concept hierarchy tree based on semantic weight coefficients.

Algorithm of constructing a relational class hierarchy tree based on semantic weight coefficients

```
FOR each Class in Collection
    COMPUTE Semantic weight coefficient using
    equation(1)
```

---

[4] http://protege.stanford.edu

```
END LOOP
SORT all classes in Collection according to Semantic
weight coefficients
SET Collection[0] as the Root of Tree
EXTRACT all properties related to Root with "rdfs:range" into
Iterator
FOR each Item in Iterator
    EXTRACT all property values of Item at
    "rdfs:domain" into SubIterator
    FOR each SubItem in SubIterator
        IF SubItem is a class and SubItem is-in
        Collection
            APPEND SubItem to Root in Tree
            Call subTree(SubItem, Collection) (recursive)
        ELSE
            Do nothing
        END IF
    END LOOP
END LOOP
```

4   Evaluations and Conclusion

Based on the main theory of WOMap Model, we have developed an intelligent Semantic Web application called Semantic Browser [13], which has been successfully used by China Academy of Traditional Chinese Medicine as an intelligent platform for knowledge sharing and exchange. The concept maps in Semantic Browser are constructed based on the world's largest TCM ontology called Traditional Chinese Medical Language System [12] (TCMLS) and users can interact with concept maps dynamically to explore and share large-scale TCM knowledge in the Web.

In this paper, we describe an interactive visual model WOMap for Web ontologies. The WOMap Model has enhanced Web ontologies in several aspects (visualization, interaction and exploration) based on concept maps. End-users can interact with the WOMap Model to explore and share large-scale domain knowledge through intelligent operations in an intuitive and universal view towards Web information resources. Of course our model is far from being complete and we need to extend it in several dimensions cover more aspects of Web ontology in the future work.

Acknowledgement

References

1. B.R Gaines, M.L.G. Shaw.: WebMap: Concept Mapping on the Web. The Second International WWW Conference (1995)
2. B.R Gaines, M.L.G. Shaw.: Concept maps indexing multimedia knowledge bases. AAAI-94 Workshop: Indexing and Reuse in Multimedia Systems, Menlo Park, California (1994)

3. D. Corbett.: Interoperability of Ontologies Using Conceptual Graph Theory. 12th International Conference on Conceptual Structures (2004)
4. D.H. Jonassen, T. Reeves et al.: Concept Mapping as Cognitive Learning and Assessment Tools. Journal of Interactive Learning Research (1997), 8 (3/4), 289-308
5. F. Manola, E. Miller.: RDF Primer, W3C Recommendation (2004)
6. J.D. Novak, D.B. Gowin.: Learning how to learn. Cambridge: Cambridge University Press (1984)
7. J.D. Novak.: Clarify With Concept Maps: A Tool for Students and Teachers Alike, The Science Teacher (1991), 58(7), 45-49
8. M. Carnot, B. Dunn et al.: Concept Maps vs. Web Pages for Information Searching and Browsing. Proceedings of the 2003 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on Enablement through Technology (2003)
9. O. Gerbe, G.W. Mineau.: The CG Formalism as an Ontolingua for Web-Oriented Representation Languages. 10th International Conference on Conceptual Structures (2002)
10. T. Berners-Lee, J. Hendler, O. Lassila.: The Semantic Web, Scientific American (2001)
11. T. Berners-Lee.: Conceptual Graphs and the Semantic Web, W3C Design Issues (2001)
12. X. Zhou, Z. Wu, A. Yin et al.: Ontology Development for Unified Traditional Chinese Medical Language System. Journal of Artificial Intelligence in Medicine (2004), Vol.32, Issue 1, Pages 15-27
13. Z. Wu, Y. Mao, H. Chen.: Semantic View for Grid Services. Proceedings of the International Conference on Services Computing (2004), 329-335

RDF-Based Ontology View for Relational Schema Mediation in Semantic Web

Huajun Chen[1], Zhaohui Wu[1], and Yuxin Mao[1]

Zhejiang University, College of Computer Science, Hangzhou, 310027, China
{huajunsir,wzh,maoyx}@zju.edu.cn

Abstract. One fundamental issue for semantic web application is how to define the mapping between relational model and RDF model, so that the legacy data in relational databases can be integrated into semantic web. In this paper, we propose an view-based approach to mediate relational schema using RDF-based ontology. We formally define the *Ontology View*, and precisely define the semantics of answering queries using ontology view. With our approach, we highlight the important role played by *RDF Blank Node* in defining semantic mappings and representing the incomplete part of relational schema.

1   Introduction

For many semantic web applications, it is required to integrate legacy relational databases using RDF-based ontologies, so that the relational data can be retrieved and manipulated by formulating RDF queries upon the ontologies. Essentially speaking, this is the problem of uniformly querying many disparate relational data sources through one common virtual interface. In this case, relationship or mappings must first be established between the source schemas and the target common schema. A typical approach to define this kind of mappings is called answering query using view [1] [2].In this paper, we study the problem of view-based relational schema mediation using RDF-based ontology[1]. While most of the preceding work has been focused on the relational case, and recently the XML case, we consider views associated with RDF ontologies.

In more specific, we introduce the *Ontology View* approach to defining mapping between source relational schema and target RDF ontologies, and precisely define what it means to answer a target RDF query, given a set of such kind of ontology views. We define a *Target RDF Graph Instance* that satisfies all the requirements with respect to the given views and source instances, and we take the semantics of query answering to be the result of evaluating the query on this *RDF Graph Instance*. With our approach, we highlight the important role played by *RDF Blank Node* in defining semantic mappings and representing the incomplete information or hidden semantics of relational schema.

Related Work. There are a lot of semantic tools that concern mapping RDF with relational model. Some of them deal with the issue on using RDBMS as RDF triple storage, such as Sesame [6], KAON [5], Jena [4], etc. Others deal with the issue on exposing relational data as RDF, such as D2RMap [7],D2RQ [3] and RDF Gateway[2]. Our work resembles the latter case such as D2RMap, etc. Comparing with them, we employee a view-based approach which has been proven to be a flexible way to define semantic mappings. Furthermore, the formal aspects of the query answering including the query semantics and query complexity are not reported for these work. [9] considers the problem of answering query using views for semantic web,but their approach is more description-logic-oriented.

The rest of the paper is organized as follows: Section 2 defines the *Ontology View* by some examples. Section 3 deals with the issue of the semantics of answering query using *Ontology View*. We conclude in Section 4.

2   Ontology View

We start our discussion with a simple use case: suppose both W3C organization and Zhejiang University (abbreviated as ZJU) have a legacy relational database about their employees and projects, and we would like to integrate them by the FOAF ontology[3], so that we can query these relational databases by formulating RDF queries upon the FOAF ontology.

The mapping scenario in Figure 1 illustrates two source relational schemas (W3C, and ZJU),a target RDF schema (a part of the foaf ontology), and two mappings between them. Information in the two sources may overlap:the same project may appear in both sources.Graphically, the mappings are described by the arrows that go between the "mapped" schema elements. Essentially speaking, a RDF view is a mapping assertion of the form $Q^S \rightarrow Q^T$,where Q^S is a query over the source relational schema and Q^T is a query over the RDF ontology. These mappings specify a containment assertion: for each tuple returned by Q^S, there must exist a corresponding tuple in Q^T. For instance, the right part of the view *V1* can be viewed as a RDF query Q^T over the foaf ontology, and the view specifies that: for each tuple of the *w3c:emp* table, there must exist a corresponding RDF graph instance that matches the query pattern of Q^T.

We formally define the *Ontology View* as below.

Definition 1.RDF Triple Assume there is an infinite set U, called RDF URI references; an infinite set B, called the Blank Nodes; and an infinite set L, called RDF literals. A triple $(s, p, o) \in (U \vee B) \times (U) \times (U \vee B \vee L)$ is called an RDF triple. In such a triple, s is called the subject,p the predicate and o the object.We use UBL to denote the union of the set of U, B, and L.

Definition 2.RDF Graph An RDF graph is a set of RDF triples.

Definition 3.Ontology View Let Var be a set of variable names disjoint from UBL. A typical Ontology View is like the form:

---

[2] RDF Gateway: http://www.intellidimension.com

[3] Friend of a Friend (FOAF) project:http://www.foaf-project.org/
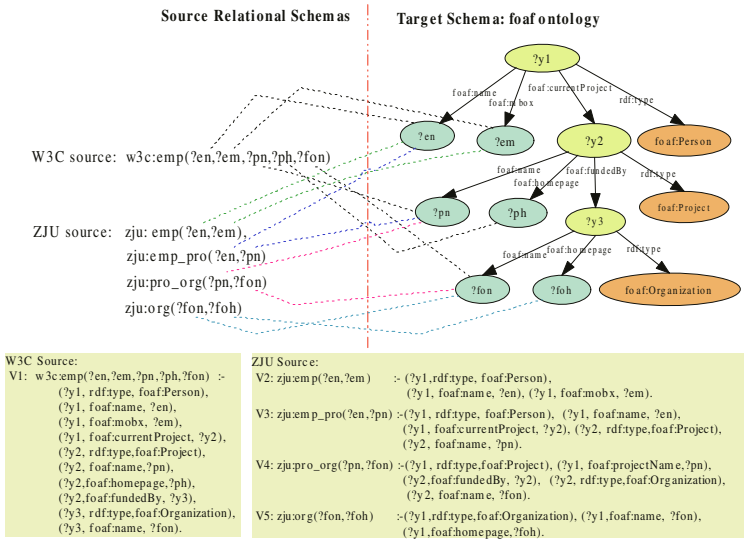
Fig. 1. The Ontology View Example: The symbols $?en, ?em, ?pn, ?ph, ?fon, ?foh$ are variables and represent, respectively, "employee name", "mail box", "project name", "project homepage", "funding organization name", "funding organization homepage". The foaf ontology consists of three classes: *foaf:Person, foaf:Project, foaf:Organization.*

$R(\bar{X}) : G(\bar{X}, \bar{Y})$, where:

1. $R(\bar{X})$is called the head of the view, and R is a relational predicate ;

2. $G(\bar{X}, \bar{Y})$ is called the body of the view, and G is a RDF graph with some nodes replaced by variables in Var.

3. The $\bar{X}, \bar{Y}$ contain either variables or constants. The variables in $\bar{X}$ are called distinguished variables, and the variables in $\bar{Y}$ are called existential variables. We require that the view be safe, that is to say, every variable that appears in the head must also appear in the body. And we often denote individual existential variables as "?y1,?y2,...".

3   The Semantics of Ontology View

In this section, we consider the issue of query semantics using ontology view. The fundamental problem we want to address is: *given a set of source instances I,i.e., a set of source relations, and given a set of Ontology Views such as $V1, V2$,etc., what should the answers to a target RDF query be?*

3.1   RDF Query

Roughly speaking, a RDF query can be viewed as a RDF graph with some nodes replaced by variables. The following example *q1* illustrates a query over the target

RDF ontology in Figure 1. The queries are written in SPARQL[4] query notation. This query asks for all tuples with person name (?en), mail box (?em), project name (?pn), project homepage, the homepage of the organization funding the project(?foh). There is one optional blocks in the *where* part. According to the SPARQL specification, the *OPTIONAL* predicate specifies that if the optional part does not lead to any solutions, the variables in the optional block can be left unbound.

```
(q1):select ?en,?em,?y2,?pn,?ph,?foh where
     (?y1 rdf:type foaf:Person) (?y1 foaf:name ?en)
     (?y1 foaf:mbox ?em) (?y1 foaf:currentProject ?y2)
     (?y2 rdf:type foaf:Project) (?y2 foaf:name ?pn)
     (?y2 foaf:homepage ?ph) (?y2 foaf:fundedBy ?y3)
     (?y3 rdf:type foaf:Organization)
     OPTIONAL(?y3 foaf:homepage ?foh)
```

We then formally defined the RDF query we consider in this paper as below.

Definition 4. RDF Query Let Var be a set of variable names disjoint from UBL. A RDF query is like the form: $H(\bar{X}) : -G(\bar{X}, \bar{Y})$, plus a set of constraint C, where:
1. $H(\bar{X})$ is called the head of the query;
2. $G(\bar{X}, \bar{Y})$ is called the body of the query, and G is a RDF graph with some nodes replaced by variables in Var.
3. The $\bar{X}, \bar{Y}$ contain either variables or constants. The variables in $\bar{X}$ are called distinguished variables, and the variables in $\bar{Y}$ are called existential variables.
4. The C is a subset of the variables occurring in H. The variables in C can be left unbound if no solution can be found.

3.2   The Query Semantics

To define the semantics of answering query using view,one possible approach that has been extensively studied in the relational literature, is to consider the target instance, which is yielded by applying the view definition onto the source instances, as an *incomplete databases* [2][8]. There are often a number of *possible databases D* that are consistent with this *incomplete database*. Then the query semantics is to take the intersection of $q1(D)$ over all such possible D. This intersection is called the set of the *certain answer*.

We take a similar but somewhat different approach,which is more RDF-inspired. We define the semantics by constructing a *Target RDF Graph Instance G* based on the source instances and ontolgoy views, then define the result of a target RDF query to be the result of evaluating it on G.

The construction process goes like this way: for each tuple in the sources, we add a set of RDF triples in the target so that the ontology view is satisfied.For

---

source 1 :
    w3c:emp("Dan Brickley","danbri@w3.org","SWAD","http://swad.org","EU") ; w3c:emp("huajun","huajunsir@zju.edu.cn","DartGrid","http://dart.org", "NSFC").

source 2:
    zju:emp("huajun","huajunsir@zju.edu.cn"), zju:emp_pro("huajun","DartGrid"),   zju:emp("huajun","huajunsir@zju.edu.cn"), zju:emp_pro("huajun","TCM-Grid"),
    zju:pro_org("DartGrid","NSFC"), zju:org("NSFC","http://www.nsfc.gov.cn");   zju:pro_org("TCM-Grid","China863"), zju:org("China863","http://www.863.org.cn").

Source Instance

Target Instance   Applying Ontology View

_:bn1 rdf:type foaf:Person;
    foaf:name "Dan Brickley";
    foaf:mbox "danbri@w3.org";
    foaf:currentProject _:bn2.
_:bn2 rdf:type foaf:Project;
    foaf:name "SWAD";
    foaf:homepage "http://swad.org";
    foaf:fundedBy _:bn3.
_:bn3 rdf:type foaf:Organization;
    foaf:name "EU".

_:bn4 rdf:type foaf:Person;
    foaf:name "huajun";
    foaf:mbox "huajunsir@zju.edu.cn";
    foaf:currentProject _:bn5;
    foaf:currentProject _:bn6.
_:bn5 rdf:type foaf:Project;
    foaf:name "DartGrid";
    foaf:homepage "http://dart.org"
    foaf:fundedBy _:bn7.

_:bn7 rdf:type foaf:Organization;
    foaf:name "NSFC";
    foaf:homepage "http://www.nsfc.gov.cn".
_:bn6 rdf:type foaf:Project;
    foaf:name "TCM-Grid";
    foaf:fundedBy _:bn8.
_:bn8 rdf:type foaf:Organization;
    foaf:name "China863";
    foaf:homepage "http://www.863.org.cn".

Fig. 2. Convert the source relational instances into target RDF graph instances. We also describe the RDF graph using the N3 notation at the below of the figure.

example, applying the ontology views such as $V1, V2, etc.$ on each tuple in both w3c and zju sources (in the top of Figure 2), will yields a RDF graph instance as in Figure 2.

The key notions is the *skolem functions* we introduced to generate the blank node IDs in the target RDF instance.As can be seen, corresponding to each existential variable $?y \in \bar{Y}$ in the view, we generate a new blank node ID. For examples, _:bn1,_:bn2 in Figure 2 are both newly generated blank node IDs corresponding to the variables $?y1, ?y2$ in $V1$. This treatment of the existential variable is in accordance with the RDF semantics, since blank nodes can be viewed as existential variables. In general, we associate each RDF class in the target ontology with a unique *Skolem Function* that can generate blank node ID at that type. For instances, we associate the RDF classes with the following skolem functions respectively: *foaf:Person: SF1(?en); foaf:Project: SF2(?pn); foaf:Organization:SF3(?fon).*

The choice of function parameters depends on the constraints user want to set on the target schema. For example,$SF1(?en)$ set a new constraint that says: "if

two instances have the same value slot for the property *foaf:name*, then they are equal and have the same blank node ID". This is somewhat similar to the *Primary Key Constraint* in relational database, and is useful for merging instances stemming from different sources. For example, for project name "DartGrid", the same blank node ID will be generated for both W3C and ZJU sources, so that the data from different sources can be merged together.

Blank Nodes. For many legacy database, the semantics of the relational schemas are often not represented explicitly enough for integration. Take the w3c:emp table as the example again, the emp table imply the semantics that says "there is some project whose name is ..., whose homepage is ...", and "for each project, there is a funding organization whose name is ...", although these semantics are not explicitly described in the table schema. These semantics can be well captured by RDF blank nodes, since blank nodes fit quite well to represent the meaning of "there is something". As matter of fact, this is one of the cases of representing incomplete information. The "incomplete" problem [8] has been discovered for a long time, and considered as a very important issue with related to data integration system, and it is more acute for semantic web applications because they are open-ended system where closed world assumption will fail.

The evaluation of *q1* on this RDF graph *G* produces the following set of triples. Note that the query require the variable ?*ph* (the project homepage) MUST be bound to an instance, but the variable ?foh (the organization homepage) can be left unbound, namely, is nullable.

Table 1. Query answers after evaluating *q1* on the RDF instance in Figure 2.

Person.name	Person.mail-box	Project	Pro.name	Pro.homepage	Org.homepage
Dan Brickley	danbri@w3.org	_:bn2	SWAD	http://swad.org	NULL
Huajun	huajunsir@zju.edu.cn	_:bn5	DartGrid	http://dart.org	http://nsfc.gov.cn

Based upon the observation above, we give the formal details regarding the semantics for target RDF query answering.

Definition 5 Target RDF Instance Let I be a source instance, V be a set of RDF views, we call a RDF graph G a *Target RDF Instance* with respect to I, V, if G satisfies the constraints of V for the given I, where: 1.G is generated by applying the RDF views on I; 2.The blank nodes in G are generated using the skolem functions associated with the RDF classes.

Definition 6 Query Answer Let q be a RDF query, then the set of the query answer of q with respect to I, V, denoted by $answer_V(q, I)$, is the set of all tuples t such that $t \in q(G)$ for the Target RDF Instance G.

Theorem 1. Given the ontology view language and RDF query language defined in Definition 3 and 4, the problem of answering query using materialized ontology view is NP-complete.

This query semantics is different with the *certain answer* in relational literatures for two practical reasons: Firstly,the query answer can contain nulls during evaluation, because of the *OPTIONAL* predicate used in RDF query, secondly, The query answer can contain newly generated blank node IDs.

Theorem 1 answers the question of the complexity of the problem of answering query using ontology view. Informally speaking, given a set of source instances, we can convert the source data into corresponding RDF instance in polynomial time. Since the query semantics of answering query using view is defined as evaluating the query upon the target RDF instance, these two problems have the same computational complexity. It has been proven that the problem of querying RDF graph with blank nodes is NP-Complete [10].Therefore,the problem of answering query using materialized ontology view is NP-complete.

4 Conclusion

This paper presents a view-based approach to mediate relational schema using RDF-based ontology.We define a *Target RDF Graph Instance* that satisfies all the requirements with respect to the given views and source instances, and take the semantics of query answering to be the result of evaluating the query on this *RDF Graph Instance*. With our approach, we highlight the important role played by the *RDF blank nodes* in defining the semantic mapping and representing. Some of the open questions that remain to be answered are: extension to a larger class of RDF queries and ontology views that would include aggregation, recursion and predefined constraints such as subclass and subproperty.

References

1. A. Y. Halevy. Answering queries using views: A survey. VLDB Journal, 10(4), 2001.6.
2. Serge Abiteboul, Complexity of Answering Queries Using Materialized Views, in PODS, 1998.
3. Christian Bizer, Andy Seaborne D2RQ -Treating Non-RDF Databases as Virtual RDF Graphs.Poster at ISWC2004.
4. Dave Reynolds Jena relational database interface, part of the documentation for the Jena Semantic Web Toolkit, HP Labs Semantic Web Activity.
5. E. Bozsak, Marc Ehrig, et al. KAON SERVER - A Semantic Web Management System. at WWW2003.
6. Jeen Broekstra, Arjohn Kampman, and Frank van Harmelen. Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema PDF, in ISWC 2002.
7. Chris Bizer, Freie. D2R MAP - A Database to RDF Mapping Language, Poster, International World Wide Web Conference 2003. 15. Y. Arens, C.A. Knoblock,
8. R. van der Meyden. Logical Approaches to Incomplete Information: A Survey. In Logics for Databases and Information Systems, pages 307-356. Kluwer, 1998.
9. Francois Goasdoue. Answering Queries using Views: a KRDB Perspective for the Semantic Web. ACM Transaction on Internet Technology, Vol.V, June2003,P1-22.
10. Claudio Gutierrez, Carlos Hurtadouc, and Alberto O.Mendelzonut, "Foundations of Semantic Web Databases," in PODS 2004.

Essentialized Conceptual Structures in Ontology Modeling

Patryk Burek

Department of Computer Science, University of Leipzig, Germany
`burek@informatik.uni-leipzig.de`

Abstract. Psychology and cognitive science show that human concepts possess particular structures (conceptual structures). However, in the process of ontology modeling information concerning the structure of human concepts is lost. In ontologies concepts are typically represented as undifferentiated collections of necessary (or necessary and sufficient) conditions. The lack of representation of conceptual structure may cause ontologies to be inadequate and may limit their usability for human users. We present an attempt to bring ontology modeling closer to theories of conceptual structures, in particular to psychological essentialism. A metaontology is developed to support the representation of conceptual structure, in particular the distinction between essential and merely necessary conditions.

1 Introduction and Motivation

Ontology modeling (OM) is the process of developing a domain model (an ontology) upon not or only partially formalized expert's knowledge. Thus, in order to enable machine processing of ontologies, knowledge must be formalized. However, knowledge, when formalized, may be accidentally changed or deformed; in particular cognitive aspects of knowledge are lost in the process of OM.

Much work has been devoted to support the correctness of knowledge modeling, inter alia Brachman's analysis of *is_a* [4]; Woods' analysis of subsumption [17] and the role of links in semantic networks [16]; Wand's work devoted to the notion of relationship [15]; the metaontology of OntoClean for building ontologically correct taxonomies developed by Guarino and Welty [10].

In contrast to the above our motivation is not to clean ontologies from incorrect usage of ontological constructs. Instead, we aim to enable cognitive adequacy of ontologies. By a cognitively adequate ontology we mean an ontology, in which the structure of ontological categories resembles the structure of human concepts [18]. Cognitively inadequate ontologies, and models in general, may have a limited usability for human users. Moreover, an explicit representation of conceptual structures permits the identification of subtle differences in the meaning of concepts, and thus makes ontologies more precise.

Our second motivation is to make ontological concepts more comprehensive for human users. It seems that formal definitions of ontological concepts are hardly comprehensible for human users. In contrast the explicit representation of the conceptual structure of essentialized concepts permits to select essential information from the body of a concept's specification and reveals explanatory dependencies in the structure of concepts.

Two assumptions underlie our approach. The first is the assumption of psychological essentialism, according to which humans do not represent concepts as collections of undifferentiated necessary (or necessary and sufficient) conditions. Rather, concepts exhibit a complex structure, organized around essential properties. The second assumption concerns the notion of *essence*, which is not taken in a modal sense as a necessary property but as an explanation of nonessential characteristics of a category.

To represent conceptual structure we have developed a metaontology introduced in the next section. This metaontology provides means to represent the distinction between *essential* and *peculiar* characteristics and their explanatory interdependence. Those notions are introduced in section 3 and illustrated with an example in section 4. In sections 5 and 6 related works, conclusions and future work are presented.

2   Metaontology

2.1   Scope and Purpose of Metaontology

We introduce a metaontological framework for representing conceptual structure in ontologies. The framework is intended to be a metaontology, which can be understood as an ontology devoted to ontologies. Ontology is defined as an explicit specification of a conceptualization [8] or as an explicit, partial account of a conceptualization [9]. Thus, metaontology may be understood as an explicit, partial specification of an ontology. Our metaontology is partial, since we do not aim at providing a full specification of ontologies, but only at representing the structure of human concepts underlying the ontological categories, in particular the essentialized concepts. In this sense the framework differs from other metaontological approaches, like for example SOU Information Flow Framework [11] or OntoClean, which do not share this purposes.

2.2   Common Ontology Pattern

Typically, ontologies are considered as structures of concepts, relations, functions, axioms and instances [8], where relations, functions and axioms define concepts by necessary (and sufficient) conditions. That general paradigm is adopted by methodologies of ontology development (see [6] for an overview) and by ontology representation formalisms. For example, in the Web Ontology Language (OWL) and in the underlying description logics (DL) concepts are defined by means of *inclusions* and *equalities* [2]. An analogous pattern is adopted in modeling paradigms used in software engineering. For example, in the Unified Modeling Language (UML) [14] classes, corresponding to ontological concepts, are defined by means of attributes, operations and associations, which can be understood as necessary conditions for class membership.

2.3   Cognitive Ontology Pattern

Since our purpose is to focus on conceptual structures we take a different perspective on ontologies than in the pattern sketched above. We define an ontology as follows:

Def. 1. An ontology $O = $ *(Cat, Char, char, cat-str)* consists of a set *Cat* of categories, a set *Char* of characteristics, a characterizing relation $char \subset Cat \times Char$, which as-

signs to each category c certain characteristic ch, which characterizes it, and a function $cat\text{-}str: Cat \times Char \to Cat \times Char$, which reflects the structure of a category by assigning to its characteristic some other characteristic of the same or other category.

In this sense, an ontology is a structure of characterized categories, where categories correspond in general to mental concepts, whose meaning is revealed by characteristics. By a category we understand each element of an ontology that is characterized - each which has an explicit specification. In this sense the notion of category is more general than the notion of concept in the pattern discussed in the previous subsection. The set of categories is not disjoint with the set of relations, functions and instances, since all of them may be characterized by some characteristic and thus all of them are meaningful elements of an ontology.

A characteristic is a part of an explicit specification of a category; it is what is said about a category. Characteristics are assigned to categories by a binary characterizing relation $char(c,ch)$, where characteristic ch, $Char(ch)$, characterizes category c, $Cat(c)$. One characteristic may characterize many categories and each category may be characterized by one or more characteristics.

The assumption taken is that not all characteristics of a category have the same status. Therefore categories are not mere collections of undifferentiated characteristics, but rather they have an internal structure. To represent a category's structure we use a category structure relation $cat\text{-}str(c_1,ch_1,c_2,ch_2)$, where $Char(ch_1)$ and $Char(ch_2)$ characterize respectively $Cat(c_1)$ and $Cat(c_2)$.

In a logical context characteristics ought to be understood as axioms, or meaningful parts of axioms. In UML, attributes, methods and associations, defining some UML class should be taken as characteristics. The notion of characteristic is intended to collect various definitional constructs under one umbrella, like UML associations and attributes, DL restrictions, subsumption and others.

3   Conceptual Structure: Essential Characteristics

In modeling formalisms categories are represented as collections of undifferentiated characteristics, which means that all characteristics of a category have equal definitional status. For example, none of the characteristics composing a DL axiom or a UML class have a distinguished status. However, humans do not represent all concepts as undifferentiated collections of mere necessary (or necessary and sufficient) conditions, but human concepts rather may have various intra-structures (see [12] for an overview). We consider one of such conceptual structures, namely the one postulated by psychological essentialism.

3.1   Psychological Essentialism

According to psychological essentialism, two types of properties formulate human concepts – so-called essential and surface properties. Surface properties are considered as caused and constrained by deeper – essential features [1]. Psychological essentialism states that people act as if concepts have some essential, deep properties that are both criteria for concept membership and responsible (preferably in a causal sense) for other surface features of concepts [13]. We leave aside the argument of the

deep character of essence, skipped also in some forms of psychological essentialism [3], and concentrate on the causal and explanatory character of essence.

3.2 Essence as Explanation

Since essence is crucial for category membership, the question arises of what essence is. One possible answer is given by modal essentialism, where necessary characteristics are identified as essential. However this solution seems to be too general, since not all necessary characteristics are essential. For example, the characteristic of *being a member of a singleton {Socrates}*, although necessary for *Socrates*, should not be considered as essential [7]. Rather, we say that the set of essential characteristics is a subset of the set of necessary characteristics of a given category. We call nonessential necessary characteristics *peculiar.*

We understand essence as such a necessary characteristic of a category that has a prior status in comparison to the remaining necessary characteristics of that category. To denote necessary characteristic *ch* of category *c* we use a binary relation *nec-char(c,ch)*, which is a subrelation of relation *char(c,ch)*. Then, we introduce two subrelations of *nec-char(c,ch)*: essential characterization *ess(c,ch)*, and peculiar characterization *pec(c,ch)*[1]. In the first case *ch* is an essential characteristic of category *c*, in the second *ch* is a peculiar characteristic of *c*.

The priority of essence over mere necessary characteristics is founded on the relation of explanation. In this sense, essence is an explanation of some or all of the characteristics of a category:

Def. 2. A necessary characteristic *ch* of a category *c*, *nec-char(c,ch)*, essentially characterizes *c* iff *ch* explains at least one of the remaining necessary characteristics of *c*, and it is not explained itself by any other necessary characteristic of *c*. A characteristic *ch* essentially characterizing a category *c* is called an essence of *c* and denoted by *ess(c,ch)*.

We introduce an explanation relation, *ex(c,ch,c',ch')*, as a subrelation of the relation depicting category structure *cat-str(c,ch,c',ch')*, with the meaning that a characteristic *ch* of a category *c* explains characteristic *ch'* of category *c'*. A characteristic *ch* is an explanation of *ch'* if *ch* provides an answer for a why-question concerning *ch'*.

Preliminarily, we consider two types of explanation: functional explanation *f-ex(c,ch,c',ch')* and causal explanation *c-ex(c,ch,c',ch')*. Roughly speaking, a causal explanation of *x* is such that provides a cause of *x*, while functional explanation provides a function or purpose of *x*.

4 Example

Let us now illustrate the notions introduced by means of their application to defining categories in ontologies. Consider the following two toy-categories h_1 (Heart$_1$) and h_2 (Heart$_2$) defined as follows in DL:

---

[1] Not a characteristic as such is necessary, essential or peculiar, but a characteristic in the context of a category it characterizes. Thus, to be precise one should say that a characteristic characterizes a category essentially, not that a characteristic is essential. To simplify the language, in the current paper, we also use the second form as an abbreviation of the first

```
Heart₁ ⊑ ∀locatedIn.Chest ⊓ ∀pumps.Blood ⊓ =4hasPart.Cavity
Heart₂ ⊑ ∀locatedIn.Chest ⊓ ∀pumps.Blood ⊓ =4hasPart.Cavity
```

From the above we may conclude that extensionally and intensionally h_1 and h_2 are equal, since they have the same set of referents and share the same characteristics – *is located in chest; pumps blood; has four cavities* respectively denoted by ch_1, ch_2 and ch_3. We will demonstrate that although these categories are extensionally and intensionally equal they may differ nontrivially, since their internal structures may differ.

For a functional biologist the function of the heart serves as an explanation of its structure and location. For her it is essential for the heart that it pumps blood. Because of its function the heart is located in the chest and has a particular structure. Thus on the metalevel the structure of category h_1 may be represented as follows (see figure 1 for a graphical representation): *nec-char(h₁,ch₁) ∧ nec-char(h₁,ch₂) ∧ nec-char(h₁,ch₃) ∧ f-ex(h₁,ch₂,h₁,ch₁) ∧ f-ex(h₁,ch₂,h₁,ch₃) ∧ ess(h₁,ch₂) ∧ pec(h₁,ch₁) ∧ pec(h₁,ch₃)*.

For a nonfunctional biologist the function is not an explanation of the remaining features of the heart. The role of such an explanation is played by, for example, an additional characteristic concerning the structure of an organism's DNA, denoted by ch_4. The structure of the DNA could explain why the heart is characterized in a given way. Thus, the characteristic *ch4*, characterizing *organism o*, provides a causal explanation for the characteristics ch_1, ch_2 and ch_3. The internal structure of category h_2 is as follows: *nec-char(h₂,ch₁) ∧ nec-char(h₂,ch₂) ∧ nec-char(h₂,ch₃) ∧ c-ex(o,ch₄,h₂,ch₁) ∧ c-ex(o,ch₄,h₂,ch₂) ∧ c-ex(o,ch₄,h₂,ch₃) ∧ pec(h₂,ch₁) ∧ pec(h₂,ch₂) ∧ pec(h₂,ch₃)*.

Although ch_4 explains all of the characteristics of h_2, in accordance with def. 2, it is not considered its essence, since it does not characterize h_2 but *organism o, char(o,ch₄)*. Thus, we say that ch_1, ch_2 and ch_3 are peculiar characteristics of h_2 and ch_4 provides an external causal explanation for h_2. We see that the structures of the categories h_1 and h_2 are different and the categories, although intensionally and extensionally equal, differ on the metalevel.

The above example shows that the representation of conceptual structure can give a deeper understanding of categories. Some of the differences between categories concern essential properties. If essences are not represented in the model, these subtle differences are lost. Moreover, the representation of explanatory relations within categories makes them more intelligible to human users. If our toy-definitions would be extended to full specifications, then finding crucial information turns out to be complicated if not impossible for human users. In such cases essential characteristics may provide guidance in grasping the meaning of a whole category.
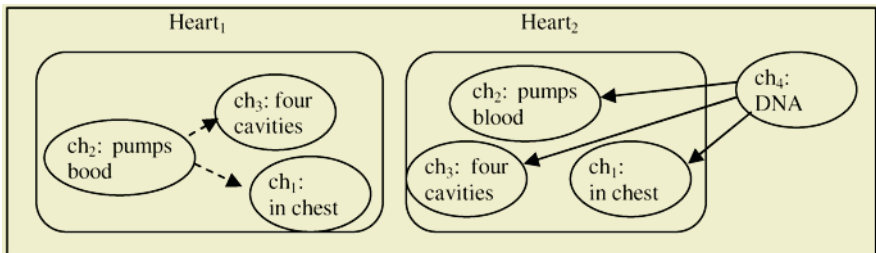
Fig. 1. Categories represented as directed graphs, where nodes represent characteristics, continuous arrows - causal explanations, and dashed arrows - functional explanations

5   Related Work

Essence is generally considered in a modal sense in knowledge modeling. An exception is Brachman's approach, where definitional properties of concepts are distinguished from incidental properties [5], [2]. That distinction corresponds to our distinction of essential and peculiar characteristics. However, Brachman suggests not including non-essential knowledge in concept's definition. Instead, he suggests modeling it by means of rules in CLASSIC [5], or in the ABox in DL [2]. We do not exclude peculiar characteristics from category specification, since they still remain necessary conditions of category membership.

6   Conclusion and Future Work

In ontology modeling not only philosophical correctness of ontologies is an issue, but also their cognitive adequacy. If ontologies are intended to represent people's knowledge, then preserving the structure of human concepts may be of value.

In the current paper a metaontology for representing the structure of concepts in ontologies is introduced. The metaontology is founded on psychological essentialism, which gives evidences that humans use essences in mental representations of concepts. Further, we refer to non-modal essentialism, according to which essence is not identified with necessary characteristics, but is taken as an explanation.

This approach provides means to represent the explanatory structure of categories and to identify and represent their essential characteristics. Since the information about conceptual structure is represented on the metalevel, the adoption of the framework does not result in an increase of the complexity of ontologies as such.

The identification of essential characteristics has several advantages: (i) it permits precise modeling of human concepts; (ii) it separates essential from merely necessary knowledge, which enables one to grasp complex categories; (iii) the explicit specification of explanatory dependencies between characteristics of categories makes them more intelligible for human users.

It is not clear whether psychological essentialism does account for the structure of all human concepts. In some cases other theories may perform better. Thus, to fulfill the postulated requirement of cognitive adequacy the framework should be extended to handle other theories of conceptual structures as well. Moreover it should be extended to handle more types of explanation and be elaborated in a form ready-to-use for ontology modeling, providing the methodology for identification of essence.

Acknowledgements

References

1. Ahn, W., Kalish, C., Gelman, S., Medin, D., Luhmann, C., Atran, S., Coley, J., Shafto P.: Why essences are essential in the psychology of concepts. Cognition, 82 (2001) 59-69

2. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.(eds.): Description Logic Handbook. Cambridge University Press (2003)
3. Barrett, H.C.: On the functional origins of essentialism. Mind and Society, 3(2) (2001) 1-30
4. Brachman, R.: What isa is and isn't: an analysis of taxonomic links in semantic networks. IEEE Computer, 16(10) (1983) 30-36
5. Brachman, R., McGuinness, D., Patel-Schneider, P., Resnick L.: Living with CLASSIC: When and how to use a KL-ONE-like language. In Sowa J.(ed.): Principles of Semantic Networks. Morgan Kaufmann, San Mateo, California (1991) 401-456
6. Fernández-López, M.: Overview of Methodologies for Building Ontologies. In IJCAI99 Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends, Stockholm (1999)
7. Fine, K.: Essence and modality. In Tomberlin J.(ed.): Philosophical Perspectives VIII: Logic and Language. Ridgeview Publishing Company Atascadero, California (1994) 1-15
8. Gruber T.: A translation approach to portable ontology specifications. Knowledge Acquisition, 5 (1993) 199-220
9. Guarino, N., Giaretta, P.: Ontologies and Knowledge Bases: Towards a Terminological Clarification. In Mars, N.: Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing. IOS Press, Amsterdam (1995) 25-32
10. Guarino, N., Welty, C.: Ontological Analysis of Taxonomic Relationships. In Leander, A., Storey, V. (eds.): Proc. ER-2000. Springer-Verlag (2000)
11. 11.Kent, R.: The iff foundation for ontological knowledge organization. In Williamson, N.J., Beghtol, C. (eds.): Knowledge Organization and Classification in International Information Retrieval. Haworth Press, New York (2004) 187–203
12. Margolis, E., Laurence, S.: Concepts and Cognitive Science. In Margolis, E., Laurence S. (eds.): Concepts: Core Readings. Bradford Books/MIT Press, Cambridge (1999) 3-81
13. Medin, D., Ortony A.: Psychological essentialism. In Vosniadou, S., Ortony A.(eds.): Similarity and analogical reasoning. Cambridge University Press, NY (1989) 179–196
14. Object Management Group: OMG Unified Modeling Language Specification. V. 1.5 (2003)
15. Wand, Y., Storey, V., Weber R.: An ontological analysis of the relationship construct in conceptual modeling. In ACM Transactions on Database Systems, 24(4) (1999) 494-528
16. Woods W.: What's in a Link: Foundations for Semantic Networks. In Bobrow, D., Collins A. (eds.): Representation and Understanding. Academic Press, NY (1975) 35-82
17. Woods W.: Understanding subsumption and taxonomy: A framework for progress. In Sowa, J. (ed.): Principles of Semantic Networks. Morgan Kaufmann, CA (1991) 45-94
18. Zhang, J.: Representations of health concepts: a cognitive perspective. Journal of Biomedical Informatics 35 (1) (2002) 17-24

Turning Mass Media to Your Media:
Intelligent Search with Customized Results

Jun Lai and Ben Soh

Department of Computer Science and Computer Engineering, La Trobe University
Bundoora, Melbourne, VIC 3086 Australia
{jun,ben}@cs.latrobe.edu.au

Abstract. In this paper we propose an intelligent web search method with customized results. This approach adopts a cosine method to calculate the similarity between document profile and customer profile. The document profile is derived from the similarity score of documents. The customers' search history is captured to generate customer profile. Then the customized search results are recommended to the end users based upon the similarity between document profile and customer profile.

1 Introduction

As the information on the web grows exponentially, it has become more and more difficult for the users to find the information they want. Customers using search engines not only expect high quality information, but also wish that the information presented were customized. For instance, an 18-year-old male who is looking for cars on the web is probably looking for something different from one that would be chosen by a 50-year-old woman. How to return the most relevant information to the end user has posed some key challenges for search engines. However most search engines and information filtering tools are typically not customized to individual users, they tend not to deliver an appropriate amount of preferred information. To solve these issues, recommender systems have been advocated and some personalized approaches to web search have been proposed. These systems recommend music, movies and products using either content-based or collaborative recommendation [9][10][11]. However, they are not customizing documents in search result. Some search engines allow user to specify the number of web pages to be retrieved, domains (e.g. .gov, .edu, .com) to be included in the search results, number of Internet Spiders to be used, and so on. However, the search result is not customized; different user will still get the same result if the query entered is the same. PAWS-Cluster [1] tries to personalize web search by comparing the returned documents and user's profile. However, the user's profile is derived based on words that appeared in the user's documents on the desktop computer, which can not accurately present user's preference.

We propose in this paper a personalized approach to web search result. This approach applies an algorithm to compute a similarity score of document based on keywords and Dublin core 15 elements and the document profile is generated. The customer profile is derived based upon customer searching and browsing behaviors. Then document profile and customer profile are analyzed and customized to present search results to users.

2   Related Work

2.1   Background

Personalized recommendation has a long history. There are massive efforts which have been made to the task of getting personalized results. Content-based recommendation and collaborative recommendation are two fundamental recommendation methods. Content-based filtering searches for items similar to those the user prefers based on a comparison of content using text-learning methods.

In collaborative filtering, documents are recommended for a user based on the likes of other users with similar tastes. User profiles are used to compare with each other. Groups of similar profiles are identified and users belonging to one group will be presented the same set of documents.

2.2   Prior Related Work

Personalized music recommendation has been done in Ringo [9] using collaborative recommendation. Movies are recommended in [10] based on mixed information of social and content categories to predict users' preferences about a movie. Personalized products in online shopping are completed in [11]. This system learns customers' preferences toward specific product features. The recommendation is then made based on product features. However, further work on personalized document search result is still lacking.

PHOAKS (People Helping One Another Know Stuff) [12] recommend relevant, high-quality information on the web to users by automatically recognizing and redistribution recommendations of web resources based on group profiles instead of individual profile. People belong to the same group will be recommended the same information. Siteseer [13] is a personalized navigation tool for the web. It is a web-page recommendation system that uses an individual's bookmarks and the organization of bookmarks within folders for predicting and recommending relevant pages. The drawback is those web pages that may be of interest to user will probably not be recommended if they are not bookmarked.

PAWS-Cluster (Personalized Adaptive Web Search with clusters) [1] uses personalized information on the client side to aid web search, then the returned documents will be compared with the user's profile which is acquired by using the words that appeared in the user's documents on the desk-top computer as the base of the profile. The user's preference cannot be accurately presented by the documents on the desktop. To this end, most customized recommendation systems are neither recommending personalized web search results nor having accurate user's preferences to produce customized web search result.

3   Customized Search Results

Every individual has a very unique way of searching and interacting with information. The way the user searches can be used to construct customer profile. Document profile is generated based on Dublin Core elements. Then the Cosine method is adopted to measure the similarity between customer profile and document profile. From the similarity, customized search results are presented.

3.1   System Design

Figure 1 shows the architecture of the proposed system. This customized searching system consists of three components: customer profile, document profile and profile comparison.
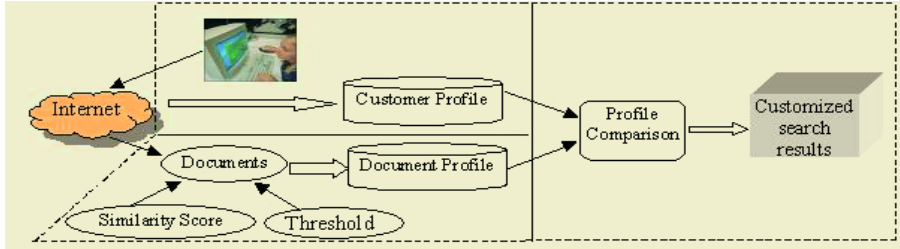
Fig. 1. Architecture of customized searching system

3.2   Document Profile (DP)

Currently, there are a number of metadata standards proposed for web pages. Among them are two well-publicized, solid efforts: the Dublin Core Metadata standard and the Warwick frame-word. The Dublin Core is a 15-Element Metadata Element Set proposed to facilitate fast and accurate information retrieval on the Internet [3]. Keywords are given different weights ranging from 0 to 1 (inclusive), depending on which element the keyword is in. Using sensitivity analysis, the weights are calibrated by the system designer, based on the importance and relevance of the keyword concerned.

Similarity score (SS) is used to form DP. To compute the similarity score of documents, we initially select some keywords appearing in those documents. Each keyword is assigned a weight, ranging form 0 to 1, as mentioned previously. For example, the weight in the element Title should be higher than that in the element Description. The number of times a word appears in the document also affects the value of relevance among documents.

The similarity score is computed as the sum of all products of keyword weight and the number of times that keyword appears in the document, using the following formula [14]:

$$SS_{(d,\ Kj)} = \sum_{j=1}^{n}(W_{Kj} * Count_{Kj}) \tag{1}$$

where:
- SS is the similarity score of a document based on the keyword K_j.
- d denotes a document ID.
- K_j is the keyword appearing in the element j of the document $(1 \leq j \leq n)$.
- W_{Kj} is the pre-defined weight of the keyword K_j appearing in the element j, determined by a system administrator via sensitivity analysis.
- $Count_{Kj}$ is the number of times that K_j appears in the element j of the document.

Table 1 shows an example of the number of times a keyword appearing in the elements (namely Title, Subject, Description, Content, and Reference) of a particular document, which is indicated by document ID (d) associated with keyword of filtering. The similarity scores (SS) for all the documents are calculated and showed in the last row using formula (1).

Table 1. Similarity score of documents

d	21	45	567	789
Keyword	Filtering	Filtering	Filtering	Filtering
Title	1	1	0	0
Subject	2	1	0	1
Description	2	7	3	7
Content	10	4	1	4
Reference	9	8	1	7
SS	12.59	11.89	3	10.5

Furthermore, the similarity score of a document based on various keywords can also be computed. Table 2 shows the SS of document ID 21 in terms of keywords "filtering", "clustering", "data" and "mean deviation"(keywords can be phrases as well):

Table 2. Similarity score of document No.21 in terms of various keywords

D	21	21	21	21
Keyword	Filtering	Clustering	Data	Mean deviation
Title	1	1	0	0
Subject	2	1	0	1
Description	2	7	1	6
Content	10	5	3	4
Reference	9	6	1	5
SS	12.59	11.59	2.6	9

Based on the similarity score, the DP is derived as follows [14]:

$$DP_{(d)} = \{ (K_j, SS_{(d, Kj)} | K_j \in C, SS_{(d, Kj)} \geq SS_{threshold} \} \tag{2}$$

where:

- d denotes a document ID.
- K_j denotes a keyword.
- C is a set of all keywords to which the document can be related.
- $SS_{(d, Kj)}$ is the similarity score for document d for the keyword K_j.
- $SS_{threshold}$ is the minimum similarity score acceptable for a document to relate to that keyword.

For example, if a given threshold is 9, we draw the conclusion that document 21 is irrelevant to keyword Data and the profile of document 21 is:

$DP_{(21)}$ = {(filtering, 12.59), (clustering, 11.59), (mean deviation, 9)}.

3.3   Customer Profile (CP)

In this approach, we define the CP as follows:

$$CP_{(n)} = \{ (K_i, D_i, T_{(Di, Ki)} \mid K_i \in C, \ 0 < i < N, 0 \leq T_{(Di, Ki)} < \infty) \} \tag{3}$$

where:

- n denotes a customer ID.
- K_i denotes a keyword or phrase
- C is a set of keywords or phrases to which the document can be related.
- D_i is the document whose ID is i.
- $T_{(Di, Ki)}$ is the time of user spent on D_i based on K_i as search keyword.

Equation 3 shows that a customer's level of interest in a particular document is based on the customer's searching and browsing history. The history includes: (i) the keywords or phrases the customer usually searches for, (ii) the document the customer hit in the search result, and (iii) the time that the customer spends on a particular document. A particular customer's level of interest is quantified with respect to those of the general users in accordance with the average customer profile (shown in Equation 4). As mentioned before, for a new customer using our search engine, the average customer profile (CPA) is adopted to him/her to start with.

$$CPA_{(n)} = \frac{CP_{(n)}}{1/N \sum_{n=1}^{N} CP_{(n)}(K_i, D_i, T_{(Di, Ki)})} \tag{4}$$

3.4   Personalized Search Results

When a customer performs search on the web, his/her browsing and searching behaviors can be captured by the web site. In our proposed approach, customer needs to login, so that his/her profile can be built up and updated. For a new customer or those customers who do not login, the CPA is applied. Personalized search result is then presented with the customer preferred rank based on the comparison of CP and DP.

When a document is retrieved from a search engine, a similarity is computed between the DP and CP. We apply the cosine similarity measure [8] to calculate the similarity between these two profiles. Each of CP and DP can be represented as an n-dimensional vector $[V_1, V_2, \ldots V_n]$.

DP contains a set of keywords where vector V_i represents the significance of the ith word in the vocabulary. One common way to compute the significance is to multiply the number of times the ith word appears in the document by the inverse document frequency of the ith word. The inverse of document frequency is one divided by the number of times the word appears in the entire document. We use all indexed documents in the search engine as the entire document. Here we only consider keywords in DP, because the similarity scores of keywords are only for filtering the keywords if they are above the given threshold. For example, the $DP_{(21)}$ in section 3.2 can also be presented as $DP_{(21)} = \{$filtering, clustering, mean deviation$\}$.

So does CP. CP contains a set of time that customer spent on a particular document based on a keyword as search query. The significance of $T_{(Di, Ki)}$ is to multiply the time

that customer spent on D_i by the inverse of total time that customer spent on all documents. Thus the similarity between two profiles is calculated as the inner product of the CP and DP vectors shown in the following formula.

$$S = \sum_{i=1}^{N} \frac{<V_{ic}V_{id}>}{|CP||DP|} \tag{5}$$

Where

$$CP = [\, V_{1c}, V_{2c}, \dots V_{nc}\,], \ DP = [\, V_{1d}, V_{2d}, \dots V_{nd}\,]$$

$$<V_{ic}V_{id}> = \sum_{i=1}^{n} V_{ic}V_{id}, \ \ |CP| = \sum_{i=1}^{n} \sqrt{V_{ic}^2}, \ \ |DP| = \sum_{i=1}^{n} \sqrt{V_{id}^2}$$

3.5 Evaluation

The quality of a search method necessarily requires human evaluation, especially for customized search results. Therefore, we carried out an evaluation experiment with 80 users who performed searches on the web (using a search engine) by entering search keywords. In this experiment, we used an existing well-known search engine to return the search results, because building up an indexing search engine is beyond this re-search. On average, 20 queries were entered by each user. The search engine returned search results without knowing the customer profile. Then these 80 users' browsing and search behaviors were captured and saved as customer history which was used to create customer profile. Meanwhile, the returned documents from the search engine were analyzed to generate document profile. After calculating the similarity between two profiles with formula (5), we returned the search result with our re-arranged rank in that the DP with higher similarity to CP was ranked higher. The non-personalized and personalized search results are evaluated by these 80 users. 48 of them are satisfied with the personalized search results and they comment that the new rank is more accurate than the non-personalized rank. Table 3 shows the feedback from these 80 users surveyed.

Table 3. Experimental result

Feedback	Number of users
Satisfy with personalized results	42
Do not mind the re-arranged rank	24
Do not feel the difference	7
Not satisfy	7

The result of evaluation indicates that 52.5% users are satisfied with this personalized search result. 30% users do not mind the new rank. Only 8.75% users are not satisfied with the new rank and prefer the non-personalized result.

4 Conclusions and Future Work

In this paper, we propose an approach to customize search results by comparing the customer profile and document profile. Customer profile is derived from users'

searching and browsing history. Document profile is created based on an algorithm which computes the similarity score of documents. We adopt a cosine method to calculate the similarity between document profile and customer profile. The evaluation is carried out by 80 users; majority of them prefer our proposed approach of customizing and re-ranking search results to the non-customized results.

Another aspect that we will be investigating in the future is that customer interests and preferences may change with time. Thus the customer feedback can be used to update the customer profile. The future work will also include carrying out the evaluation with a great number of users who have different background with different interests and tastes.

References

1. Meng, X and Chen, Z, "Personalized Web Search With Clusters", *International Conference on Internet Computing, IC'03*, 2003, pp. 46-52.
2. Salton, G., "Another look at automatic text-retrieval systems", *Communications of the ACM, 29(7)*, 1986, pp. 648-656.
3. Dublin Core -- http://dublincore.org/documents/dces/
4. J, A Hartigan. *Clustering Algorithms*. WILEY Publication, 1975.
5. Hearst, M. TileBars, "Visualization of Term Distribution Information in Full Text Information Access", in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'95)*, 1995, pp. 59-66.
6. Veerasamy, A. and Belkin, N.J., "Evaluation of a Tool for Visualization of Information Retrieval Results", in *Proceedings of the 19th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'96)*, 1996, pp. 85-92.
7. Hearst, M. and Pedersen, J. Reexamining the Cluster Hypothesis, "Scatter/Gather on Retrieval Results", in *Proceedings of the 19th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'96)*, 1996, pp.76-84.
8. Salton, G. Automatic Text Processing. Reading, MA: Addison-Wesley Publishing, 1989.
9. Shardanand, U. & Maes, P. "Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of CHI'95 Conference on Human Factors in Computing Systems,* ACM Press, 1995, pp. 210-217
10. Basu, C., Hirsh, H. & Cohen, W. "Recommendations as classification: Using social and content-based information in recommendation". In *Proceedings of AAAI-98, 1998, American Association for Artificial Intelligence*, 1998.
11. Weng, S. S & Liu, M. J. "Personalized product recommendation in E-Commerce" in *Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service,* 2004, pp.413-420
12. Terveen, L., Hill, W., Amento, B., McDonald, D. & Creter, J. Phoaks: "A system for sharing recommendations". *Communications of the ACM*, 40(3), 1997
13. Rucker, J. & Polanco, M.J. Sitesser: "Personalized navigation for the web". *Communications of the ACM*, 40(3), 1997.
14. J, Lai and B, Soh. "Using Element And Document Profile For Information Clustering", in *Proceedings of the 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service, 2004, pp. 503-506.*
15. S. Vrettos, A. Stafylopatis., "A Fuzzy Rule-Based Agent for Web Retrieval-Filtering", *Web Intelligence: Research and Development: First Asia-Pacific Conf.*, Springer-Verlag, October, 2001, pp. 448-453.
16. Baeza-Yates, R. and Ribeiro_neto, B., *Modern Information Retrieval*, Addison-Wesley, Reading, MA, 1999.

Improving Search on WWW.HR Web Directory by Introducing Ontologies

Gordan Gledec, Maja Matijašević, and Damir Jurić

Faculty of Electrical Engineering and Computing
University of Zagreb, Croatia
{maja.matijasevic,gordan.gledec,damir.juric}@fer.hr

Abstract. In this paper we propose ontology-based improvement of the Croatian Web directory search mechanism, which is currently not capable of executing queries that take into account the structure and semantics of user's query. The proposed approach is verified by introducing an ontology in the domain (directory category) of "tourism". We address three problems related to the search mechanism: low recall, high recall and low precision and vocabulary mismatch. The results show significant improvements in terms of subjective relevance and quality of results.

1 Introduction

The directory of Croatian Web servers, named WWW.HR, was first introduced a decade ago [1]. As the directory grew in terms of both the number and variety of sites included, it became increasingly difficult to find the information needed, especially for inexperienced users not familiar with formulating logical queries.

Recent study of user queries [2] on Croatian Web directory identified several problems related to the number and relevance of search results. These problems clearly indicate the need for new search mechanisms that will enable semantic search on the directory. In this paper, we address these problems by introducing ontology-based improvements into the search engine.

In the last few years, with the introduction of Semantic Web [3], a lot of research is under way in the field of ontology-based search [4], [5], [6]. Some visionary ideas on how Web search will evolve toward question answering systems can be found in [7].

Our research is inspired by Navigli and Velardi [4], who show that introducing other types of semantic information derivable from an ontology into the Web site search mechanism is effective at improving search results. To the best of our knowledge, this is the first study that analyses benefits of ontology-based search improvements related to the Croatian language and the Croatian Web space, however, the approach and method applied in this work are rather general and not language-specific.

The remainder of this paper is organised as follows. Section 2 explains the structure of Croatian Web directory and current search mechanisms. Section 3 describes the idea of ontology-based directory search. Section 4 explains how the ontology was implemented. Section 5 presents the results. Finally, Section 6 concludes the paper and gives references for future work.

2   WWW.HR Directory and Search

The WWW.HR is a Web-based information service supported by the Croatian Academic and Research Network – CARNet. Established in 1994, WWW.HR tends to be a thematic portal, providing general information concerning Croatia and a directory of Croatian Web sites. The directory is organized into a hierarchical structure similar to Yahoo (*http://www.yahoo.com/*) or DMOZ (*http://www.dmoz.org/*), starting from broad, more general categories, to narrower, more specific subjects, leading finally to a Web page that (presumably) has information on the subject. The top level of directory contains 14 categories, namely: About Croatia, Arts & Culture, Business & Economy, Computers & Networking, Education, Entertainment, Events, Law & Politics, News & Media, Organizations, Science & Research, Sport & Recreation, Society, and, Tourism & Traveling.

The directory is updated and expanded based on site submissions. New sites are submitted through Web interface. The following information about each site is recorded in the directory database and used by the master search index file:

– site name (in English and Croatian),
– site description (in English and Croatian),
– site URL,
– category name in which the site is submitted (in English and Croatian).

The master search index file also contains META keywords and description extracted from the Web page headers. When the user enters the query into the search form, the search index is queried and matching results returned to the user. The types of user queries supported are:

– all keywords entered (logical *and* between the terms, this is the default query),
– logical expressions (using *and*, *or*, *not*, + and - operators),
– keywords ending with wildcards (an asterisk representing any keyword suffix),
– phrases (any character string in quotation marks).

In order to gain knowledge of how users search our directory, we conducted the study of user queries which revealed that more than 60% of queries consisted of a single term, typically resulting in a rather large number (>20) matches. The type of query terms analysis revealed that 70% of all terms were nouns. Other problems discovered in our study may be summarized as:

(1) no recall: almost 30% of all queries returned no relevant results;

(2) low precision and high recall: almost 40% of queries returned 15 or more results;

(3) vocabulary mismatch: results of the query highly depend on the vocabulary (particular terms) the users use in formulating their query.

If a user enters a search term which is not used within either the title, or the description or the META tag(s) the query returns no matching results. If the query is too general, this results in high recall, so the user still has to scroll through a large amount of results to find the relevant one.

To improve the search in terms of semantics, we first had to build an ontology. We selected one of the top-level categories, namely "Tourism", as a case study, since Web log statistics show that this directory was most frequently accessed, either by browsing or searching.

3   Idea of Ontology-Based Directory Search

In "classical" search, the user enters the query term(s), and receives result(s) as a list of matching directory entries (Fig. 2, steps 1-2-9-10). The idea of enhancing this with the ontology-based search is illustrated in Fig. 2, steps 1 through 10. The user enters the query (1) into the search form, and submits it to the search engine (2). The engine consults the ontology (3) and extracts the relevant knowledge (4). The list of terms is extracted from the ontology and presented to the user (5) who can select the relevant terms (6) he wants to include in the modified query (7). Upon the submission of the modified query (8), the search engine retrieves results from the search index (9) and presented to the user (10).
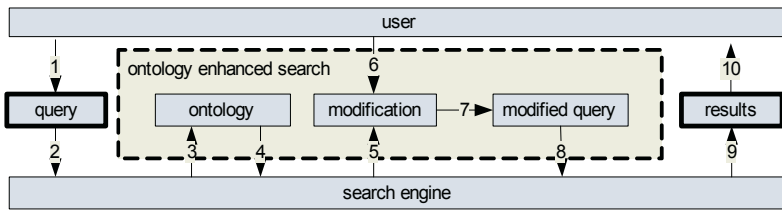
Fig. 1. The process of query modification using ontologies

In order to implement this solution, the ontology has to be built. Since the directory is rather general, and the user is a public user, this directory would be covered by many domain-related ontologies. The knowledge of user behaviour and expectations, gained from the study described in Section 2, was used as a guidance.

Building the ontology is the most difficult and time-consuming task, since no automated tools are used and the process involves only human knowledge. As such, the process of ontology creation can be split into three phases:

(a) domain consideration,
(b) planning the domain,
(c) coding the ontology.

Domain consideration involves decisions like what should be included as a part of the ontology and what should be omitted. Planning the domain requires consultation with relevant sources of information. In the case of directory based search, the already existing categorization and subcategorisation in the directory, as well as the analysis of the most frequently used keywords used to search the directory may be used. Still, certain parts of knowledge must come from the ontology creator, whose role is to synthesize knowledge from different sources, which is still possible until ontologies become rather small and domain specific. Coding the ontology in a tehnicaly acceptable format is the last step in the process of creating the ontology.

4   Implementation

Our ontology was created taking into account the existing categorization of the "Tourism" section and keywords that users use while searching the directory. The knowledge defined in the ontology is represented by subclasses of the query term, alternative terms or synonyms, upclasses and possible restrictions.

In the domain consideration phase we focus only on important concepts which are usualy described in terms of nouns. Many of these already exist as a part of the directory categories. The top-level directory of „Tourism" comprises of 14 categories with aditional subcategories under each of them: Accommodation, Food and Beverage, Cultural and Historical Monuments, Agencies and Associations, Gastronomy, Traveling and Transport, Guides, Web Cameras, Islands, Yachting, Nature, Spas, Places and Regions.

In the phase of planning the domain, statistics obtained from the Web server access logs were used to decide which concepts to include in the ontology. In our study, most of the main classes within the ontology were extracted from already existing categorisation in the directory. One superclass was created ("tourist"), and related to other classes by relations such as "to see", "to offer", "to give information about" or "located in", as shown in Fig. 1. For ontology coding, we used the W3C Ontology Web Language [9].
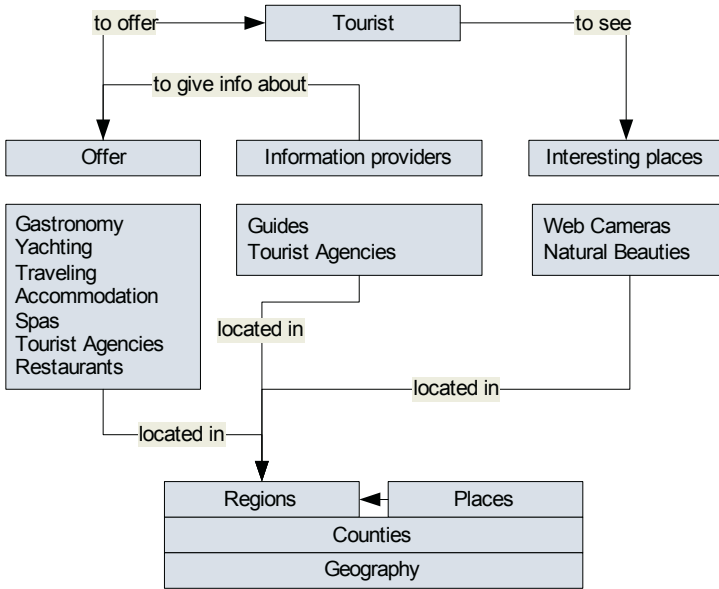
Fig. 2. Ontology "tourism"

This enhancement to the "classical" search was implemented by a set of custom made perl scripts.

Fig. 3. shows the user interface of the enhanced search page, containing the form presented to the user in step (5) on Fig. 2. The form includes options for the modifying the query by selecting alternative terms and synonyms.

5   Results

We addressed the problems associated with search mechanisms, low recall, high recall and low precision and vocabulary mismatch by conducting three types of tests.

Fig. 3. The form which enables the user to modify the query

The first test was designed to address the problem of low recall. The idea to approach this problem was to modify the original user query to include all subclasses of the query term. When the query was modified with all subclasses included, the search obviously retrieved more results than the original search. For example, searching for query *prirodne ljepote* (natural beauties) retrieved 43 results, but with subclasses like *nacionalni parkovi* (national parks) and *rezervat* (reservation) included, it retrieved 76 results (Fig. 4 – 1a). Another example is *autprijevoz* (car transport). The search for this term retrieves only 3 results, but when subclasses like autobusna agencija (bus agency), rent-a-car, *autobusni prijevoz* (bus transport), *kamionski prijevoz* (truck transport) and similar are included in the search,. the search engine returns 56 results (Fig. 4 – 1b).

The second test aimed at addressing the problem of high recall and low precision based on restricting the original query. Searching with restrictions resulted in less matches because of specific requirements (eg. location), so only pages matching these requirements were taken into account. An example is the query for the term *aerodrom* (airport), which retrieved 9 results, but the ontology found restriction for location and offered a choice of cities (Zagreb, Zadar, Split) to the user, so the number of matching results can be much smaller (Fig. 4 – 2b). Another example is the search for the term *apartman* (appartment) – the search engine returns hundreds of results, but with restrictions that include the keyword *kuhinja* (kitchen) the number of retrieved results is restricted to managable 11 (Fig. 4 – 2b).

The third test addressed the problem of vocabulary mismatch by modifying the original query by including synonyms and alternative words extracted from the

ontology. Croatian language is very rich in synonyms and dialects. When synonims or alternative words were included in our ontology, the search retrieved more relevant results than the original query, as expected. For example, the word "car" has several forms in Croatian language: *automobil* (automobile), *auto* (auto), *vozilo* (vehicle), to name just a few. The search returned 43 results for the keyword *automobil*. When the query was modified to include all three forms, the number of retrieved results was 100 (Fig. 4 – 1a). The query for the word *dvorac* (castle) returned 11 results, but when it was modified with words that have similar meanings - *kula* (tower), *palača* (palace), *zamak* (court), *utvrda* (fortress), *dvor* (court) - the number of matching results grew to 41 when all keywords were selected for the final query (Fig. 4 – 3b).
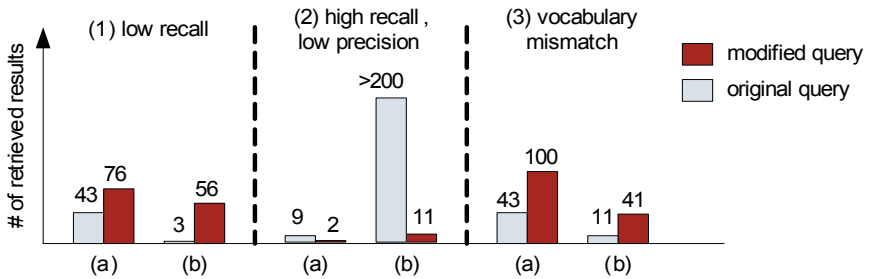
Fig. 4. The comparison between "classical" (string-based) and ontology-based search

6 Conclusions and Future Work

The results may be summarized as follows. The query modifications that involved the inclusion of synonyms and subclasses returned many more relevant results than the original search, and the modification with restrictions returned fewer of results, but with greater relevance. (The relevance of the retrieved results throughout this paper was assessed subjectively.)

It may also be noted that combining the proposed methods of query modifications may increase search efficiency. The query can be modified by first expanding it with the synonyms and then restricting it to give more focused results. For example, if the original query searches for the term *aerodrom* (airport), it could first be expanded by adding synonym *zracna luka*, and then restriction to a preferable location, for example Split, could be applied to narrow the number of matching results.

Our future work will focus on developing ontologies for other frequently accessed parts of the directory. Also, we need to improve a rather rudimentary user interface to enable users to modify their queries in a more usable fashion. Additionally, we need to perform a more formal user survey to assess the relevance of results retrieved through ontology-based search.

References

1. Jurić, J., Matijašević, M., Mikuc, M. First experiences with the World-Wide-Web in Croatia, Proceedings of the 2nd International Conference on Telecommunication Information Systems, TIS '94, pp. 90-95, Žilina, Slovakia (1994)

2. Gledec, G., Ljubi, I. How Users Search WWW.HR Web Directory? Proceedings of the 6th CARNet Users Conference, CD-ROM edition. Zagreb (2004)
3. Berners-Lee, T., Hendler, J., Lassila, O. The Semantic Web, Scientific American, vol. 284, No. 5. pp. 34-43 (2001)
4. Navigli, R., Velardi, P. An Analysis of Ontology-based Query Expansion Strategies, Proceedings of the 14th European Conference on Machine Learning (ECML 2003), Cavtat-Dubrovnik, Croatia (2003)
5. Ding, L., Finin, T., Joshi, A., Peng, Y., Scott Cost, R., Sachs, J., Pan, R., Reddivari, P., Doshi, V. Swoogle: A Semantic Web Search and Metadata Engine, Proceedings of the 13th ACM Conference on Information and Knowledge Management (CIKM'04), Washington DC (2004)
6. Modica, G., Gal, A., Jamil, H. M.; The Use of Machine-Generated Ontologies in Dynamic Information Seeking, Proceedings of the 6th International Conference on Cooperative Information Systems (CoopIS 2001), Springer-Verlag LNCS series, September 5-7, Trento, Italy (2001)
7. Zadeh, A.L.: Web Intelligence, World Knowledge and Fuzzy Logic – The Concept of Web IQ (WIQ), Proceedings of the 8th International Conference KES 2004, Wellington, New Zealand (2004)
8. Antoniou, G., van Harmelen F.: A Semantic Web Primer. The MIT Press (2004)
9. OWL - Web Ontology Language Reference, URL: http://www.w3.org/TR/owl-ref/ (2004)

Designing a Tool for Configuring an Intelligent and Flexible Web-Based System

Diego Magro and Anna Goy

Dipartimento di Informatica, Università di Torino
Corso Svizzera 185, 10149 Torino, Italy
{magro,goy}@di.unito.it

Abstract. Several *intelligent* Web-based systems need complex reasoning mechanisms and problem solving capabilities in order to perform their tasks. These systems should protect their users against the complexity of their inference engine. Such a protection should be offered both to the final users and to the domain experts that instantiate the Web systems on the different particular domains. We claim that to this purpose an intermediate representation layer, with the role of filling the gap between the implementation-oriented view of the domain (needed by the reasoning module) and the human-oriented view of the same domain (suitable both for the final users and for the domain experts), is required. This intermediate layer also provides a representation which is independent from the specific reasoning approach adopted. In the paper we discuss these issues by referring to a specific example, i.e. the STAR system (a Web-based system which exploits a configuration engine to support the user in organizing a personalized tour) together with STAR-IT, a tool that supports the instantiation of STAR on different domains.

1   Introduction

People surfing the Web require an active and intelligent support, even in complex tasks (e.g., the Web should provide financial or medical advice, leisure and entertainment offers, travel plans and so on). For this reason, Web-based systems need to be equipped with complex reasoning mechanisms and problem solving capabilities.

To avoid ad hoc solutions, the exploitation of existing problem solvers is recommended, also to ensure a certain degree of generality and flexibility which guarantees the possibility to configure (in the following we will use the term *instantiate*) the system on different domains.

Current systems exploiting problem solving techniques (e.g., configuration systems [1, 2]) are typically designed for expert users, and they are based on an implementation-oriented representation of the domain. However, when these services are offered on the Web, the interaction needs of typical Web surfers, which include a large number of non-expert users requiring a natural and user-friendly interaction, cannot be ignored.

This scenario poses three main challenges to be faced when designing and developing an intelligent Web-based system that supports the user in a complex problem solving task:

- Non-expert users should not be exposed to the complexity of the knowledge representation and of the reasoning mechanisms used by the problem solver. In fact, both the knowledge representation and the reasoning mechanisms could be quite complex, expecially if the system exploits an existing problem solver (a configurator, a planner, or whatever), which was designed to be general with respect to the specific application domain.
- The system should not be committed to any specific approach, i.e., it should be designed in order to work with different problem solvers of the same type (e.g. different planners, possibly representing the domain knowledge in different ways), as well as with different types of problem solvers, since a same complex task may have more than one sound formalizations (e.g., some tasks may be formalized both as planning problems or as a scheduling problems). In other words, the reasoning engine should be wrapped within an independent module, in order to support the interoperability of the system with other modules, implementing different reasoning approaches.
- Another important aspect concerns the instantiation of the Web system on a particular domain and its maintenance. Such tasks are usually performed by a domain expert who is not aware of the reasoning techniques nor of the particular knowledge representation used by the problem solver. Thus, a tool supporting the domain expert in these tasks is needed.

We claim that this three goals require an intermediate representation layer, with the role of filling the gap between the implementation-oriented view of the domain, needed by the reasoning module, and the human-oriented view of the same domain, needed by the frontend of both the instantiation tool and the final system. This intermediate layer also provides a representation which is independent from the specific reasoning approach adopted. In other words, such a layer will guarantee both the tranparency of the reasoning mechanisms for the user (domain expert and final user) and the independency of the frontend of the system (intantiation tool and final system) from the specific reasoning techniques adopted.

The design of such a layer requires a domain ontology, based on the concepts used by people to reason about that domain, coupled with a translation module, that provides the mapping between these concepts and the representation needed by the actual reasoning engine.

In the following we will discuss in more detail this solution by referring to a specific example, i.e. STAR (Smart Tourist Agenda Recommender), a Web-based system which exploits a configuration engine to support the user in organizing a personalized tour in a given geographical area. In particular, we will discuss the main functionality of STAR-IT, a tool that supports the instantiation of STAR on different domains.

In Section 2 we introduce STAR, while in Section 3 we describe the istantiation tool; Section 4 concludes the paper.

2   STAR

In the travel and tourist domain different systems have been developed to support the users to plan long trips, involving flights scheduling and hotel reservation, or even to

select a set of tourist activities (e.g. [3–5]). However, in such systems the task of organizing a coherent agenda is totally left to the user.

The organization of a tourist agenda is a task that people are quite used to deal with. However, accomplishing such a task in a satisfactory way can be complex and time consuming, since it requires to access different information sources (tourist guides, the Web, local newspapers, etc.), to consider various kinds of constraints (the available time, the physical locations in which the different activities take place, and so on), and to mentally backtrack many times (e.g. when discovering that an interesting museum is closed, that two historical buildings are too far from each other to be visited in the same morning). Within our prototype STAR [6], we have tried to automate the task of organizing a tourist agenda by defining it as a configuration problem and exploiting a configuration engine based on the $\mathcal{FPC}$ framework [7], to find a solution that fulfils user requirements.

Intuitively, a configuration problem [8] is the task of assembling a (sub)set of predefined components in order to build a complex entity that meets a set of requirements. The component types are given a priori and no new type can be introduced during the configuration process. Usually, a set of constraints define the valid combinations of components. These constraints may restrict the types or the number of components allowed as parts of the complex entity, or they may express more complex relations among the components.

STAR's architecture currently includes three main modules: the Frontend, the Backend and the Knowledge Base. The Frontend consists of two components: the User Interface exploits XML/XSL technology to dynamically generate HTML Web pages, displayed on the browser; the Frontend Manager is a Java Servlet that takes the role of a mediator between the Backend and the interaction with the user.

The Knowledge Base contains a $\mathcal{FPC}$ declarative representation of a generic agenda as a complex (i.e. configurable) entity in which tourist items (e.g. buildings, restaurants, etc.) are the basic components and denote the corresponding activities (e.g. visits to buildings, lunch in a restaurant, etc.).

The Backend encompasses three submodules: a $\mathcal{FPC}$ configurator (the reasoning engine); the translator of the user's requirements into a set of $\mathcal{FPC}$ (input) constraints; and the Backend Manager, which handles the interaction with the Frontend and coordinates the activities of the other two Backend submodules. All these submodules are implemented in Java.

STAR's reasoning engine takes as input both the knowledge base and the input constraints (representing the user requirements) in order to configure a suitable agenda to suggest to the tourist (i.e. the final user of the system).

3   STAR-IT

STAR-IT is a tool supporting the domain expert in instantiating STAR on particular domains. Two STAR applications may differ with respect to various dimensions, which represent the degrees of freedom of the instantiation activity. Here, we focus on those dimensions that are most closely related to the problem solving capability: the description of the general structure of an agenda and the definition of the properties that the final user can impose as requirements.
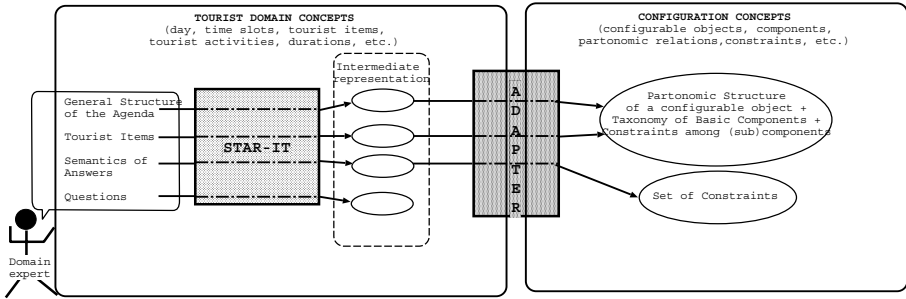
Fig. 1. Instantiation Process

We will show how the domain expert can instantiate a system like STAR without being a knowledge engineer, and in a way that is totally independent from the specific inference engine actually exploited by the system. Figure 1 depicts the instantiation process.

Thanks to STAR-IT, the domain expert can describe the general structure of an agenda and its possible content by means of common concepts, which include: the time slots of the day (e.g. morning, afternoon, etc.); the tourist items (e.g. buildings, restaurants, events, etc.) that are involved in the tourist activities (e.g. visiting buildings, eating, attending events, etc.); the possible durations of the activities; the daily opening hours and the actual tourist attractions availability in particular days; the distance between the physical locations in which the activities take place.

STAR-IT produces an intermediate representation based on these concepts, which expresses the user-oriented view of the domain and is independent from the specific reasoner. A module, called Adapter, plays the role of a bridge between the problem solver and the rest of the system. Such module takes as input the intermediate representation and it produces a knowledge base in the problem solver format (therefore the Adapter depends on the particular problem solver). In the current prototipe, a $\mathcal{FPC}$ general configurator plays the role of the problem solver and the Adapter translates the intermediate description of the general agenda into a $\mathcal{FPC}$ knowledge base. Such a knowledge base is expressed by means of the usual configuration concepts: configurable and basic components and subcomponents, partonomic relations, constraints among (sub)components and so on. The domain expert can be unaware both of the configuration ontology and of the ways the configuration concepts are actually represented within the configurator.

In the following we show a fragment of the internal representation, based on an XML language, adopted by STAR-IT at the intermediate layer:

```
<DAILY_AGENDA>
  <TIME_SLOTS>
    <TIME_SLOT>
      <TS_NAME>Morning</TS_NAME>
      <TS_MAX_DURATION>12</TS_MAX_DURATION>
      <TS_PREFERRED_DURATION>10</TS_PREFERRED_DURATION>
      <ACTIVITIES>
        <ACTIVITY>
```

```
        <A_NAME>havingBreakfast</A_NAME>
        <A_MIN_NUM>0<A_MIN_NUM> <A_MAX_NUM>1</A_MAX_NUM>
        <A_ITEM_TYPE>Bar</A_ITEM_TYPE>
        ...
      </ACTIVITY>
      <ACTIVITY>
        <A_NAME>visitingAttractions</A_NAME>
        <A_MIN_NUM>0<A_MIN_NUM> <A_MAX_NUM>10</A_MAX_NUM>
        <A_ITEM_TYPE>TouristAttraction</A_ITEM_TYPE>
        ...
      </ACTIVITY>
      <ACTIVITY> <A_NAME>havingLunch</A_NAME>...</ACTIVITY>
      <ACTIVITY> <A_NAME>doingShopping</A_NAME>...</ACTIVITY>
      <ACTIVITY> <A_NAME>attendingEvent</A_NAME>...</ACTIVITY>
    </ACTIVITIES>
  </TIME_SLOT>
  <TIME_SLOT> <TS_NAME>Afternoon</TS_NAME>...</TIME_SLOT>
  <TIME_SLOT> <TS_NAME>Evening</TS_NAME>...</TIME_SLOT>
  </TIME_SLOTS>
  ...
</DAILY_AGENDA>
```

The part enclosed in the DAILY_AGENDA tags describes the general structure of a one-day agenda. This fragment, for instance, says that a day is partitioned in three time slots (Morning, Afternoon and Evening). The morning should better not exceed 10 time units (the domain expert can define her own time unit; in this example, each time unit roughly corresponds to half an hour) and it must not be longer than 12 time units. The set of all the possible activities is specified for each time slot of the day; e.g. in the morning a tourist can have breakfast, visit up to 10 tourist attractions and so on. The tourist items that can be involved in the activities are described in the section enclosed between the AVAILABLE_TOURIST_ITEMS tags and they are organized into a taxonomy, as shown in the following:

```
<AVAILABLE_TOURIST_ITEMS>
  <CLASSES>
    <CLASS>
      <C_NAME>Bar</C_NAME> <SUBCLASS_OF>FoodPlace</SUBCLASS_OF>
      ...
    </CLASS>
    ...
  </CLASSES>
  <ELEMENTS>
    <ELEMENT>
      <E_NAME>HappyBar</E_NAME> <IS_A>Bar</IS_A>
    </ELEMENT>
  <ELEMENTS>
<AVAILABLE_TOURIST_ITEMS>
```

The Adapter translates this user-oriented representation into the representation required by the configurator. In particular, in the current prototype, it builds a $\mathcal{FPC}$

partonomy representing the configurable daily agenda which, in this example, has three configurable direct components corresponding to the morning, the afternoon and the evening. These components have a set of partonomic roles representing the allowed activities and that can be filled with the tourist items, which are the basic components of the configuration. Moreover, a set of $\mathcal{FPC}$ constraints restrict the set of valid configurations; e.g. the maximum and the preferred duration of the morning is coded by two weighted constraints (one with $k = 10$ and $w \neq \infty$ and the other with $k = 12$ and $w = \infty$), as follows:

$\Sigma(\{\langle havingBreakfast, duration\rangle, \langle visitingAttractions, duration\rangle,$
$\langle havingLunch, duration\rangle, \langle doingShopping, duration\rangle,$
$\langle attendingEvent, duration\rangle\}) \leq k, weight = w$

STAR-IT supports the domain expert also in defining the possible (final) user's requirements on a tourist agenda, i.e. which properties of the agenda are influenced by user's preferences and needs (e.g. having/not having breakfast, visiting historical museums, etc.), expressed in an initial interview (Web form) that STAR proposes to the user.

STAR-IT enables the domain expert to list the set of questions that STAR will ask to the final user, along with a finite set of possible answers. Each answer represents a user requirement whose meaning is specified during the instantiation phase.

To express the semantics of each answer, the domain expert has to identify which time slots, which activities, and which kinds of tourist items the requirement refers to; moreover, she has to specify the actual content of the requirement. For instance, in the following we show the intermediate representation of a requirement stating that the tourist is interested to visit many baroque buildings either in the morning or in the afternoon.

```
<REQUIREMENT>
  <TIME_SLOTS>
    <TIME_SLOT>Morning</TIME_SLOT>
    <TIME_SLOT>Afternoon</TIME_SLOT>
  </TIME_SLOTS>
  <ACTIVITIES>
    <ACTIVITY>visitingAttractions</ACTIVITY>
  </ACTIVITIES>
  <TOURIST_ITEMS>
    <ITEM_TYPE>
      <CLASS>Building</CLASS>
      <RESTRICTIONS>
        <RESTRICTION>
          <ATTRIBUTE>Style</ATTRIBUTE> <VALUE>Baroque</VALUE>
        <RESTRICTION>
      </RESTRICTIONS>
    </ITEM_TYPE>
  </TOURIST_ITEMS>
  <CONTENT>many</CONTENT>
</REQUIREMENT>
```

In the current prototype, the intermediate representation of each requirement is translated into a set of $\mathcal{FPC}$ constraints.

4   Conclusions

In this paper we have presented STAR-IT, a tool supporting the instantiation of STAR (a Web-based system supporting the user in organizing a personalized tourist agenda) on new domains, and we have discussed how STAR-IT faces two main challenges: (i) it hides the technological complexity to the user (i.e. the domain expert), who needs not to be aware of the actual inference mechanisms underlying the system; (ii) it is not committed to a specific reasoning module, therefore modules based on different inference techniques can be easily plugged in the system. In particular, we have shown how an intermediate representation layer, representing the human-oriented view of the domain, can be exploited to achieve these goals.

A first prototype of STAR-IT has been implemented (as a Web-based Java application) and it has been used to instantiate a STAR application on the city of Turin. The current prototype is still only a research prototype and feedback from real users is needed in order to enhance the usability of the tool and to refine the approach.

References

1. Mailharro, D.: A classification and constraint-based framework for configuration. AI EDAM **12** (1998) 383–397
2. Fleischanderl, G., Friedrich, G.E., Haselböck, A., Schreiner, H., Stumptner, M.: Configuring large systems using generative constraint satisfaction. IEEE Intelligent Systems (1998) 59–68
3. Torrens, M., Faltings, B., Pu, P.: *SmartClients*: Constraint satisfaction as a paradigm for scaleable intelligent information systems. Int. Journal of Constraints **7** (2002) 49–69
4. C.Knoblock: Building software agents for planning, monitoring, and optimizing travel. In: Proc. of ENTER 2004. (2004)
5. Ricci, F., Arslan, B., Mirzadeh, N., Venturini, A.: Itr: a case-based travel advisory system. In: Proc. of ECCBR 2002. (2002) 613–627
6. Goy, A., Magro, D.: Dynamic configuration of a personalized tourist agenda. In: Proc. of IADIS Int. Conf. WWW/Internet 2004. (2004) 619–626
7. Magro, D., Torasso, P.: Decomposition strategies for configuration problems. AIEDAM, Special Issue on Configuration **17** (2003) 51–73
8. Sabin, D., Weigel, R.: Product configuration frameworks - a survey. IEEE Intelligent Systems (1998) 42–49

Location-Sensitive Tour Guide Services Using the Semantic Web

Jong-Woo Kim[1], Ju-Yeon Kim[1], Hyun-Suk Hwang[2], and Chang-Soo Kim[3]

[1] Pukyong National University
Interdisciplinary Program of Information Security, Korea
{jwkim73,jykim}@mail1.pknu.ac.kr
[2] Pukyong National University, Institute of Engineering Research, Korea
hhs@mail1.pknu.ac.kr
[3] Pukyong National University, Dept. of Computer Science, Korea
cskim@pknu.ac.kr

Abstract. The tour guide services are a LBS application that provides the information based on the location of users. The recent researches of the tour guide services focus on the context-sensitive computing, but the tour guide services based on HTML contents can provide only the static information. Therefore, the goals of our research are to semantically search the information. For the purpose, we describe the ontologies of the tour guide services, and express the tour information as the RDF contents based on the ontologies. In this paper, we present the architecture of the tour information Services based on the Semantic Web technologies and describe the process of the implementation.

1 Introduction

The Location Based Services (LBS) provide the context sensitive information based on the mobile users' location. The LBS includes the services such as local maps, local weather, traffic condition, tour guides, and shopping guides. For example, travellers can search not only the location of hotels and ATM machines near by user's current location but also the addition information using the LBS system when they visit a city first. Therefore, LBS provides different results as the users' position even though users request the same services[11]. Tour guide services are one of the most useful services. The tour guide services provide the additional information such as pictures and detail descriptions of users' interesting place as well as simple location information. Especially, the researches of context sensitive computing has recently increased to provide the adaptive information to users' situation. The objective of the context sensitive tour guide services is to provide the different information as the location moving, traffic situation, weather and cost even though users request the same information[2, 8, 10].

In this paper, we focus on how we get the needed information of users. The existing tour guide services have provided static information and must store the information as the situation estimated in advance because the existing researchers are interested in providing the information. Therefore, we try to apply

the technologies of the Semantic Web to tour guide services to generate more dynamic information as the requirements of users when creating the information. The Semantic Web[1] is a technology to add information on Web to well-defined meaning, to enable computers as well as people to understand the meaning of the documents easily, and to automate the works such as information search, interpretation, and integration.

We describe the Semantic information of the tour guide services using the technologies of the Semantic Web and propose a method which can semantically search the information for the requirement of users. This paper is organized in the following manner. In section 2, we present related work, followed by overview of tour guide services and the Semantic Web. In section 3, we present the architecture of tour guide services based on the Semantic Web and implement a simple prototype to realize the architecture in section 4. The final section summarizes and presents the future work.

2   Related Work

There have been a number of researches to provide context sensitive Tour Guide. Abowd et al.[2] has developed the Cyberguide, which is handheld electronic tourist guide system that supplies the user with information based on user's location. This system provides the following components: a map component for cartographer services, information component for librarian services, positioning component for navigator service and communication component for messenger service. Davis et al.[8] has an ongoing project GUIDE to investigate electronic tourist guides in a practical real-world environment. They have been building and testing different versions of electronic tourist guides for the city of Lancaster over the past few years. Their current approach uses wireless communication on a pen-based tablet computer. They intended to consider not only the location but also the visitor's interests, the city's attractions, mobility constraints, available time, weather and cost as the context for their context sensitive tour guide system. Simcock et al.[10] focus on the software support for location-based applications. They are not just interested in the location but also other elements, attractions and equipment near by. Their main aim is to develop a context sensitive travel expo application. The application consists of a map mode, a guide mode, and an attraction mode. It provides map service and tourism information service through HTML-based sound, images, and text.

The recent researches in Tour Guide Services focus on the context sensitive computing. Although these systems provide context sensitive Tour Guide information based on location through HTML-based content, there still remain a challenge to overcome the limitations of their approache, static information.

The Semantic Web[1] is a technology to add information on Web to well-defined meaning, to enable computers as well as people to understand meaning of the documents easily, and to automate the works such as information search, interpretation, and integration. The automated agents and dedicate search engines can be a high level of automation and intelligence because the Semantic Web documents have a meaning that a computer can interpret, contrary to exist-

ing HTML-based Web documentations. The Semantic Web consists of Uniform Resource Identifier (URI), UNICODE, Resource Description Framework (RDF), RDF Schema, and Ontology hierarchically. The RDF is an XML-based language to describe resources like images, audio files, or concepts available via the Web. The RDF model is called as a triple because it has three parts, subject, predicate, and object. The subject and object means a resource, and the predicate defines their relation. The ontology defines the common words and concepts used to describe and represent an area of knowledge. The Semantic Web documents must be represented as the RDF which defines common data of special domains based on an ontology which represents data relation. The contents on this structure can be searched, integrated, and inferred semantically.

The systems using the Semantic Web have been developed in practical fields, and its effectiveness was also presented. Tanen et al.[9] have developed the Courseware Watchdog which is a comprehensive approach for supporting the learning need of individual in the environment of fast change using Ontology. Lee et al.[7] presented an architecture how a service flow can be automatically composed by syntactic, semantic, and pragmatic knowledge in medical services called "Semantic Medical Services(SMS)". Hunter et al.[3] presented a Web-services-based system which they have developed to enable organizations to semi-automatically preserve their digital collections using ontologies.

3   Semantic Web Based Tour Guide Services

In this section, we use the technologies of the Semantic Web to provide the Semantic information for the requirement of users as expressing and searching the information of tour guide services semantically. In our previous work, we defined an ontology for hotel information and provided the information using the method which is to crawl the required information of users from HTML-based contents[5]. However, this HTML-based contents had a problem that it can not completely search the semantic information. Therefore, we use RDF(Resource Description Framework) which is a language designed to the Semantic Web to express tour contents. And we design the structure of the Tour Guide System for supporting the semantic information. We use an ontology of hotel domain to describe the semantic contents.

The tour guide services provide the tour information and map services for touring as shown in Figure 1. The users input the search values including location, room type, meal type, and the range of the price through the Web browser and request the results of the search to the tour guide server (Figure 1(a)). Next, The tour guide server executes the requested search and shows the lists of hotels agreed with the condition (Figure 1(b)). The map service provides the location information of the hotel if they select a hotel item in the provided results (Figure 1(c)).

The tour guide system includes the three components such as the Tour Guide Client, the Guide Server, and the Tour Contents Server(Figure 2). First, we define the ontology of the hotel domain, and the Tour Contents Server generates the RDF contents as the defined hotel ontology. The Guide Agent of the Guide
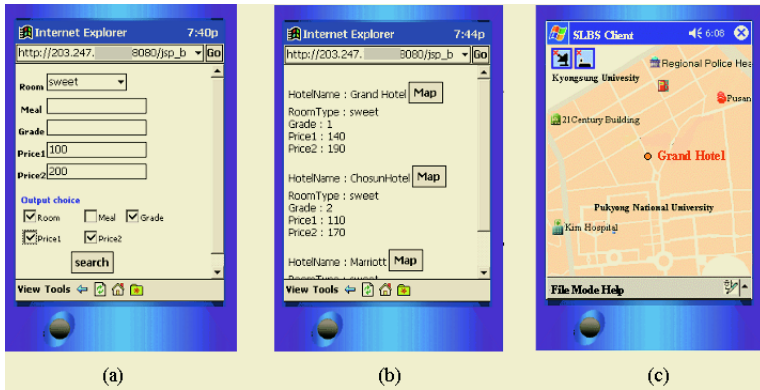
Fig. 1. User interface of Location Sensitive Tour Guide using Semantic Web.
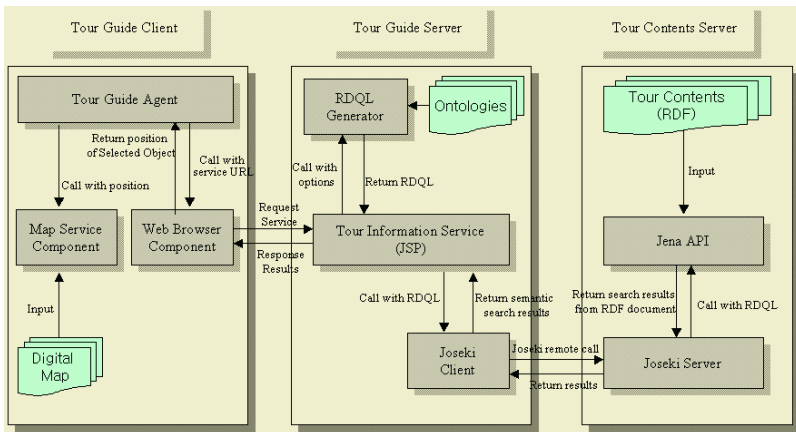
Fig. 2. Architecture of Semantic Web Based Tour Guide System.

Client attains the location of the users and provides the interface for guiding the tourists using the map service component and Web browser component. The map service component uses the mobile GIS model developed in our previous work[6] to provide the map services. We proposed the map reduction that is a method to provide the map services, and the Simple Spatial data Format(SSF) that is an efficient digital map format. The Tour Guide Agent uses the Web browser component to search the requested information of the users. The Tour Information Service of the Tour Guide Server executes the practical searching to find semantic information. The Web browser is an interface to transmit and attain the query for the searching.

The Tour Information Service(TIS) generates the query statements of the Resource Description Query Language(RDQL) to search RDF documents with the users' search condition inputted from the Tour Guide Client. Next, the TIS
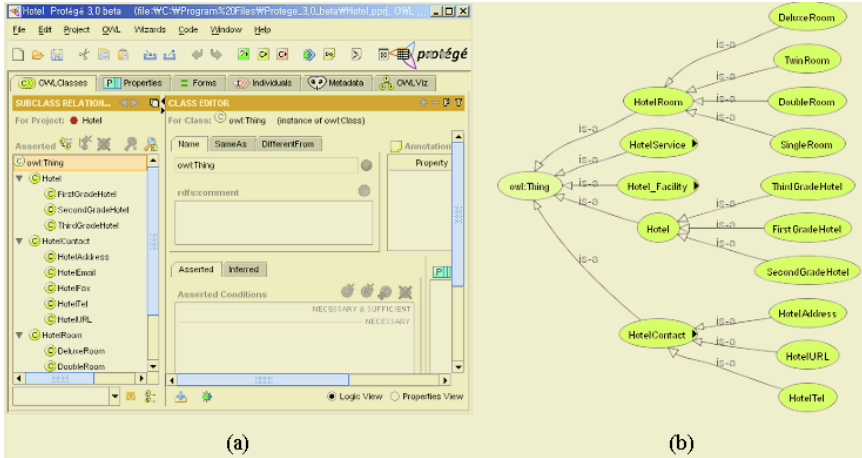
Fig. 3. Ontology modeling example in hotel domain. (a) Protege-2000, (b) OWLViz.

searches the RDF contents stored in the Tour contents Server by the Joseki Client/Server , which is a API for supporting the remote query, and send the search results to the Tour Guide Client. We will explain the detail method for describing and searching the tour information in section 4.

4   Realizing Tour Information Services

To provide tour information, we use two steps: expressing the semantic contents and searching the contents. We use the hotel domain to realize the two steps for the tour information services.

First, we need to define ontologies for a specific domain and generate RDF contents for real instances based on the defined ontologies. We use a Protege-2000 tool to create an OWL ontology model of hotel domain. The Web Ontology Language(OWL) is a standard ontology language from the World Wide Web Consortium(W3C). Protege-2000[1] is an integrated software tool used by system developers and domain experts to develop knowledge-based systems. The applications developed with protege-2000 are used in the problem-solving and the decision-making in a particular domain. Protege-2000 is an ontology editor and a knowledge-based editor and provide the OWL plug-in , which is a function for modeling the objects and translating to OWL type. An example of the hotel ontology created with the Protege-2000 OWL plug-in[2] is shown in Figure 3(a) and a View of OWLViz[3] is shown in Figure 3(b).

We define the classes with inheritance relations in OWL classes tab, and defined data properties of defined classes and relations between classes in prop-

---

[1] http://protege.stanford.edu/

[2] http://protege.stanford.edu/plugins/owl/index.html

[3] http://www.co-ode.org/downloads/owlviz/
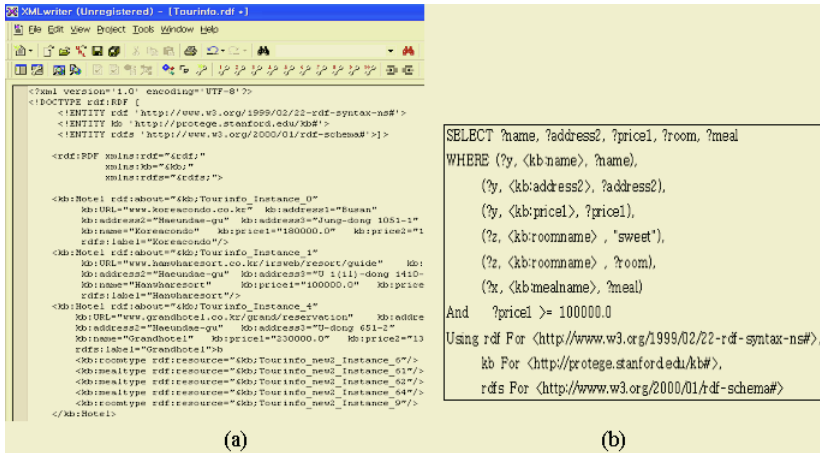
Fig. 4. Example of instances(RDF) and RDQL query for semantic search.

erties tab (Figure 3(a)). we use the OWLViz to express the hierarchies of the defined hotel ontology (Figure 3(b)). Also, We created the RDF contents after inputting the instances in the individual tab(Figure 4(a)).

We use the Jena Toolkit[4] to search only requested information of users from the RDF-based contents like the Figure 3. The Jena Toolkit makes it easier to develop applications that use the Semantic Web information model and languages. Jena Architecture's heart is the RDF API, which supports the creation, mainpulation, and query of RDF documents. Joseki is a Java client and server that implements the Jena network API over HTTP[4]. We can semantically search the instances of RDF documents through RDQL, a Qurey Language for RDF. We use JSP(Java Server Page) to provide user interface for information service. It use RDQL Generator(Java Class) which is used to create user's input RDQL and Joseki Client API which is used to semantically search the tour information. Figure 4(a) shows RDF Contents generated based on the defined ontology(Figure 3), and Figure 4(b) shows an example of the RDQL query statement created using the RDQL generator.

5 Conclusion

In this paper, we use the Semantic Web technologies to improve the problem of the existing tour guide services, which can provide only the defined format of information. We present the architecture of the Tour Guide Services and describe the process of the implementation. First, we defined the ontologies of the Tour information and generated RDF contents based on the defined ontologies. Next, we executed the semantic search using Jena API and Joseki Client/Server API. The interface between the Mobile Client and the Tour Guide Server is executed by JSP. We use our previous work to do the map services for Mobile GIS. We

implemented a simple prototype of the hotel domain and showed the efficiency of the Tour Information Service based on the Semantic Web technologies. The advantages of the Tour Guide Services using the Semantic Web technologies include (1) the exact tour information (2) the interoperability between the server systems with the contents on the various platform (3) the user-tailored tour information. However, it is required to define the standard ontologies for the Tour Guide Services to provide the inference function based on the well-defined ontologies.

Acknowledgement

References

1. Berners-Lee, T., Hendler, J. and Lassila, O.: The Semantic Web. Scientific American (2001)
2. G. Abowd, C. Atkeson, J. Hong, S. Long, R. Kooper, M. Pinkerton, G. Marchionini, Cyberguide: A Mobile Context-Aware Tour. Technical Report GIT-96-06, Georgia Institute of Technology, (1997)
3. Hunter, J., Choudhury, S.: A Semi-Automated Digital Preservation System based on Semantic Web Services. JCDL'04, ACM (2004)
4. Jeremy J. Carroll, Lan Dickinson, and Chris Dollin.: Jena: Implementing the Semantic Web Recommendations. ACM (2004)
5. Kim, J.W., Kim, C. S., Gautam, A., Lee, Y.: Location-Based Tour Guide System Using Mobile GIS and Web Crawling. Lecture Notes in Computer Science, Vol.3428. Springer-Verlag, Berlin Heidelberg (2005)
6. Kim, J. W. Park, S. S. Kim, C. S. Lee, Y.: The Efficient Web-Based Mobile GIS Service System through Reduction of Digital Map. Lecture Notes in Computer Science, Vol.3043. Springer-Verlag, Berlin Heidelberg (2004)
7. Lee, Y., Patel, C., Chun, S. A., Geller, J.: Compositional Knowledge Management for Medical Services on Semantic Web. WWW2004. ACM (2004)
8. N. Davis, K. Cheverst, K. Mitchell, A. Efrat, Using and Determining Location in a Context-Sensitive Tour Guide. Cdomputer, Volume 34, Issue 8, IEEE, (2001) 35-41
9. Tane, J. Schmitz, C., Stumme, G.: Semantic Resource Management for the Web: An E-Learning Application. WWW2004, ACM (2004)
10. T. Simcock, S, Hillenbrand, B. Thomas, Developing a Location Based Tourist Guide Application. Proceedings of the Australasian Information Security Workshop conference on ACSW Frontiers 2003, Australian Computer Society, (2003)
11. Shiowyang Wu, Kun-Ta Wu. : Dynamic Data Management for Location Based Services in Mobile Environments. Proc. 2003 Int. Conf. Database Engineering and Applications symposium. IEEE (2003)

Semantic Discovery of Web Services

Hongen Lu

Department of Computer Science and Computer Engineering
La Trobe University
Bundoora, Melbourne
VIC 3086, Australia
helu@cs.latrobe.edu.au

Abstract. Web service is the next ware of Internet computing. Discovery web services is becoming a new challenge due to the increasing number of available services on the World Wide Web. In this paper, I investigate the semantics discovery of Web services based on domain ontology. Multiple service discovery strategies are defined and explained. The semantics service discovery methods provide a new way to locate and utilize published Web services. The approach is flexible and extensible to accomplish complex web service requests.

1 Introduction

There have been a number of efforts to add semantics to the discovery process. An early work in this area has been the creation of DAML-S [1], which uses a DAML+OIL based ontology for describing Web Services. The latest draft release [2] of DAML-S uses WSDL (Web Service Description Language) in conjunction with DAML-S for Web Service descriptions. Web Services Description Language (WSDL) is an emerging communication standard used by Web Services. Management of the interaction between the broker and the service provider based on WSDL is difficult for non-trivial tasks [5]. Previously, an XML based language is proposed to describe the semantics of Web services [3]. Base on this language, I investigate the semantics discovery of Web services based on domain ontology. Multiple service discovery strategies are defined and explained. The semantics service discovery methods provide a new way to locate and utilize published Web services. The approach is flexible and extensible to accomplish complex web service requests.

2 Web Services Description

To discover to suitable web services, a language is needed to describe the request. In this section, the factors required to represent knowledge in Web services description are described. The following factors are identified to describe the semantics of Web services, and an XML based language for service description is given in [3].

- Inputs: This part specifies the objects that a Web Service takes as inputs;
- Outputs: This part of the representation specifies the objects that will be the outputs generated by this service;
- Input Constraints: This part defines the constraints that expected to hold before this capability can be applied, i.e. the preconditions of this capability;
- Output Constraints: This defines the postconditions after the capability has been performed;
- Privacy: Privacy should also be considered, since some service providers or consumers may not want their identities to be revealed to others, whom they know nothing or little about;
- Quality: Quality is always a concern of service consumers. Different service providers might provide the same service, but the qualities of their services may vary a lot. By specifying the quality of the required service, consumers have the choice of selecting the service according to their requirement.

3   Web Service Discovery

The signature of a web service is defined as following. In this definition, the sequence of variables types, specified in domain ontology terms, taken and given by a web service decides its signature.

Definition 1. Web Service Signature *The signature of a web services is the sequences of types of its input and output variables.*

3.1   Signature Discovery

Definition 2. Signature Discovery *Let $\mathcal{C}$ be a service description in ACDL (Agent Capability Description Language) containing: an input specification $I^{\mathcal{C}}$ containing the variables $v_1, \ldots, v_n$, and output specification $O^{\mathcal{C}}$. Let $\mathcal{T}$ be a service request in ACDL with input specification $I^{\mathcal{T}}$ containing variables $u_1, \ldots, u_m$, and output specification $O^{\mathcal{T}}$. $\mathcal{C}$ is a signature discovery web service of $\mathcal{T}$, if*

$$I^{\mathcal{T}} \preceq_{st} I^{\mathcal{C}} \text{ and}$$
$$O^{\mathcal{C}} \preceq_{st} O^{\mathcal{T}}$$

where $I^{\mathcal{T}} \preceq_{st} I^{\mathcal{C}}$ means $\forall v_i \in I^{\mathcal{C}} \exists u_j \in I^{\mathcal{T}}$ that $u_j \preceq_{st} v_i$ and for $i \neq k$, $u_j \preceq_{st} v_i$, and $u_l \preceq_{st} v_k$, $j \neq l$.

This discovery strategy is the basic and necessary requirement in the process, since without signature compatibility a requested service's input variables can not be accepted by the provider. From the definition, it is worth noting that the signature discovery relation is not commutative; that is, if service $\mathcal{C}$ is signature matched with $\mathcal{T}$, it is not necessary that $\mathcal{T}$ is signature matched against $\mathcal{C}$.

An Example of Signature Discovery. Following an example is given to illustrate the above definitions and the signature algorithm. Considering the following agent capability description of SportsFinder [4], an information agent to extract sports results from the web:

```
(capability
    :cap-id SportsFinder
    :constraint-language fopl
    :input ( (Team ?team) )
    :input-constraint (
            (elt ?team TeamName)
            (CompeteIn ?team Sports) )
    :output ( (TeamScore ?score) )
    :output-constraint ( (Has Score ?team ?score) ) )
```

The above description shows that SportsFinder can find out the scores of a sports team. Suppose the mediator has already received the above service advertisement. Then some information agent sends the following service request to the mediator:

```
(capability
    :cap-id SoccerResult
    :constraint-language fopl
    :input ( (SoccerTeam ?soccer_team) )
    :input-constraint (
            (elt ?soccer_team TeamName)
            (CompeteIn ?soccer_team Soccer) )
    :output ( (Score ?result) )
    :output-constraint ( Has Score ?soccer_team ?result) ) )
```

When applying the signature discovery algorithm on these two service descriptions, we have input variable `soccer_team` as a subtype of `team` and output type `TeamScore` as a subtype of `Score`, thus `SoccerResult` is signature matched against `SportsFinder`. That means the service of `SportsFinder` can take the variables of the request as input, and its output is compatible with the variables' types of request. From this example it is easy to understand that signature match relation is not commutative. For the above two descriptions, service `SportsFinder` is not signature matched with `SoccerResult`, although vice versa.

3.2 Constraint Discovery

Definition 3. Constraint Discovery *Let $\mathcal{C}$ be a capability description with input constraints $C_I^{\mathcal{C}} = \{ C_{I_1}^{\mathcal{C}}, \ldots, C_{I_{k_{\mathcal{C}}}}^{\mathcal{C}} \}$ and output constraints $C_O^{\mathcal{C}} = \{ C_{O_1}^{\mathcal{C}}, \ldots, C_{O_{l_{\mathcal{C}}}}^{\mathcal{C}} \}$. Let $C_I^{\mathcal{T}} = \{ C_{I_1}^{\mathcal{T}}, \ldots, C_{I_{k_{\mathcal{T}}}}^{\mathcal{T}} \}$ and $C_O^{\mathcal{T}} = \{ C_{O_1}^{\mathcal{T}}, \ldots, C_{O_{k_{\mathcal{T}}}}^{\mathcal{T}} \}$ be the input and output constraints respectively of service $\mathcal{T}$. We define $\mathcal{T}$ is constraint matched with $\mathcal{C}$ if*

$$C_I^{\mathcal{T}} \preceq_\theta C_I^{\mathcal{C}} \text{ and}$$
$$C_O^{\mathcal{C}} \preceq_\theta C_O^{\mathcal{T}}$$

where $\preceq_\theta$ denotes the θ-subsumption relation between constraints. For $C_I^{\mathcal{T}} \preceq_\theta C_I^{\mathcal{C}}$ means $\forall C_{I_i}^{\mathcal{T}} \in C_I^{\mathcal{T}} \exists C_{I_j}^{\mathcal{C}} \in C_I^{\mathcal{C}}$ that $C_{I_i}^{\mathcal{T}} \preceq_\theta C_{I_j}^{\mathcal{C}}$ and for $i \neq k$, $C_{I_i}^{\mathcal{T}} \preceq_\theta C_{I_j}^{\mathcal{C}}$, and $C_{I_k}^{\mathcal{T}} \preceq_\theta C_{I_l}^{\mathcal{C}}$, we have $j \neq l$.

In this definition a θ-subsumption relation is introduced between constraints, the definition is given next.

θ-Subsumption Between Constraints. Since all the constraints are given in `constraint-language`, the details of θ-subsumption depends on the constraint-language. In first order predicate logic, which we used as the constraint-language in the examples, constraints are a set of clauses. In [7], the definition of clause θ- subsumption is given below.

Definition 4. Clause θ-subsumption *A clause C is θ-subsumed by another clause D, denoted as $C \preceq_\theta D$, if there exists a substitution θ such that $C \subseteq \theta(D)$.*

For example,
$$(\text{CompeteIn Tigers NBL}) \preceq_\theta (\text{CompeteIn ?team NBL})$$
which means when we substitute the variable `?team` with a constant `Tigers`, a Melbourne Basketball team, the two clauses are equivalent.

θ-subsumption between two constraints, each a set of clauses, is defined in terms of subsumption between single clauses. More specifically, let $C^\mathcal{C}$ and $C^\mathcal{T}$ be two constraints. Then we define $C^\mathcal{C}$ θ-subsumption $C^\mathcal{T}$, denoted as $C^\mathcal{T} \preceq_\theta C^\mathcal{C}$, if every clause in $C^\mathcal{T}$ is θ-subsumed by a clause in $C^\mathcal{C}$.

Algorithm and an Example. It is straightforward to have an algorithm to check if two capability descriptions are constraint matched. Next an example is given for constraint discovery. Considering the following capability description, which has been registered to the mediator.

```
(capability
    :cap-id Golf
    :constaint-language fopl
    :input ( (GolfPlayer ?player)
             (Tournament ?tour) )
    :input-constraint (
            (elt ?player PlayerName)
            (CompeteIn ?player ?tour) )
    :output ( (Rank ?rank) )
    :output-constraint (
            (Has Rank ?player ?rank)
            (RankingIn ?rank ?tour) ) )
```

The above advertisement describes an information agent with the capability to find out the position of a golf player in a tournament. Input variables are the player's name and the tournament that the player attended; output is the ranking of the golf player. Suppose an agent requests a service of finding a golf player's position in the PGA Golf Tournaments, then it will send a service request as following to the mediator:

```
(capability
    :cap-id GolfPGA
    :constaint-language fopl
    :input ( (GolfPlayer ?player) )
    :input-constraint (
            (elt ?player PlayerName)
            (CompeteIn ?player PGA) )
    :output ( (Rank ?rank) )
    :output-constraint (
            (Has Rank ?player ?rank)
            (RankingIn ?rank PGA) ) )
```

According to the constraint definition, we know that if we apply a θ substitution: θ: `?tour` ← `PGA` to the constraints in `Golf` service description, the above two descriptions will have the same constraints. So the mediator can give that service request `GolfPGA` is constraint matched with `Golf`, that is the service of `Golf` is applicable to the request of `GolfPGA`. It is worth noting that we can use the same process as above to check the `io-constraint` fields of service descriptions.

3.3   Partial Discovery

Definition 5. Partial Discovery *Let $\mathcal{C}$ be a service description in our ACDL containing: an input specification $I^{\mathcal{C}}$ containing variables $V^{\mathcal{C}}_{I_1}, \ldots, V^{\mathcal{C}}_{I_{n_{\mathcal{C}}}}$, and output specification $O^{\mathcal{C}}$ with variables $V^{\mathcal{C}}_{O_1}, \ldots, V^{\mathcal{C}}_{O_{m_{\mathcal{C}}}}$, and $\mathcal{C}$'s input constraints $C^{\mathcal{C}}_I = \{ C^{\mathcal{C}}_{I_1}, \ldots, C^{\mathcal{C}}_{I_{k_{\mathcal{C}}}} \}$ and output constraints $C^{\mathcal{C}}_O = \{ C^{\mathcal{C}}_{O_1}, \ldots, C^{\mathcal{C}}_{O_{l_{\mathcal{C}}}} \}$. Let $\mathcal{T}$ be another agent service with the correspondent description parts as: input $I^{\mathcal{T}}$ containing variables $V^{\mathcal{T}}_{I_1}, \ldots, V^{\mathcal{C}}_{I_{n_{\mathcal{T}}}}$, and output specification $O^{\mathcal{T}}$ with variables $V^{\mathcal{T}}_{O_1}, \ldots, V^{\mathcal{T}}_{O_{m_{\mathcal{T}}}}$, and $\mathcal{T}$'s input constraints $C^{\mathcal{T}}_I = \{ C^{\mathcal{T}}_{I_1}, \ldots, C^{\mathcal{T}}_{I_{k_{\mathcal{T}}}} \}$ and output constraints $C^{\mathcal{T}}_O = \{ C^{\mathcal{T}}_{O_1}, \ldots, C^{\mathcal{T}}_{O_{l_{\mathcal{T}}}} \}$. We define $\mathcal{T}$ is partial matched with $\mathcal{C}$ if*

$$\exists V^{\mathcal{T}}_{I_i} \in I^{\mathcal{T}}, \exists V^{\mathcal{C}}_{I_j} \in I^{\mathcal{C}} \text{ that } V^{\mathcal{T}}_{I_i} \preceq_{st} V^{\mathcal{C}}_{I_j}$$
$$\exists V^{\mathcal{C}}_{O_j} \in O^{\mathcal{C}}, \exists V^{\mathcal{T}}_{O_i} \in O^{\mathcal{T}} \text{ that } V^{\mathcal{C}}_{O_j} \preceq_{st} V^{\mathcal{T}}_{O_i}$$
$$\exists C^{\mathcal{T}}_{I_i} \in C^{\mathcal{T}}_I, \exists C^{\mathcal{C}}_{I_j} \in C^{\mathcal{C}}_I \text{ that } C^{\mathcal{T}}_{I_i} \preceq_{\theta} C^{\mathcal{C}}_{I_j}$$
$$\exists C^{\mathcal{C}}_{O_j} \in C^{\mathcal{C}}_O, \exists C^{\mathcal{T}}_{O_i} \in C^{\mathcal{T}}_O, \text{ that } C^{\mathcal{C}}_{O_j} \preceq_{\theta} C^{\mathcal{T}}_{O_i}$$

If two services are not completely matched, but they have some functions in common, then we call this kind of discovery *partial discovery*. In some circumstances, partial matched services is applicable if the unmatched parts are not critical. The partial discovery means for two capability description, if some of their input, output variables have subtype relations, and there are constraint clauses in their input and output constraint specifications that are θ- subsumptioned, these two services are partial matched. Semantically, in some circumstances, i.e. the unmatched variables and constraints are irrelevant; the partial

matched service is applicable. In **Algorithm 1**, a method is given to evaluate if two service descriptions are partial matched. In the algorithm, we can also find out the matched variables and constraints in these two services.

Algorithm 1 Given two service description $\mathcal{C}$ and $\mathcal{T}$ with their input, output and constraints specifications as in **Definition 5**, this algorithm checks out if the two services are partial matched, and returns the matched fields.

$\quad inSetC \leftarrow I^{\mathcal{C}}; inSetT \leftarrow I^{\mathcal{T}}; inSet \leftarrow Null;$
$\quad$ **for** $v \in inSetC$ **do** {
$\quad\quad$ **for** $u \in inSetT$ **do**
$\quad\quad\quad$ **if** $(u \preceq_{st} v)$ **then** {
$\quad\quad\quad\quad inSetC \leftarrow inSetC - \{v\};$
$\quad\quad\quad\quad inSetT \leftarrow inSetT - \{u\};$
$\quad\quad\quad\quad inSet \leftarrow inSet + \{(u, v)\};$ } }
$\quad outSetC \leftarrow O^{\mathcal{C}}; outSetT \leftarrow O^{\mathcal{T}}; outSet \leftarrow Null;$
$\quad$ **for** $v \in outSetT$ **do** {
$\quad\quad$ **for** $u \in outSetC$ **do**
$\quad\quad\quad$ **if** $(u \preceq_{st} v)$ **then** {
$\quad\quad\quad\quad outSetC \leftarrow outSetC - \{v\};$
$\quad\quad\quad\quad outSetT \leftarrow outSetT - \{u\};$
$\quad\quad\quad\quad outSet \leftarrow outSet + \{(u, v)\};$ } }
$\quad inConSetC \leftarrow C_I^{\mathcal{C}}; inConSetT \leftarrow C_I^{\mathcal{T}}; inConSet \leftarrow Null;$
$\quad$ **for** $v \in inConSetC$ **do** {
$\quad\quad$ **for** $u \in inConSetT$ **do**
$\quad\quad\quad$ **if** $(u \preceq_{\theta} v)$ **then** {
$\quad\quad\quad\quad inConSetC \leftarrow inConSetC - \{v\};$
$\quad\quad\quad\quad inConSetT \leftarrow inConSetT - \{u\};$
$\quad\quad\quad\quad inConSet \leftarrow inConSet + \{(u, v)\};$ } }
$\quad outConSetC \leftarrow C_O^{\mathcal{C}}; outConSetT \leftarrow C_O^{\mathcal{T}}; outConSet \leftarrow Null;$
$\quad$ **for** $v \in outConSetT$ **do** {
$\quad\quad$ **for** $u \in outConSetC$ **do**
$\quad\quad\quad$ **if** $(u \preceq_{\theta} v)$ **then** {
$\quad\quad\quad\quad outConSetC \leftarrow outConSetC - \{v\};$
$\quad\quad\quad\quad outConSetT \leftarrow outConSetT - \{u\};$
$\quad\quad\quad\quad outConSet \leftarrow outConSet + \{(u, v)\};$ } }
$\quad$ **if** $(notEmpty(inSet) \wedge notEmpty(outSet)$
$\quad\quad \wedge\ notEmpty(inConSet) \wedge notEmpty(outConSet)$
$\quad$ **then return true else return false;**

Fig. 1. Partial Discovery Algorithm

3.4   Dynamic Selection of Web Service Discovery Strategy

Next question is how to decide which strategy to apply? Given the complexity of Web services and the dynamic changing environment, it is difficult to decide which strategy is the best for which circumstance. However, to select a suitable service discovery strategy, the following factors, not restricted to, should be considered:

- The features of a service request, such as the input, output variables' types and constraints;
- The attributes of service advertisements. For example, the privacy the service provider will take account of;
- The current load of matchmaking and other environment attributes.

Different developers can utilise various decision models to select a service discovery strategy or a combination of above strategies to satisfy their requirements.

4   Conclusion

This paper addresses the web service discovering problem. First I describe the problem and process of service discovering. An XML based language for web service describing is presented. The constraints parts in this language is independent, which users can plug-in any constraint language, such as first order predicate logic language. This provides flexibility and compatibility. Multiple service discovering strategies are given according to different levels of requirements and various features of service advertisements and requests. In the discovery strategies, I consider multiple attributes of the services and the features of service requirements and advertisements. Multiple discovery methods give users more choices and they are flexible to deal with the changing environment.

References

1. Ankolenkar, A., Burstein, M., Hobbs, J.R., Lassila, O., Martin, D.L., McDermott, D., McIlraith, S.A., Narayanan, S., Paolucci, M., Payne T.R., and Sycara, K. The DAML Services Coalition, "DAML-S: Web Service Description for the Semantic Web", The First International Semantic Web Conference (ISWC), Sardinia, Italy, June 2002.
2. DAML-S 0.7 Draft Release, 2002.
3. Lu, H., Sajjanhar, A., Meaningful UDDI Semantics Description, International Conference on Computer Applications in Industry and Engineering, Las Vegas, USA, November 2003.
4. Lu, H. and Sterling, L and Wyatt, A., Knowledge Discovery in SportsFinder: An Agent to Extract Sports Results from the Web, Methodologies for Knowledge Discovery and Data Mining, Third Pacific-Asia Conference (PAKDD-99) Proceedings, pp. 469-473, Beijing, 1999.
5. W3C. Web Services Definition Language. http://www.w3.org/TR/wsdl, 2002.
6. Sycra, K., Paolucci, M., Soudry, J., Srinivasan, N., Dynamic Discovery and Coordination of Agent-Based Semantic Web Services, IEEE Internet Computing, May/June, 2004.
7. Sycara, K., Lu, J., Klusch, M., Interoperability among Heterogeneous Software Agents on the Internet, Technical Report of the Robotics Institute, Carnegie Mellon University, 1998.

Making Sense of Ubiquitous Data Streams –
A Fuzzy Logic Approach

Osnat Horovitz, Mohamed Medhat Gaber, and Shonali Krishnaswamy

School of Computer Science and Software Engineering, Monash University
osnat.horovitz@gmail.com,
{Mohamed.Medhat.Gaber,Shonali.Krishnaswamy}
@infotech.monash.edu.au

Abstract. There is currently a growing new focus in data mining - Ubiquitous Data Mining (UDM). UDM is the process of mining data streams in a ubiquitous environment, on resource constrained devices [KPP02]. UDM is widely applied in facilitating real-time decision making in mobile and highly dynamic environments/applications, such as road safety and mobile stock portfolio monitoring. A significant challenge in these contexts is the interpretation and analysis of results produced through unsupervised techniques (which are invaluable since little is known about the streamed data). We propose a novel fuzzy approach that leverages the significant benefits of UDM clustering and supplements the interpretation and use of these results through using expert/background knowledge.

1 Introduction

The abundance of mobile devices, such as PDAs and cellular phones, coupled with the progress made in wireless communication, has made the possibility of performing mobile data mining, which is the analyzing and mining of data streams from handheld devices, significantly more feasible. Consequently, considerable effort has focused on researching and developing this field which was termed Ubiquitous Data Mining (UDM). UDM is the process of mining data streams in a ubiquitous environment, on resource constrained devices [KPP02]. It is a currently emerging focus in data mining, encouraged and supported by the constantly growing computational abilities of mobile devices, which have resulted in systems that are capable of performing time-critical, intelligent, data analysis on streams of continuous data, thereby facilitating the "anytime anywhere" [KLZ02] data mining paradigm.

As the significance of UDM is the combination of real-time data stream mining, in mobile and dynamic environments, on resource constrained devices, standard data mining algorithms must be modified appropriately. In this paper we present an approach to mining and analyzing data streams online, inspired by the principles of fuzzy logic. We apply our approach to a specific scenario in the field of road safety. More specifically, we aim to mine on-board vehicle sensory data for the purpose of detecting, and alerting to, dangerous driving behavior.

Typical scenarios where UDM techniques have been, applied/discussed include [Gro98] [Sin01] [CSK01] [KBL04] [KPP02]. The major challenge of mining data streams is the fact that, unlike mining relational databases, with data streams only one pass over the data is possible. This requires the development of specialized algo-

rithms. Various algorithms which tackle the particular issues of online clustering of data streams have been designed [GMM00] [GuG04] [GKZ04] [CCF97]. Due to the unbounded nature of data streams, and the resource constrained devices often used to process them, all these clustering algorithms are, to a greater or lesser extent, incremental.

Various data mining techniques have been applied to different aspects of road safety [FMM03] [Sin01] [Sin01a] [OlP00]. Because the field of road safety deals with individual vehicles, constantly on the move, it is evident that Ubiquitous Data Mining is especially suited for it. Using mobile technology sensory data, in the form of data streams, can be collected and analyzed on-board the vehicles in real-time. UDM techniques have been applied to road safety in the VEDAS (VEhicle DAta Stream mining) system [KBL04], which is the work that most closely relates to our own research. VEDAS uses clustering to detect problems in a vehicle's condition and distribution statistics to monitor driver behavior.

While clustering is a very valuable technique for dealing with unknown data streams, for many UDM applications it is constrained by the difficulty in interpretation of the results for non-expert users such as mobile stockbrokers. As UDM applications are typically targeted at mobile and dynamic environments, there is a need for real-time decision making which is not facilitated by unsupervised data mining techniques, where human experts are needed to make sense of clustering results. Our aim is to support the process of increased understanding of clustering results, so that they can be used to identify, or classify, new unseen data instances in real-time UDM applications. Our system uses dynamic labeling to bridge the gap between clustering and classification, through a fuzzy approach. We employ fuzzy logic principles to combine unsupervised data mining results with human expertise and background knowledge, in the form of rules, to attain labeled classes which can be utilized for online classification of data streams. Another advantage of the fuzzy logic approach is that it gives classification results which include a degree of probability. To illustrate our approach we apply it in the field of road safety where we establish that it is feasible to detect dangerous driving behavior in real-time, so that an appropriate action can be initiated with minimal user intervention in real-time.

The remainder of the paper is organised as follows. Section two presents our fuzzy logic approach, as well as a fuzzy algorithm for online classification. Section three presents the evaluation results of our approach, applied to a specific road safety scenario. Finally, in section four, we conclude the discussion.

2 A Fuzzy Logic Approach to UDM

We have created a two-staged model for using data mining to improve road safety, which has two major stages; data synopsis, and online classification of driving behavior. In the first stage, the data synopsis stage, we focus on the online clustering of data streams, using UDM techniques. In the second stage, the online classification stage, we use classification models constructed using the data obtained in the data synopsis stage to perform online monitoring of driving behavior, so that when dangerous behavior is detected an alarm can be raised.

The main challenge facing us is, therefore, how to translate the clustering models we obtain in the data synopsis stage to comprehensive classification models we can

use in the online classification stage. We need to find a way to combine the expert knowledge we have in the field of road safety with the results of the unsupervised data mining of data streams. As a solution to this challenge, we propose to use a fuzzy logic approach to classification. According to [Men95], the uniqueness of fuzzy logic systems is their ability to combine what is described as "objective knowledge" and "subjective knowledge". In the case of our system, the "objective knowledge" is the numerical models built in the data synopsis stage, while the "subjective knowledge" is expert knowledge we posses in the field of road safety and the area driving behavior.

The fuzzy approach we adopt in our model manifests itself in the two phases of the online classification stage:

1. The labeling of the clustering models according to expert knowledge, thereby making them classification models.
2. The results of online classification will show the degree of probability of a driver's behavior belonging to each of the known driving behavior classes.

We use expert knowledge to label unsupervised models is by creating a knowledge base containing rules which can be applied to the unsupervised clusters, as shown in Figure 1. For example, a rule might indicate that the more times a driver exceeds the speed limit, the higher the likelihood of the driver being drunk. Initially, the rules will comprise of available expert knowledge but over time the knowledge base will be updated and refined according to the results of the data synopsis.
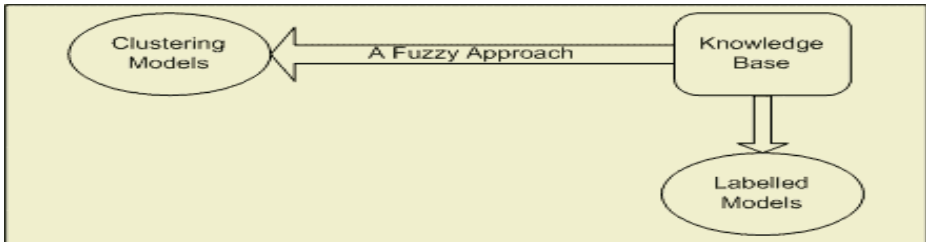
Fig. 1. Clustering to classification through fuzzy logic

The labeling of the unsupervised models of driving behavior will create, in effect, classification models of dangerous driving behavior. Once a sufficiently comprehensive number of robust supervised driving behavior models exist, the online classification stage can begin. The dangerous driving behavior models will be used as the means to monitor and classify the incoming data streams on an on-board vehicle device, and take appropriate action as deemed necessary. At the same time, the data synopsis will continue, and the models will be continually refined and updated. Updated models will be transmitted from the central server to the on-board vehicle devices periodically.

It is important to note that although in this particular case the labelling of clustering models is done on the central server, due to the fact that data gathered from various vehicles is used to construct the models, in other scenarios, such as personalised stock portfolio monitoring, the labelling can be done on-board the mobile device, thus facilitating real-time decision making.

2.1   A Fuzzy Classification Algorithm

The other aspect of our fuzzy logic approach to classification is that the results of classifying a data stream sample, which is the driver's current driving behavior, do not merely show the class to which the data sample is closest, but give the probability of that data sample belonging to any of the known driving behavior classes. In other words, we determine the degree of membership of a new data sample to each of the given classes.

When using the supervised model for online classification, three factors should be considered:

1. The distance of the data sample from the class centre.
2. The weight of the class centre.
3. The distance of the data sample from the other class centers.

Taking these factors into account, this is the proposed algorithm:

1. For each of the cluster centers:
    a. Calculate the distance of the data sample to the cluster centre.
    b. Multiply the inverse distance to the cluster centre by the weight of the cluster centre, divided by the combined weight of all the cluster centers.

$$C_x = \frac{1}{dist} \times \frac{weight}{\sum_{i=1}^{numberofcenters} weight}$$

   Where *dist* represents the distance from the new data sample to a cluster centre, and *weight* represents the number of data samples in a particular cluster.
2. To calculate the probability of the data sample belonging to each of the clusters, divide the result of the first step by the combined results of the first step of all the cluster centers.

$$\text{Degree of membership} = \frac{Cx}{\sum_{i=1}^{numberofcenters} Ci}$$

This algorithm gives us the probability of the data sample belonging to a certain class. Different actions can be taken according to the degree of membership and to the behaviour class, like raising an alarm or notifying other vehicles in the vicinity.

3   Evaluation of the Model in a Road Safety Scenario

As we discussed, we have created a model for using UDM techniques, in combination with fuzzy logic principles, in the field of road safety. In this section we will present the results of applying the supervised models to online classification of data streams, in order to detect, and alert to, dangerous driving behavior. Our goal is to be able to make sense of previously unseen data.

For the purpose of evaluation, our algorithm is used in the case of drunk driving behavior. The first step of the algorithm involves randomly generating data streams which emulate the kind of sensory data that can be obtained in a moving vehicle. Incremental clustering is then performed on the unseen data, on the on-board device,

using the LWC algorithm [GKZ04], which is a lightweight one-look incremental clustering algorithm designed for operating in real time on resource constrained devices. The generated data is taken from a study presented in [MBF00], where participants used a driving simulator with deferent degrees of blood alcohol concentration. The following response measures were recorded in the study: reaction time to peripheral signals (sec), correct responses to peripheral signals (number), speed deviation (mph), lane position deviation (ft), collisions (number) and times over speed limit (number). In the second step of the algorithm, we transmit the clustering results to the server, where, using fuzzy logic principles, we combine these results with expert knowledge. This results in labeled classes of drunk driving behavior, ordered by the degree of drunkenness; from the least drunk to the most drunk. The third step of the algorithm is the online classification. The labeled classification models are sent back to the on-board device, where they are used to classify new, unknown, data samples. The testing of the on-board device was done on iPAQ with 64 MB, running Microsoft Windows CE version 3.0.9348 with StrongArm processor. The program has been developed using Microsoft embedded Visual C++ 4.0.

As discussed, in the first step of the algorithm we cluster generated data streams. Using the data recorded in the [MBF00] study, we generate simulated sensory data for four types of drunk driving behavior; sober, borderline, drunk and very drunk. The data is then clustered on-board the device, where the weight of a cluster centre represents the number of items in that cluster.

In the second step of the algorithm, the clustering centers are sent to the central server where, using expert knowledge, they are ordered by degree of drunkenness. The expert knowledge includes rules such as:

- The *higher* the number of "Correct responses to peripheral signals", the *less* drunk the driver.
- The *higher* the number of "Times over speed limit" ", the *more* drunk the driver.

In the third stage of the algorithm the classification models are used on-board the device to classify new data items. Here again we apply the fuzzy logic approach. Instead of merely classifying a driver's behavior to one of the available classes, we find its degree of membership, or the probability it of belonging, to each of the classes. Table 1 shows the results of the classification of randomly generated new data samples. Each of the data sample was generated from one of the four original classes of driving behavior mentioned above (Observed Behavior). The classifying application is, of course, unaware of these classes, but it allows us to test our expected results. The attributes for which data was generated include: *Number of Correct Responses, Number of Collisions, Time over Speed Limit, Reaction Time, Speed Deviation and Lane Deviation.*

We can make several observations from the results of the algorithm. Firstly, let us consider the results of the first stage and second stage of the algorithm. After the online clustering and offline labeling we ended up with three classes of driving behavior with varying degrees of drunkenness, from least drunk to most drunk, even though we originally generated the data as four distinct driving behavior classes. This is because the difference between the borderline and the drunk driving behavior is difficult to distinguish, while the distinction between these two classes and the sober and very drunk behavior classes is very clear.

Table 1. Classification of different unseen data samples

Trial#	Observed Behaviour	EXPECTED RESULTS		
		Class1 (least drunk)	Class2	Class 3 (most drunk)
1	Sober	79.4%	12.5%	8.1%
2	Drunk	21.5%	52.2%	26.3%
3	Very Drunk	9.2%	9.5%	81.3%
4	Borderline	26.6%	61.3%	12.1%
5	Very Drunk	9.9%	10.1%	80%
6	Drunk	20.8%	55.8%	23.4%

How then do we know which driver is borderline drunk and which driver certainly is drunk? This is helped by the results of the third stage of the algorithm. Because we have the degree of membership to each class of behavior, if a new data sample has the highest degree of membership to class number 2 we can determine the level of drunkenness by observing in which of the other classes that sample has a higher degree of membership.

It is evident that our method of combining UDM techniques with more traditional data mining technique and a fuzzy logic approach provides results which are easily comprehensible, and which allow for an immediate, as well as appropriate, respond, without the need of human intervention.

4   Conclusions

UDM is a growing focus in data mining, and has much potential. UDM applications often require the ability to perform real-time decision making based on quick and accurate analysis of incoming data streams. While unsupervised data mining techniques provide insight into unknown data streams, they necessitate human intervention for the purpose of interpreting results. We, therefore, seek to enhance the clustering process by adding expert and background knowledge to it, thereby making instant analysis of results and appropriate action taking possible.

Road safety is clearly an area where UDM techniques can have a significant benefit. In our model we create a data synopsis of driving behaviour by using UDM techniques for online incremental clustering of data streams. A fuzzy logic approach is then applied in order to facilitate the understanding of these clusters. We combine unsupervised data mining results with available expert knowledge to achieve classification models of driving behaviour. The classification models are then used on-board a vehicle, on a mobile device, where for each new data sample we show the probability of it belonging to any of the known driving behaviour classes. In the case of dangerous driving behaviour being detected, an appropriate action is taken.

References

[CCF97] Charikar, M., Chekuri, C., Feder T. and Motwani, R., Incremental Clustering and Dynamic Information Retrieval, in proceedings of the 29th annual ACM Symposium on Theory of Computing, 1997.

[CSK01] Chen, R., Sivakumar, K., Kargupta, H., An Approach to Online Bayesian Learning from Multiple Data Streams, Workshop on Ubiquitous Data Mining for Mobile and Distributed Environments, Held in Conjunction with Joint 12th European Conference on Machine Learning (ECML'01) and 5th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'01), Freiburg, Germany, September 3-7, 2001.

[FMM03] Flach, P.A., Mladenic, D., Moyle, S., Raeymaekers, S., Rauch, J., Rawles, S., Ribeiro, R., Sclep, G., Struyf, J., Todorovski, L., Torgo, L., Blockeel, H., Wettschereck, D., Wu, S., Gartner, T., Grobelnik, M., Kavsek, B., Kejkula, M., Krzywania, D., Lavrac, N. and Ljubic, P., On the road to knowledge: mining 21 years of UK traffic accident reports. In: Data Mining and Decision Support: Aspects of Integration and Collaboration, pages 143--155. Kluwer Academic Publishers, January 2003.

[GKZ04] Gaber, M. M., Krishnaswamy, S., Zaslavsky, A., Cost-Efficient Mining Techniques for Data Streams, Australasian Workshop on Data Mining and Web Intelligence (DMWI2004), Dunedin, New Zealand, 2004.

[GMM00] Guha, S., Mishra, N., Motwani, R., and O'Callaghan, L., Clustering data streams, in Proc. FOCS, p. 359–366, 2000.

[Gro98] Grossman, R., Supporting the Data Mining Process with Next Generation Data Mining Systems, Enterprise Systems, August 1998.

[GuG04] Gupta, C., and Grossman, R. L., GenIc: A Single Pass Generalized Incremental Algorithm for Clustering, International Conference Data Mining, SIAM 2004.

[KBL04] Kargupta, H., Bhargava, R., Liu, K., Powers, M., Blair, P., Bushra, S., Dull, J., Sarkar, K., Klein, M., Vasa, M., and Handy, D., VEDAS: A Mobile and Distributed Data Stream Mining System for Real-Time Vehicle Monitoring, Proceedings of SIAM International Conference on Data Mining, April 2004.

[KPP02] Kargupta, H., Park, B., Pittie, S., Liu, L., Kushraj, D., and Sarkar, K., MobiMine: Monitoring the Stock Market from a PDA. ACM SIGKDD Explorations, Volume 3, Issue 2. Pages 37--46. ACM Press, January 2002.

[KLZ02] Krishnaswamy, S., Loke, S., and Zaslavsky, A., Towards Anytime Anywhere Data Mining E-Services, Proceedings of the Australian Data Mining Workshop (ADM'02) at the 15th Australian Joint Conference on Artificial Intelligence, (eds) S.J. Simoff, G.J. Williams, and M. Hegland. Canberra, Australia, December 2002, pp. 47 - 56, Published by the University of Technology Sydney, 2002.

[MBF00] Moskowitz, H., Burns, M., Fiorentino, D., Smiley, A., Zador, P., Driver Characteristics and Impairment at Various BACs, Southern California Research Institute, August 2000.

[Men95] Mendel, J. M., Fuzzy logic systems for engineering: A tutorial, Proceedings of the IEEE, 83(3):345–377, March 1995.

[OlP00] Oliver, N. and Pentland, A. P., Graphical Models for Driver Behavior Recognition in a SmartCar, MIT IEEE intelligent vehicles symposium, 2000.

[Sin01] Singh, S., Identification of Driver and Vehicle Characteristics through Data Mining the Highway Crash, National Highway Traffic Safety Administration, USA, 2001.

[Sin01a] Singh, S., A Sampling Strategy for Rear-End Pre-Crash Data Collection, National Highway Traffic Safety Administration, USA, 2001.

σ-SCLOPE: Clustering Categorical Streams Using Attribute Selection

Poh Hean Yap and Kok-Leong Ong

School of Information Technology, Deakin University
Waurn Ponds, Victoria 3217, Australia
{phya,leong}@deakin.edu.au

Abstract. Clustering is a difficult problem especially when we consider the task in the context of a data stream of categorical attributes. In this paper, we propose σ-SCLOPE, a novel algorithm based on SCLOPE's intuitive observation about cluster histograms. Unlike SCLOPE however, our algorithm consumes less memory per window and has a better clustering runtime for the same data stream in a given window. This positions σ-SCLOPE as a more attractive option over SCLOPE if a minor lost of clustering accuracy is insignificant in the application.

1 Introduction

In recent years, the data in many organizations take the form of continuous streams, rather than finite stored data sets. This possess a challenge for data mining, and motivates a new class of problem call *data streams* [1–3]. Designing algorithms for data streams is a challenging task: (a) there is a sequential one-pass constraint on the access of the data; (b) and it must work under bounded (i.e., fixed) memory with respect to the data stream.

Also, the continuity of data streams motivate time-sensitive data mining queries that many existing algorithms do not adequately support. For example, an analyst may want to compare the clusters, found in one window of the stream, with clusters found in another window of the same stream. Or, an analyst may be interested in finding out how a particular cluster evolves over the lifetime of the stream. Hence, there is an increasing interest to revisit data mining problems in the context of this new model and application.

In this paper, we study the problem of clustering a data stream of categorical attributes. Data streams of such nature, e.g., transactions, database records, Web logs, etc., are becoming common in many organizations [4]. Yet, clustering a categorical data stream remains a difficult problem. Besides the dimensionality and sparsity issue inherent in categorical data sets, there are now additional stream-related constraints. Our contribution towards this problem is the σ-SCLOPE algorithm that is an improvement over the recently proposed SLCOPE inspired by two other works: the CluStream [5] framework, and the CLOPE [4] algorithm.

Our proposed strategy in σ-SCLOPE has been proven to improve the performance of SCLOPE in terms of memory consumption and runtime. This is

important since the efficiency of the algorithm is extremely important in the data stream environment. Our strategy is to introduce attribute selection to eliminate redundant nodes in an otherwise complete FP-Tree structure used in SCLOPE. Using basic frequency ordering and support thresholds, we effectively obtain a smaller FP-Tree while maintaining most of the accuracy in clustering. Furthermore, because the frequency order is updated as the stream is seen, clustering accuracy improves from window to window.

2   Algorithm Design

We begin our problem discussion by giving the following scenario. Consider a set of records $\mathcal{R}_1, \ldots, \mathcal{R}_i, \ldots$ collected in data stream $\mathcal{D}$ at time periods $t_1, \ldots,$ $t_i, \ldots$, such that each record $\mathcal{R} \in \mathcal{D}$ is a vector containing attributes drawn from $\mathcal{A} = \{a_1, \ldots, a_j\}$. A clustering $\mathcal{C}_1, \ldots, \mathcal{C}_k$ on $\mathcal{D}_{(t_p, t_q)}$ is a partition of records $\mathcal{R}_x, \mathcal{R}_y, \ldots$ seen between t_p and t_q (inclusive), such that $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_k = \mathcal{D}_{(t_p, t_q)}$ and $\mathcal{C}_\alpha \neq \varnothing$ and $\forall \alpha, \beta \in [1; k)$, $\mathcal{C}_\alpha \cap \mathcal{C}_\beta = \varnothing$. We need to perform clustering for records in each window $W_{(t_p, t_q)}$, where each window contains a set of records $\mathcal{R}_i, \ldots, \mathcal{R}_j, \ldots$ captured at a specified time periods t_p and t_q.

2.1   Observation About FP-Tree

A compact FP-Tree structure can be designed based on the following observations:

- One scan on $\mathcal{D}_{(t_1, t_2)}$ is necessary to obtain attributes captured at each specified time window. It is essential to allow only a single sequential access to the database in order to meet the one pass constraint in data streams environment.
- If we started constructing FP-Tree for first window with a random order of attributes, we can avoid repeated database scans.
- If an ordered frequency of attributes is used in constructing the tree structure, a smaller FP-Tree will be created.
- Each node in FP-Tree represents a singleton attribute. If descending ordered frequency of attributes is used, more frequently occurring items will be inserted into the tree before the less frequent ones. Hence sharing of nodes will be more likely to occur.
- Records with different attributes create a new subtree in the structure while records with the same attributes are merged with existing nodes. Sharing of the same prefix paths can contribute to a natural clustering of the records.
- Since only frequent items will contribute to frequent patterns mining, if an attribute selection is applied in constructing the tree structure, it will further reduce the size of its resulting FP-Tree.

2.2   Proposed Solution

With the above observations, we now begin to construct our FP-Tree structure for first window, W_1. The following sample records are recorded between t_1 and t_2: $\{\langle c, b, a \rangle, \langle a, b, d \rangle, \langle c, d, e \rangle, \langle e, b, g \rangle, \langle a, b, c \rangle, \langle f, e, d \rangle, \langle g, f, d \rangle, \langle a, b, c \rangle\}$.

Table 1. Singleton frequency for attributes in window W_1

Attributes	c	b	a	d	e	g	f
Singleton Frequency	4	5	4	4	3	2	2
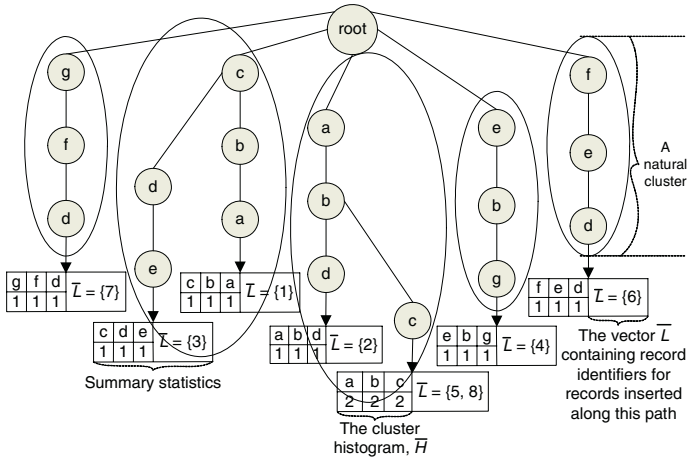
Fig. 1. The FP-Tree for first window, W_1

Upon arriving , the singleton frequency of each incoming attribute is recorded in Table 1. Then each attribute is inserted as nodes in the FP-Tree as shown in Figure 1. Each individual path in the FP-Tree created leads to a natural clustering of its attributes, forming a cluster Histogram $\overline{H}$ and a vector $\overline{L}$ containing the record identifiers that is used in distinguishing records that had contributed to that cluster. At the end of window W_1, the ordering of attributes is updated as shown in Table 2. This updated descending ordered singleton frequency for attributes in W_1 would be used in constructing FP-Tree in the next window W_2.

In W_2, suppose we have a different set of records between t_2 and t_3: $\{\langle e, g, f \rangle,$ $\langle c, d, a \rangle, \langle a, b, c \rangle, \langle a, e, d \rangle, \langle c, d, e \rangle, \langle a, d, c \rangle, \langle e, d, c \rangle, \langle b, c, a \rangle\}$. Similar to W_1, the singleton frequency of each incoming attribute is being recorded before inserting into the FP-Tree. However, a descending ordered of singleton frequency obtained from Table 2 is used in the tree construction this time, instead of random order as applied in W_1. Node with highest frequency in each record is first inserted into the tree following by the less frequent ones. Then, records with different attributes from the existing paths create a new sub-tree in the structure while records with the same attributes are merged with existing nodes.

By looking at information from Table 2, we can also see that items 'g' and 'f' are the least frequent items, which might not be of interest to user who wishes to infer interesting patterns from a large database. Hence, in addition to the use of ordered frequency of attributes, we apply an user specified support threshold, e.g., $\sigma=2$. Thus attributes 'g' and 'f' will be eliminated from the tree construction, resulting in FP-Tree as shown in Figure 2(a).

Table 2. Descending ordered singleton frequency for attributes in window W_1

Attributes	b	c	a	d	e	g	f
Descending Ordered Singleton Frequency	5	4	4	4	3	2	2
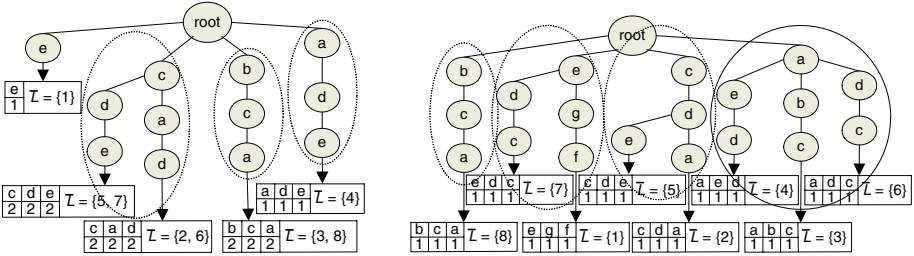
Fig. 2. Construction of FP-Tree for W_2 (a) with (b) without the use of descending ordered frequency of attributes and support threshold, σ

At the end of window W_2, the ordering of attributes is again updated as shown in Table 3 and this order will then be used in constructing the next window, W_3. As we can see, frequency order changes as the stream progresses. If we apply this case as transactions in retail businesses, popularity of the items would be the cause of changes in frequency. For example, item 'b' was sold at a promotional rate during time period t_1 and t_2. After the offer ended between t_2 and t_3, its frequency dropped from the top sales position to the value of σ, i.e. 2. Thus, continuous storage of all attributes is required in order to manage the progressive changes in data stream while not losing potentially valuable information. The same methods can be applied in $W_i, \dots, W_j, \dots$.

To illustrate the difference between using and not using an ordered frequency attributes and support threshold σ at time period t_2 and t_3, we supply the same set of records we had for window W_2: $\{\langle e, g, f \rangle, \langle c, d, a \rangle, \langle a, b, c \rangle, \langle a, e, d \rangle, \langle c, d, e \rangle, \langle a, d, c \rangle, \langle e, d, c \rangle, \langle b, c, a \rangle\}$.

We eliminated the need to store frequency of attributes and had then re-constructed the FP-Tree structure for W_2 without utilizing the ordered frequency of attributes and support threshold σ. Difference between the original tree for W_2 and this re-constructed tree structure is shown in Figure 2.

From the figure, we can see that a new FP-tree with more nodes is created in Figure 2(b) as compared to Figure 2(a). The number of unique micro-cluster is increased from 5 to 8, and there are less overlaps between the nodes. Since more overlaps means better intra-cluster similarity, we can see that merging of micro-clusters with similar attributes would be faster in Figure 2(a) than Figure 2(b), as shown by dotted circles in Figure 2.

Observation 1. *Given a set of records $\mathcal{R}_i, \dots, \mathcal{R}_j$ at time period t_i and t_{i+1}, by constructing the FP-Tree, we are able to store the complete set of its summary statistics as long as the frequency of all attributes in records $\mathcal{R}_i, \dots, \mathcal{R}_j$ are above the value of given support threshold.*

Table 3. Descending ordered singleton frequency for attributes in window W_2

Attributes	c	a	d	e	b	g	f
Descending Ordered Singleton Frequency	6	5	5	4	2	1	1

Consider the following set of records, $\{\langle a, b \rangle, \langle a, b \rangle, \langle c, d \rangle\}$. We have frequency of attributes 'a' = 2, 'b' = 2, 'c' = 1, and 'd' =1. If σ is 2, items 'c' and 'd' would be eliminated from the FP-Tree construction because their frequency is below the support threshold. Thus an FP-Tree with only one path consisting of nodes 'a' and 'b' would be created. Now, let σ be 1, and an FP-Tree with 2 unique paths would be created. One with the node 'a' followed by node 'b', and one with nodes 'c' and 'd'. Since the value of support threshold is 1, all attributes, as long as they exist in the given records, would be considered as frequent items and thus be inserted into the FP-Tree.

Observation 2. *Given an user specified support threshold value, items fall under σ are items which are not of interest to user.*

Since the value of σ is specified by user, user determines the minimum target frequency of attributes that user wish to include in their final analysis. In other words, the user specified support threshold determines items which are not of interest to the analyst at time period t_i and t_{i+1}. There is no lost of valuable information by constructing FP-Tree based on the given σ value simply because these items that fell below the targeted value are not of interest to user. One can say that there is still lost in data by keeping only the summary statistics for the large items. However, minimal lost can be compensated with the huge amount of time and memory space being saved by eliminating the need to store non-frequent items in the FP-Tree. Observation 3 estimates the amount of space being saved.

Observation 3. *Given any number of records $\mathcal{R}_i, \ldots, \mathcal{R}_j$ at time period t_i and t_{i+1}, amount of space being saved by applying the use of support threshold σ can range from a single node for a 2 items records to an order of magnitude for a huge database.*

Let us look at Figure 2 again. We can see a space saving of $5 \div 8 * 100 = 62.5\%$ in the number of unique clusters and $12 \div 19 * 100 = 63.16\%$ in the number of nodes on the FP-Tree for a database with only 8 records and 7 unique attributes. Now consider a database with millions of records, e.g. Walmart with 20 millions sales transaction records in a single day, amount of space saved can be tremendous.

3   Cluster Discovery

Once the construction of FP-Tree is completed, the analyst performs clustering over the summary statistics generated, using the offline macro-clustering component at different time-horizons. Since the offline component does not require

access to the data, its design is not constrained by the one-pass requirement. Please refer to SCLOPE [6] for full description on its offline macro-clustering component and online micro-clustering component algorithms.

A typical time-sensitive cluster discovery begins with the analyst entering the time-horizon h, and the repulsion r. The time-horizon of interest usually spans one or more windows, and determines the micro-clusters involved in the analysis. On the other hand, the repulsion controls the intra-cluster similarity required, and is part of the clustering criterion called *profit*. Its definition, given below, is the same as CLOPE.

$$\texttt{profit}(\{\mathcal{C}_1,\ldots,\mathcal{C}_k\}) = \left[\sum_{i=1}^{k}\left(\frac{\texttt{size}(\mathcal{C}_i)}{\texttt{width}(\mathcal{C}_i)^r} \times |\mathcal{C}_i|\right)\right] \times \left(\sum_{i=1}^{k}|\mathcal{C}_i|\right)^{-1}$$

In addition, we have also practiced the use of pyramidal timeframe as applied in SCLOPE. Pyramidal timeframe is able to store summary statistics at different time periods at different levels of granularity, as the stream progresses. Therefore, as data in the stream becomes outdated, its summary statistics looses details. This method of organization provides an efficient trade-off between the storage requirements and the quality of clusters from different time horizons. At the same time, it also facilities the answering of time-sensitive queries posed by the analyst.

The most interesting aspect of the offline macro-clustering component of SCLOPE is its design for time-sensitive data mining queries. When used together with the pyramidal timeframe, we can analyze different parts of the data stream, by retrieving statistics of different granularity to produce the clustering we need. And since this is the an offline component (which can run independent of the data stream), our design favors accuracy over efficiency, i.e., it makes multiple iterations through the statistics, and cluster using the profit criterion.

Nevertheless, our offline component is still fast despite the fact that it is based on the design of CLOPE and SCLOPE. The rationale behind this speed is that CLOPE works with one record at a time while SCLOPE works with a group of records at a time. In our algorithm, each micro-cluster is treated as a pseudo-record, and are clustered accordingly to the given r value that in turn, determines the number of clusters k. Since the number of pseudo-records are very much lower than the physical records, it takes less time to converge on the clustering criterion.

4   Related Works

Recent years, the emergence of data streams environment have been one of the hit topics among researches in data mining algorithms [3, 7]. CluStream [5], HPStream [8] and VFDT [9] are among the works related to mining within the high speed data streams.

While these algorithms have stressed on the importance of mining high speed data streams, this paper had utilized methods in clustering data with the use of

FP-Tree structure in dealing with the constrains of data streams. Most related works including FP-Growth [10], the foremost proposer on using FP-Tree structure in storing summary of data, CLOPE [4], which has introduced clustering in categorizing high dimensional data, and SCLOPE [6] which has inspired us with the idea of combining clustering with FP-Tree structure.

5   Conclusions

In this paper, we proposed a possible attribute selection strategies based on support threshold. However, to be effective in delivering knowledge, different domains may require different attribute selection strategies. For example, in our discussion of high-change support threshold of an item from one window to another, we may have to consider alternative strategies. This will be part of our future work, which we expect to report in the near future.

References

1. Babcock, B., Babu, S., Datar, M., Motwani, R., Widom, J.: Models and Issues in Data Stream Systems. In: Proc. ACM Symp. PODS. (2002)
2. Bradley, P.S., Gehrke, J., Ramakrishnan, R., Srikant, R.: Philosophies and Advances in Scaling Mining Algorithms to Large Databases. Communications of the ACM (2002)
3. Hulten, G., Domingos, P.: Catching Up with the Data: Research Issues in Mining Data Streams. In: Workshop on Research Issues in Data Mining and Knowledge Discovery, Santa Barbara, CA (2001)
4. Yang, Y., Guan, X., You, J.: CLOPE: A Fast and Effective Clustering Algorithm for Transactional Data. In: Proc. SIGKDD, Edmonton, Canada (2002)
5. Aggarwal, C., Han, J., Wang, J., Yu, P.S.: A Framework for Clustering Evolving Data Streams. In: Proc. VLDB, Berlin, Germany (2003)
6. Ong, K.L., Li, W., Ng, W.K., Lim, E.P.: SCLOPE: An Algorithm for Clustering Data Streams of Categorical Attributes. In: Proc., Zaragoza, Spain (2004) 209
7. Garofalakis, M., Gehrke, J., Rastogi, R.: Querying and Mining Data Streams: You Only Get One Look. In: ACM SIGMOD, Madison, Winconsin, USA (2003)
8. Aggarwal, C., Han, J., Wang, J., Yu, P.S.: A Framework for Projected Clustering of High Dimensional Data Streams. In: Proc. VLDB, Toronto, Canada (2004)
9. Domingos, P., Hulten, G.: Mining High-Speed Data Streams. In: Proc. ACM SIGKDD, Boston, MA (2000) 71–80
10. Han, J., Pei, J., Yin, Y.: Mining Frequent Patterns without Candidate Generation. In: Proc. SIGMOD, Dallas, Texas, USA (2000)

Extraction of Gene/Protein Interaction from Text Documents with Relation Kernel*

Jae-Hong Eom and Byoung-Tak Zhang

Biointelligence Lab., School of Computer Science and Engineering
Seoul National University, Seoul 151-744, South Korea
{jheom,btzhang}@bi.snu.ac.kr

Abstract. Even though there are many databases for gene/protein interactions, most such data still exist only in the biomedical literature. They are spread in biomedical literature written in natural languages and they require much effort such as data mining for constructing well-structured data forms. As genomic research advances, knowledge discovery from a large collection of scientific papers is becoming more important for efficient biological and biomedical researches. In this paper, we present a relation kernel based interaction extraction method to resolve this problem. We extract gene/protein interactions of Yeast (*S.cerevisiae*) from text documents with relation kernel. Kernel for relation extraction is constructed with predefined interaction corpus and set of interaction patterns. Proposed relation kernel for interaction extraction only exploits shallow parsed documents. Experimental results show that the proposed kernel method achieves a recall rate of 78.3% and precision rate of 79.9% for gene/protein interaction extraction without full parsing efforts.

1 Introduction

With the advancement of genomic technology and genome–wide analysis of organisms, one of the great challenges to post–genomic biology is to understand how genetic information of proteins results in the predetermined action of gene products, both temporally and spatially, to accomplish biological function and how they act together with each other to build an organism. Also, it is known that gene/protein interactions are fundamental biochemical reactions in the organisms and play an important role since they determine the biological processes [1]. Therefore, comprehensive description and detailed analysis of these interactions would significantly contribute to the understanding of many important biological phenomena and problems.

After the completion of the genome sequence of *S.cerevisiae* (budding yeast), many researchers have undertaken the task of functionally analyzing the yeast genome comprising more than 6,300 proteins (YPD) [2] and abundant interaction data have been produced by many research groups and many machine learning-based promising methods have been successfully applied to the analysis of these data.

Recently, there also have been many accomplishments in literature data mining for these kinds of biology applications. Many of these applications focus on extracting gene/protein interactions that are scattered throughout the biomedical literature to

construct well-refined interaction databases of research results. Many research projects have been devised to collect information on gene/protein interactions and many databases have been constructed to store such data. However, most of data in these databases were accumulated manually which need high costs. Nevertheless, scientists continue to publish their discoveries on new gene/protein interactions and modified interactions of previous ones on various domains in scientific papers without submitting their results to specific public databases [3]. As a result, we can infer that most gene/protein interaction information still exists only in these papers.

Even though there are many databases for gene/protein interactions, most such data still exist only in the biomedical literatures. They are spread in biomedical literature written in natural languages and they require much effort such as data mining for constructing well-structured data forms. As genomic research advances, furthermore, knowledge discovery from a large collection of scientific papers is becoming more important for efficient biological and biomedical researches.

Thus, how to extract gene/protein interactions from biomedical literature has been an active research subject. Many approaches widely adopted natural language processing (NLP) techniques to resolve this problem. These methods can be regarded as parsing based methods. Both full and shallow parsing strategies have been attempted.

Yakushiji *et al*. [4] used a general full parser with grammars for biomedical domain to extract interaction events by filling sentences into augmented structures. Park *et al*. [5] proposed bidirectional incremental full parsing with combinatory categorical grammar (CCG) which localizes target verbs and then scans the left and right neighborhood of the verb respectively to find interaction events in the sentences. Temkin *et al*. [6] also utilized a lexical analyzer and context-free grammar (CFG) to extract gene, protein and small molecule interactions with a recall rate of 63.9% and precision rate of 70.2%. As a similar method, preposition-based parsing to generate templates also proposed by Leroy *et al*. [7] and they achieved precision of 70% for biomedical literature abstract processing. For a partial parsing method, Pustejovsky *et al*. [8] used the relational parsing for the inhibition relation with recall rate of 57%. But, these methods are inherently complicated, requiring many resources, and performances are not satisfactory.

In this paper, we present a relation kernel based interaction extraction method to resolve these interaction extraction problems. In the proposed method, we extract gene/protein interactions of *S.cerevisiae* from documents. To doing this, we define kernel for gene/protein relation extraction, so called '*relation kernel*', with predefined interaction corpus and set of interaction patterns. Proposed kernel method for interaction extraction only exploits shallow parsed documents without full parsing efforts.

This paper is organized as follows. In Section 2, the basic concept of kernel method and its types are described. In Section 3, relation kernel for protein-protein interaction extraction is described. In Section 4, we show experimental results of interaction extraction. Finally, Section 5 we present concluding remarks and future directions.

2   Kernel Method

An object can be transformed into a collection of features $f_1,...,f_N$ which produce a N-dimensional feature vectors. In many cases, however, it is difficult to express data via

features. In most NLP problems, for example, feature based representations produce inherently local representations of objects and it is computationally infeasible to generate features involving long-range dependencies [9].

Kernel methods are an attractive alternative to feature-based methods. Kernel methods retain the original representation of objects and use the object in algorithms only via computing a kernel function between a pair of objects. A kernel function is a similarity function which has certain properties. That is, kernel function K over the object space X is binary function $K: X \times X \rightarrow [0, \infty]$ mapping a pair of objects x, y X to their similarity score $K(x, y)$. This is embedding procedure of data items (e.g., genes, proteins, molecular compounds, etc.) into a vector space F, called a *feature space*, and searching for linear relation in such a space. This embedding is defined implicitly, by specifying an inner product for the feature space via a symmetric and positive semidefinite *kernel function*: $K(x, y) = \langle \Phi(x), \Phi(y) \rangle$, where $\Phi(x)$ and $\Phi(y)$ are the embeddings of data items x and y [9].

It can be shown that any kernel function implicitly calculates the dot product of feature vectors of objects in high-dimensional feature spaces. That is, if there exist features $f(\cdot) = (f_1(\cdot), f_2(\cdot),\ldots)$, then $f_i: X \rightarrow R$, so that $K(x, y) = \langle f(x), f(y) \rangle$. Conversely, given features $f(\cdot) = (f_1(\cdot), f_2(\cdot),\ldots)$, a function defined as a dot product of the corresponding feature vectors is necessarily a kernel function [9].

In many cases, it may be possible to compute the dot product of certain features without enumerating all the features. An excellent example is that of subsequence kernels [10]. In the subsequence kernels, the objects are strings of characters, and the kernel function computes the number of common subsequences of characters in two strings, where each subsequence match is additionally decreased by the factor reflecting how spread out the matched subsequence in the original sequences [10]. Despite of exponential number of features (subsequences), it is possible to compute the subsequence kernel in polynomial time. Therefore, one can exploits long-range features in string without enumerating the features explicitly.

There are a number of learning algorithms that can operate only using the dot product of examples. The models produced by the learning algorithms are also expressed using only dot product of example. Substituting a particular kernel functions in place of dot product defines a specific instantiation of such learning algorithms. The algorithms that process examples only via computing their dot products are sometimes called *dual* learning algorithms.

The Support Vector Machine (SVM) is known as a learning algorithm that not only allows for a dual formulation, but also provides a rigorous rationale for resisting overfitting. For a kernel-based algorithms working in extremely rich feature spaces, it is crucial to deal with the problems of overfitting. Many experimental results indicate that SVM is able to generalize very well and avoid overfitting in high dimensional feature spaces. Thus, here we use SVM and relation kernel (extended version of subsequence kernel for shallow parsing) to extract gene/protein interactions.

3   Relation Kernel for Gene/Protein Interaction

Here we use the basic notion and the idea of subsequence kernel and extend it to operate on shallow paring for gene/protein relation extraction according to the work of Lodhi *et al*. [9] and Dmitry *et al*. [10].

We can consider the sentence, "In those pathways, usually the yeast protein RAP1 suppresses the up-regulation of SH3." We can generate appropriate parse tree for this sentence and compare with the other parse trees generated from other sentences by conventional shallow parsing. The form of target relation that we want to extract in this example sentence is 'RAP – suppress – SH3' which generally has the form of 'entity 1 – action verb – entity 2.'

In this paper, we use shallow parsing only for identifying its key elements such as entities and interactions instead of providing full interpretation of the sentence because this approach may provides us a fairly robust shallow parsing and the generation ability of structured representations even for ungrammatical sentences.

Next, we convert the shallow parse tree into examples for the '*entity–interaction–entity*' relation. This relation describes about the gene/protein interactions such as gene–gene, gene–protein, and protein–protein interactions. The type of the former and the later entity includes those names of gene and protein, and the type 'relation' holds between a two 'entity' and an 'interaction.' For each example sentence, we check whether the sentence has complete relation structure which has two entities and one interaction verb structures.

Here we use the term 'example' for parse tree of sentence. Each node of example has its 'type' and 'role.' We define *Type* = {φ, GeNP, GeN, PrNP, PrN, VerbPos, VerbNeg, VerbUnk, PNP, NP, Prep, Adj, Det} and *Role* = { φ, Entity_1, ActionVerb, Entity_2, UNK}. The GeNP and GeN of 'type' is for noun phrase or noun which represents a name of gene or protein. Verbs are classified into three categories; positive (e.g., activate), negative (e.g., inhibit), and other unknown action verbs (VerbUnk). Generally, 'Type' represents the category of par-of-speech (POS) tagged results. On the other hand, 'Role' represents the role of 'type' in the sentence to construct the structure of gene/protein interaction such as '*entity–interaction–entity*.'

For two relation examples P_1, P_2, we use the similarity function $K(P_1, P_2)$ in terms of similarity function of the parent nodes. That is,

$$K(P_1, P_2) = \begin{cases} 0, \text{ if } t(P_1.p, P_2.p) = 0 \\ k(P_1.p, P_2.p) + K_c(P_1.c, P_2.c), \text{ otherwise} \end{cases} \tag{1}$$

where, $t(\cdot, \cdot)$ is a matching function and $k(\cdot, \cdot)$ is a similarity function on nodes. The matching function has '1' if two nodes have same type and role in the parsing tree and '0', otherwise. The similarity function, $k(\cdot, \cdot)$, has '1' if the two nodes have same text and '0', otherwise [9].

We also use similarity function K_c of the children node of parse tree in terms of similarities of children subsequences. That is,

$$K_c(P_1.c, P_2.c) = \sum_{\mathbf{i}, \mathbf{j}, l(i)=l(j)} \lambda^{d(\mathbf{i})+d(\mathbf{j})} K(P_1[\mathbf{i}], P_2[\mathbf{j}]) \prod_{s=1,\dots,l(\mathbf{i})} t(P_1[i_s].p, P_2[j_s].p) \tag{2}$$

where, $\mathbf{i}$ denote a sequence of indices and $d(\mathbf{i}) = l(\mathbf{i}) = i_n - i_1 + 1$ denote length of the sequence $\mathbf{i}$. In the Equation 2, the term $P[\mathbf{i}]$ represent the sequence of children $[P[i_1],\dots,P[i_n]]$ and $K(P_1[\mathbf{i}], P_2[\mathbf{j}])$ stand for summing up $K(P_1[i_s], P_2[j_s])$ for all $s = 1,\dots,l(\mathbf{i})$ [9].

By using these definitions, Equation 2 consider all subsequences of relation example children with matching parents and it accumulates the similarity for each subsequence by adding similarities of the corresponding child examples. The equation also

reflects the amount of spread of the subsequences within children sequences by decreasing overall similarity with the factor λ which has the value from 0 to 1.

Generally, the index (**i** and **j**) of two example parse tree will be different. So, we can't use Equation 2 for similarity function K_c of the children node of parse tree directly. Thus, here we use the derived recurrences of Dmitry *et al.* [9] and its original construction recurrences from Lodhi *et al.* [10] to calculate K_c. That is,

$$K_C = \sum_{q=1,\dots,\min(m,n)} K_{c,q}(m,n)$$

$$K_{c,q}(i,j) = \lambda K_{c,q}(i,j-1) + \sum_{s=1,\dots,i} t(P_1[s].p, P_2[j].p)\lambda^2 C_{q-1}(s-1,j-1,K(P_1[s],P_2[j]))$$

$$C_q(i,j,a) = aC_q(i,j) + \sum_{r=1,\dots,q} C_{q,r}(i,j)$$

$$C_q(i,j) = \lambda C_q(i,j-1) + C'_q(i,j) \qquad\qquad (3)$$

$$C'_q(i,j) = t(P_1[i],P_2[j])\lambda^2 C_{q-1}(i-1,j-1) + \lambda C'_q(i,j-1)$$

$$C_{q,r}(i,j) = \lambda C_{q,r}(i,j-1) + C'_{q,r}(i,j)$$

$$C'_{q,r}(i,j) = \begin{cases} t(P_1[i],P_2[j])\lambda^2 C_{q-1,r}(i-1,j-1) + \lambda C'_{q,r}(i,j-1) & \text{if } q \neq r \\ t(P_1[i],P_2[j])\lambda^2 K(P_1[i],P_2[j])C_{q-1}(i-1,j-1) + \lambda C'_{q,r}(i,j-1) & \text{if } q = r \end{cases}$$

where, for the number of children of P_1 and P_2, condition $m \geq n$ is assumed. The boundary conditions of calculation of K_c are as follows:

$$K_{c,q}(i,j) = 0, \text{ if } q > \min(i,j), \qquad C_q(i,j) = 0, \text{ if } q > \min(i,j)$$
$$C_0(i,j) = 1, \qquad\qquad\qquad C'_q(i,j) = 0, \text{ if } q > \min(i,j) \qquad (4)$$
$$C_{q,r}(i,j) = 0, \text{ if } q > \min(i,j) \text{ or } q < r, \quad C'_{q,r}(i,j) = 0, \text{ if } q > \min(i,j) \text{ or } q < r$$

4   Experimental Results

Data

We used 260 labeled MEDLINE records which have in any case one sentence containing at least one gene/protein interaction as positive examples (50% were used for training, 50% were used for testing). The negative examples, about 100 examples, were also collected from MEDLINE records which contain any interaction at all. The shallow parses of each record was generated by out custom shallow parsing system. We also used Brill tagger trained with GENIA biomedical corpus to guarantee more accurate POS tagging. Table 1 shows the statistics of these examples.

Table 1. Statistics of positive and negative examples of the corpus which was used for gene/protein interaction extraction task

Categories	# of Positive Relation	# of Negative Relations	Total
# of Relation	314	139	453

Algorithm Settings

For kernel learning, we used Support Vector Machine (SVM) [11] and LIBSVM algorithm implementation [12]. We set parameter   of kernel computation to 0.5 to reflect the amount of spread of interaction structure (gene names, interaction verbs,

and protein names). We evaluated kernel based relation extraxtion result in comparison with the general feature based baseline classifier, Naive Bayes. Full parsing results were used as a feature set for Naive Bayes. Here, these baseline model represent the method of full parsing based relation extraction (classification) approach.

Results

For information extraction problems, the system performance is usually reflected using the performance measure of information retrieval: *precision*, *recall*, and *F-measure*. Precision is the ration of the number of correctly predicted positive examples to the number of predicted positive examples. Recall is the ratio of the number of correctly predicted positive examples to the number of true positive examples. F-measure combines precision and recall as follows:

$$F = \frac{2 * recall * precision}{(recall + precision)} \tag{5}$$

Table 2 shows the precision, recall, and F-measure for the gene/protein relation extraction experiments.

Table 2. The precision and recall of experiments. *TP* is the number of correctly extracted true interactions and *TN* is the number of correctly extracted negative interactions in the test documents, and we calculated *Precision = TP/(TP+FP)* and *Recall = TP/(TP+FN)*. In this table, *K* is stand for kernel method and *N* is stand for Naïve Bayes. In this table, we showed the result for the major four general interaction verbs (The value of 'all verbs*' represent the total average value including the value of four interaction verbs)

Interaction Verb	Recall (%)		Precision (%)		Total *F* (%)	
	K	N	K	N	K	N
Interact	82.3	79.1	80.4	79.3	81.3	79.2
Bind	80.8	78.5	84.6	78.5	82.7	78.5
Up-regulate	76.2	72.1	82.8	74.2	79.4	73.1
Suppress	79.9	76.9	81.9	81.8	80.9	79.3
All verbs*	**78.3**	74.2	**79.9**	75.5	**79.1**	74.8

As we can see in Table 2, kernel based relation extraction outperform full parsing feature based Naïve Bayes method overall. The number of interaction verb which generally describes gene/protein interactions in the biomedical literature and recognized by this experiment was about 150. These recognized interaction verb include, for example, 'activate', 'abolish', 'accelerate', 'affect', 'alter', 'amplify', 'assemble', 'associate', 'bind', 'block', 'conjugate', 'control', 'down regulate', 'enhance', 'inactive', 'induce', 'infect', 'inhibit', 'interact', 'ligate', 'localize', 'mediate', 'modify', 'prevent', 'prohibit', 'phosphorylate', 'regulate', 'stimulate', 'suppress', 'transactive', 'ubiquitinate', 'upregulate', etc. Among these verbs some verb occurs frequently but other occurs rarely.

5   Conclusions

In this paper, we presented kernel based gene/protein relation extraction method and the result that kernel based method outperform conventional feature based (full pars-

ing information based) approach. Presented relation kernel only requires domain specific knowledge on defining matching and similarity kernel function. In this paper, we used the dictionary of gene/protein names and its aliases. So it is not difficult to expand this method to other domain by changing this simple domain knowledge.

But, this kernel method on NLP problem has one big disadvantage for practical use of algorithm because it requires high computational costs. Thus, the more efficient kernel computation method should be devised for more efficient kernel computation and easily usable functionality to the problems of various domains, and we'll try this as future works.

References

1. Deng, M., *et al*.: Inferring domain–domain interactions from protein–protein inter-actions. *Genome Res*. **12** (2002) 1540–1548.
2. Goffeau, A., *et al*.: Life with 6000 genes. *Science* **274** (1996) 546–567.
3. Huang, M., *et al*.: Discovering patterns to extract protein-protein interactions from full texts. *Bioinformatics* **20(18)** (2004) 3604–3612
4. Yakushiji, A., *et al*.: Event extraction from biomedical parsers using a full parser. In *Proc. of the 6th Pacific Symposium on Biocomputing* (PSB 2001). (2001) 408–419.
5. Park, J.C., *et al*.: Bidirectional incremental parsing for automatic pathway identification with combinatory categorical grammar. In *Proc. of the 6th Pacific Symposium on Biocomputing* (PSB 2001). (2001) 396–407.
6. Temkin, J.M., *et al*.: Extraction of protein interaction information from unstructured text using a content-free grammar. *Bioinformatics* **19(16)** (2003) 2046–2053.
7. Leroy, G., *et al*.: Filling preposition-based templates to capture information from medical abstracts. In *Proc. of the 7th Pacific Symposium on Biocomputing* (PSB 2002). (2002) 350–361.
8. Pustejovsky, J., *et al*.: Robust relational parsing over biomedical literature: extracting inhibit relations. In *Proc. of the 7th Pacific Symposium on Biocomputing* (PSB 2002). (2002) 362–373.
9. Dmitry, Z., *et al*.: Kernel methods for relation extraction. *Journal of Machine Learning Res.* **3** (2003) 1083–1106.
10. Lodi, H., *et al*.: Text classification using string kernels. *Journal of Machine Learning Res.* **2** (2002) 419–444.
11. Cortes, C, *et al*.: Support-vector networks. *Machine Learning* **20(3)** (1995) 273–297.
12. LIBSVM, http://www.csie.ntu.edu.tw/~cjlin/libsvm/

Combining an Order-Semisensitive Text Similarity and Closest Fit Approach to Textual Missing Values in Knowledge Discovery*

Yi Feng, Zhaohui Wu, and Zhongmei Zhou

College of Computer Science, Zhejiang University, Hangzhou 310027, P.R. China
{fengyi,wzh,zzm}@zju.edu.cn

Abstract. The ubiquity of textual information nowadays reflects its great significance in knowledge discovery. However, effective usage of these textual materials is always hampered by data incompleteness in real-life applications. In this paper, we apply a closest fit approach to attack textual missing values. To evaluate the closeness of texts in this application, we present an order perspective of text similarity and propose a hybrid order-semisensitive measure, M-similarity, to capture the proximity of texts. This measure combines single item matching, maximum sequence matching and potential matching and get a proper balance between usage of sequence information and efficiency. We incorporate M-similarity into two closest fit methods to missing values in textual attributes and evaluate them on data sets of Traditional Chinese Medicine (TCM). Experimental results illustrate the effectiveness of these methods with M-similarity.

1 Introduction

In the age of information explosion, knowledge discovery, also known as data mining, is drawing increasing attention in recent years. Among tons of data newly produced everyday, the majority is unstructured (such as plain text) or semi-structured (such as html, and xml). For access to these materials, the texts are stored in databases, separated by columns. Typically, the content of each column is considered as an unstructured text. Besides, a great number of databases keep information or knowledge of specified field, of which a large portion of attributes are also textual. However, the effective usage of textual attributes in knowledge discovery is always hampered by the incompleteness of data in real-life applications. Thus, it is of great necessity to attack textual missing values. This is the motivation behind our research.

   Typical methods to missing value problem include mean imputation, mode imputation and imputation by regression. However, these approaches are more applicable to structured data than unstructured text. The mean or mode of large numbers of texts is generally not meaningful for imputation due to the diversity of textual expressions and regression is usually applied to numeric or categorical data. In this paper, we apply a closest fit approach to attack this problem. This approach was firstly presented by Jerzy W. Grzymala-Busse et al[6] to fill in missing values in structured data and we extend this method to unstructured text.

In closest fit approach to textual missing values, a fundamental issue is how to evaluate the "closeness" of two given texts. The problem of text comparison is closely related to research areas like string matching, information retrieval and text mining. To incorporate text similarities in these different disciplines into one framework, we present an order perspective of text similarities in this paper. Under this perspective, we present a new order-semisensitive similarity, M-similarity, to adapt to missing value application. This similarity is based on three factors: (1) single item matching. (2) maximum sequence matching. (3) potential matching. Besides, three changeable parameters are included in M-similarity to make it more flexible.

Our lab has collaborated with that China Academy of Traditional Chinese Medicine (CATCM) since 1999 and developed a unified web accessible multi-databases query system of Traditional Chinese Medicine (TCM) bibliographic databases and specific medical databases in recent years [3]. We find the majority of data in TCM databases is semi-structured text. However, a great number of textual values in databases are missing, which prevents effective knowledge discovery. To evaluate how closest fit approach with M-similarity behaves to textual missing values in knowledge discovery, experiments on real-life TCM databases are carried out. Empirical results demonstrate the effectiveness of this method with M-similarity.

The rest of the paper is organized as follows. Section 2 gives a description of the order perspective of text similarities. Section 3 proposes our definition of M-similarity. In Section 4 this new similarity is incorporated into the closest fit methods to handle missing values in unstructured text. Section 5 provides experiments to apply this approach to TCM databases. Finally, we conclude in Section 6.

2   An Order Perspective of Text Similarity

To measure the similarity between two texts, we could consider this problem in three categories, based on the role of item order within text on comparison. We define an *item* (character, word, or phrase) as the basic unit of text comparison. The granularity of item can be determined according to language and application. Some preprocessing procedures, such as stop-word elimination, stemming and WSD, could be carried out to achieve better comparison result. After preprocessing, three categories of text similarity could be calculated.

The first category can be named order-sensitive similarity. In this category, text is usually considered as string, more precisely, string of items. To derive string similarity, a process named string alignment, or string matching, is carried out. As the literal meaning of string alignment implies, the item order within text is strictly preserved during text matching. That is, matching is carried out item by item as a pairwise alignment process. Similarities in discipline of string matching, such as edit distance[1], Hamming distanc[5], and LCS[7], can fall into this category. Most of above distances used in string alignment mainly focus on character in string. However, we could easily extend these similarities to different item level.

The second category, order-insensitive similarity, doesn't focus on matching at string level. Instead, only item co-occurrence is considered, neglecting item order within text. In discipline of information retrieval and text mining, VSM [2] and its variants are usually applied to discriminate texts more effectively, during which the information of sequence in texts is lost.

The third group is order-semisensitive one. This category differs from others in three features: (1) Block of items is used as comparison unit. (2) Only local order (the item order within each block) is preserved during matching. (3) The size of each item block is dynamically determined. Existing order-semisensitive similarities include block-moving approach presented by Tichy [8] and block edit models presented by Daniel Lopresti et al[4]. These distances are rooted on the idea of edit distance. However, in large numbers of real-world applications, such as TCM, the majority of items in text have special meaning, while the order of item in text is not fully determinant for text comparison. Thus, edit model is not so appropriate in this situation. To capture the similarity in this background more precisely and efficiently, a new order-semisensitive similarity is presented in section 3 of this paper.

3   An Order-Semisensitive Similarity

We propose a new hybrid similarity, named M-similarity, for order-semisensitive text comparison. M-similarity comprises three factors: (1) Single item matching, i.e. matching of basic unit. Unit can be character or word, depending on language feature. (2) Maximum sequence matching, i.e. multiple matching of longest substrings. In LCS distance [7], only the longest common subsequence is accounted for, neglecting other matching of substrings. In our algorithm, we find common substring as long as possible in every time of matching. (3) Potential matching. This factor considers about how possible two texts of different length will be similar when the shorter one is expanded to the length of longer one. Unlike in dynamic programming of edit distance, gap is not inserted in potential matching, which decreases the computational complexity of text comparison. To calculate text similarity involving above three factors, we define three parameters as follow:

Definition 1. W_s **: Weight of single item matching.** Every exact single item matching is taken into account during comparison, weighted by W_s. However, continuous matching with length of more than one always demonstrates more similarity between two texts. Thus we have the following weight:

Definition 2. W_{ec} **: Extra weight of continuous matching.** For every item matching in continuous matching, we consider not only the single item matching weight W_s, but also the extra weight W_{ec}. Maximum sequence matching is weighted by combining W_s and W_{ec}.

Definition 3. P_{pm} **: Possibility of two texts of different length would be totally same when the shorter one is expanded to the length of longer one.** In many databases of specified domain, the content of textual attributes for some records is input partly, i.e. incomplete. Given two texts of different length in these textual attributes, if we make the content of shorter text more complete by expanding it to length of longer one manually or computer-aidedly (by some missing value approaches, etc.), it is probable that the padding content of originally shorter text is similar to the longer text. P_{pm} represents this possibility and takes value from 0 to 1.

Based on W_s, W_{ec} and P_{pm} defined above, we could propose our definition of M-similarity. Let us assume that S is the shorter text with length N, L is the longer text with length M. Thus, M-similarity is defined as below:

$$\text{M-Similarity (S, L)} = \frac{\sum_N \lambda_{lm}(S, L, currentpos) + \lambda_{pm}(S, L)}{M \times W_s + (M-1) \times W_{ec}} \tag{1}$$

Here, $\lambda_{lm}(S, L, currentpos)$ is defined in (2), representing sum of weights for longest sequence matching at given position $currentpos$ within S. Meanwhile, sum of weights for potential matching is represented by $\lambda_{pm}(S, L)$, defined in equation (3).

$$\lambda_{lm}(S, L, currentpos) = \theta(S, L, currentpos) \times W_s \\ + [\theta(S, L, currentpos) - 1] \times W_{ec} \tag{2}$$

$$\lambda_{pm}(S, L) = [(M-N) \times W_s + (M-N-1) \times W_{ec}] \times P_{pm} \tag{3}$$

Given current position $currentpos$ within S, $\theta(S, L, currentpos)$ functions like this: it searches L from left to right for longest possible matching with substring in S starting from $S[currentpos]$, and returns the longest matching length as return value. We count each item in continuous matching by adding one W_s for each item and measure the effect of continuous matching by adding W_{ec} of $\theta(S, L, currentpos)$ - 1 times, constituting $\lambda_{lm}(S, L, currentpos)$. Then we move $currentpos$ to the next position where the value of $S[currentpos]$ can be found in L and calculate $\theta(S, L, currentpos)$ again. After scanning S from left to right like this, we obtain a cumulated score of $\lambda_{lm}(S, L, currentpos)$ that measure the degree of single item matching and maximum sequence matching. To take potential matching into account, we further add a factor $\lambda_{pm}(S, L)$ combining P_{pm} and the difference of text length. Divide this final score by the weight sum of total matching of length M, we obtain the final M-similarity between S and L, ranging from 0 to 1.

The next question is how to calculate $\theta(S, L, currentpos)$. The strategy we use is that as soon as we find the first single item matching, we proceed to check its rightly adjacent item. If this item matches again, we check the right item adjacent to current one further, and further. By this way, we could find the longest continuous matching. The consequence, however, is that we probably can't get global longest substring matching, since global optimum is likely to emerge elsewhere. If the source text is of great length and considerable item replications, we could extend to scan whole length of L and find global maximum matching. Nevertheless, in our TCM database, textual attributes are always of short length and few item replications. Thus, we can trade some accuracy for efficiency by considering only local optimum. Note that, the three adjustable parameters make M-similarity applicable to different situations. For example, we could set both W_s and P_{pm} to zero in cases where only continuous matching reflects closeness.

In the end of this section, we provide an example to calculate M-similarity: let *A, B,* … represent different items in texts, given two texts *S=ABCDEFG*, *L=DABHEFGR*, and parameters W_s =0.8, W_{ec} =0.2, P_{pm} =0.1, then *M-Similarity (S, L)* = [(2*0.8+0.2+0.8+3*0.8+2*0.2) + (8-7)*0.8*0.1] / (8*0.8+7*0.2) = 0.7026

4   Closest Fit Approach with M-Similarity

To attack missing value problem in unstructured background, we apply a closest fit method. W. Grzymala-Busse et al. [6] firstly used this approach to missing values in structured data. In closest fit methods, we search the training set, compare each case with the one having missing value, calculate similarity between them, and select the case with largest similarity. Then, we substitute the attribute value of that case for the missing attribute value. Two versions of closest fit methods could be applied. The first is to search the best candidate in entire data set, which can be named *global closest fit*. The second is to restrict search space to all cases belonging to the same category, which can be named *concept closest fit*, or *category closest fit*.

There are two ways that M-similarity could be incorporated into the closest fit approach. Firstly, when computing the similarity between the scanned case and the one with missing value, we could apply M-similarity to measure proximity between two textual attribute values. Secondly, to evaluate the accuracy of a missing value method, we always assume certain attribute values to be missing. After certain approach to missing data is carried out, a case with highest resemblance is selected and the textual attribute value of that case is used to replace the assumptive missing attribute value. Then, M-similarity could be utilized to evaluate how close the substituted text is with original one. In brief, the role M-similarity plays in closest fit approach is an essential measure to evaluate how "close" two given texts are, either in the process of candidate searching or result evaluation.

5   Experiment Results and Evaluation

Data sets used for our experiments come from CATCM. One of databases widely used in TCM is Database of Chinese Medical Formula (DCMF). As a popular medication, traditional CMF plays an important part in the health undertakings for Chinese people for several thousand years. This database contains 85988 records of CMF derived from more than 700 ancient medical books. Two crucial attributes of DCMF are ingredients and efficacy. Ingredients contains the effective components constituting formula, and Efficacy is a textual attribute describing TCM remedy principle. One of the biggest problems hampering the effective usage of DCMF is the incompleteness of data. In fact, 81.76% of data in attribute efficacy is missing. This phenomenon motivates us to develop method capable to fill in the missing data. We apply the approach presented in section 4 to handle this problem. To evaluate the accuracy of closest fit method, another data set named DCMF-2 is used. DCMF-2 is a selected version of Database of Currently Application of Herbal Medicine Formula. It contains 4646 CMF derived from modern publications. Both the attribute ingredients and efficacy are not missing for each record. Besides, DCMF-2 has an attribute indicating the category of formula, enabling the category closest fit approach.

In theory of TCM, the closeness of formula ingredients reflects the resemblance of efficacy. The attribute *ingredients* in both DCMF and DCMF-2 are semi-structured text. A system named *ingredients splitter* is developed by us in 2003 to effectively extract components of formula from this attribute. Thus, Jaccard coefficient is calculated to measure the degree of overlap between two sets of structured nominal ingredients derived from *ingredients splitter*. Here, we define the similarity between two

ingredients as the Jaccard coefficient multiplied by 100, that is, a value between 0 and 100. To fill in an assumed missing efficacy of certain formula, we search for the formula having the greatest similarity of ingredients with that case, and substitute its efficacy for missing one. M-similarity between the substitute and original efficacy is calculated subsequently to perform quantitative evaluation. The greatest similarity of ingredients mentioned above can be abbreviated as GSI for later reference.
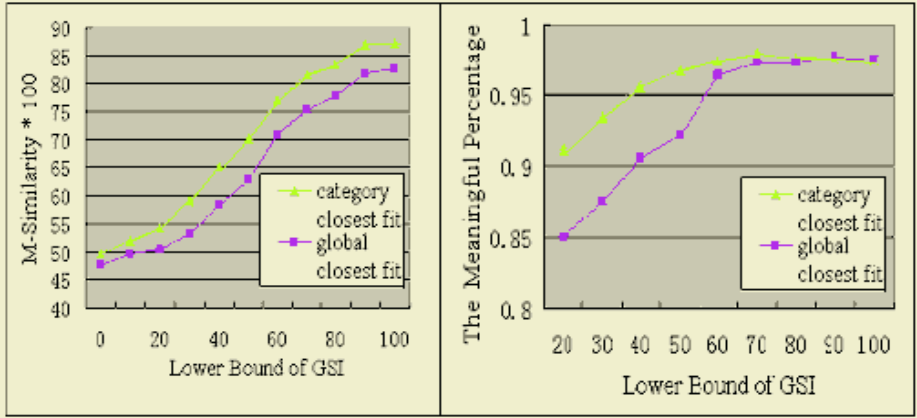

Fig. 1. Average M-Similarity (*100) for different scopes of data in DCMF-2 (Left) and The meaningful percentage of case processed by closest fit approach to DCMF-2 (Right)

Both the global closest fit approach and category closest fit one are carried in our experiments. By analyzing the characteristic of TCM text and tuning parameter empirically, we set W_s, W_{ec} and P_{pm} to 0.8, 0.2 and 0.1 respectively. In other applications, these parameters can also be tuned by experiments (we will present it in another paper). The experimental results are illustrated in Figure 1. Left chart shows how the average M-similarity (*100) changes as a function of selection scope of records. We can see that as the increase of lower bound of GSI, average M-similarity increases as well. This exhibits a phenomenon that the more similar of ingredients for two formulae, the more similar of their efficacies. Besides, we can conclude that category closest fit approach behaves better than global closest fit for each selected subset of DCMF-2, averagely exceeding 0.0511 in M-similarity.

Unlike numerical and categorical attribute, textual attribute in data mining has one distinguishing feature: the diversity of expression. Due to this feature, our approach can't totally replace manual method. However, it can serve as a valuable reference to manual work. After carefully analyzing our result, experts of TCM point out, for missing value of efficacy, those cases with M-similarity no less than 0.2 are of great meaning for manual work. They also present that our method is reasonable only if GSI is no less than 20. Thus we calculate the following value as evaluation factor: *the meaningful percentage*, that is, the percentage of case with M-similarity no less than 0.2. Right chart in Figure 1 shows how the meaningful percentage varies along with the lower bound of GSI. Among 3705 formulae (79.75% records of DCMF-2) with GSI no less than 20, nearly 85% of results derived from global closest fit approach are meaningful. This percentage reaches as high as 91.15% for category closest fit. We

can also see from Figure 1 that the meaningful percentage increases with lower bound of GSI. The reason behind it is that with the increase of lower bound of GSI, the formula used to replace missing efficacy of certain case would have more similarity of ingredients with that case, which results in more effectiveness of closest fit method. Comparing results generated from category closest fit and global closest fit, again we can draw the conclusion that category closest fit generally outperforms global one.

6 Conclusion

In this paper we have presented an order perspective of text similarities. A hybrid order-semisensitive similarity, M-similarity, is proposed subsequently to measure the equivalence of two texts. This measure is incorporated into closest fit methods to handle missing value problem in textual attributes in knowledge discovery. Two approaches based on this idea, global closest fit and category closest fit, are applied to missing value problem in data sets of Traditional Chinese Medicine. Experimental results demonstrate the effectiveness of closest fit methods with M-similarity. Due to the feature of text, missing values in textual attributes need be tackled further.

By varying the granularity of item and three parameters, M-similarity can be used in different languages and applications, not only in Chinese, English, but also in Korean, Japanese, French, German, and so on. It can serve as an effective and flexible measure for text comparison in disciplines like NLP, knowledge discovery, and text mining. Future work includes developing a systematic strategy to specify parameters of M-similarity and utilizing this order-semisensitive measure in more applications.

References

1. V. I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. Doklady Akademii Nauk SSSR, 163(4), 1965.
2. A. Singhal. Modern information retrieval: a brief overview, In IEEE Data Engineering Bulletin 24(4), 35-43, 2001
3. Xuezhong Zhou, Zhaohui Wu, Wei Lu. TCMMDB: a distributed multidatabase query system and its key technique implemention, In IEEE SMC 2001, 1095-1100, vol.2, 2001
4. D. Lopresti, A. Tomkins, Block edit models for approximate string matching, Theoretical Computer Science, 181(1): 159-179, 1997
5. D. Sankoff, J. Kruskal. Timewarps, string edits, and macromolecules: the theory and practice of sequence comparison. Addison-Wesley, 1983
6. W. Grzymala-Busse et al. A comparison of three closest fit approaches to missing attribute values in preterm birth data, International journal of intelligent systems, 17, 125-134, 2002
7. S. Needleman, C. Wunsch, A general method applicable to the search for similarities in the amino acid sequences of two proteins. J. Mol. Biol. 48, 444–453, 1970
8. F. Tichy, The string-to-string correction problem with block moves, ACM Transactions on Computer Systems, 2(4): 309–321, 1984

Support for Internet-Based Commonsense Processing – Causal Knowledge Discovery Using Japanese "If" Forms

Yali Ge, Rafal Rzepka, and Kenji Araki

Graduate School of Information Science and Technology, Hokkaido University
Kita-ku Kita 14-jo Nishi 9-chome, 060–0814 Sapporo, Japan
{geyali,kabura,araki}@media.eng.hokudai.ac.jp

Abstract. This paper introduces our method for causal knowledge retrieval from the Internet resources, its results and evaluation of using it in utterance creation process. Our system automatically retrieves commonsensical knowledge from the Web resources by using simple web-mining and information extraction techniques. For retrieving causal knowledge the system uses three of specific several Japanese "if" forms. From the results we can conclude that Japanese web pages indexed by a common search engine spiders are enough to discover common causal relationships and this knowledge can be used for making Human-Computer Interfaces sound more natural and interesting than while using classic methods.

Keywords: commonsense, causal knowledge discovery, human-computer interface.

1 Introduction

1.1 Need for Commonsense Retrieval

As it is easy to notice, the amount of accessible information increases with tremendous speed together with rapid growth of the Internet accessibility. More and more users write their blogs which say what Mr. Public did day by day but do not include information which other computer scientists would need for their machines to perform some task. For our approach - to make a system searching for data obvious for a human and are completely unknown for machines - these everyday tasks described in blogs are the clue. The "commonsense" retrieval never becomes an object of search queries as humans do not need to seek for such knowledge, they gather it through all their lives. However, the computers lack it and this is one of the reasons why people do not treat machines as intelligent partners, especially when it comes to conversation.

1.2 State of Art

There are several research projects coping with gathering commonsense as CyC [1] or OpenMind Commonsense [2]. CyC contains over a million hand-crafted

assertions and OpenMind commonsense enabled construction of a 700,000 assertion commonsense knowledge base, gathered through a web community of collaborators. But they concentrate on manual or half-automatic processing and these projects are developed only for English language. As we assume that there is too much of such knowledge to be inputted by hand, we try to make this process automatic by using simple Web-mining techniques. Our laboratory members are quite successful on achieving OpenMind results without using any human input [3][4].

2   Japanese Language Predispositions

We already have shown [5] that Japanese language has very good predispositions for text-mining for commonsense processing mostly thanks to its particles, and for that reason we concentrate on Japanese WWW resources[1]. In previous step [5] we showed that it is possible to extract simple Schankian scripts from Japanese Internet resources, this time we will prove that the Web is a vast repository for commonsensical causations which can be used for example in talking systems.

2.1   The Particles

In Japanese language there is a set of particles changing noun/pronoun's character by simple addition to the right-side of the word. For example a noun *kuruma* (a car / cars) after adding a direction indicating particle *ni* (*kuruma-ni*) suggests that following verb will be directed to the car not the opposite, automatically decreasing number of verbs candidates which could follow this noun-particle structure. One can easily build a category of verbs connected to this particular noun or a category of nouns which are glued to one particular verb by the same particle. Most popular ones particles are *wa* (Topic-Indicating), *ga* (Linking-Indicating), *no* (Possessive-Indicating), *wo* (Object-Indicating), *ni* (Direction-Indicating), *de* (Place or Means of Action-Indicating), *to* (Connective) and *mo* (Addition-Indicating). However such research does not have to be restricted to Japanese, if the similar principles could be found, an application could work with other languages – for example by using prepositions in English or regular expressions for non-gender counting in Polish.

2.2   Japanese Conditional Clauses

The causal knowledge has been a research subject but rather rarely in the perspective of being a support for commonsense processing. Papers of Sato, Kasahara and Matsuzawa al. [6] has underlined the need of commonsense processing automatization and influenced several successors. One of the most related works

---

[1] This also helps us to avoid cultural background mismatches as commonsense vary from country to country even if their citizens speak the same English language.

[7] concentrated on Japanese "if" form *tame* and newspaper corpus while in our research we use the WWW as the corpus and different Japanese "if" forms (*to,tara,eba*) as main query keywords (Japanese "if" forms have many useful functions - *to* is If/When After-Indicating, *tara* is If/When-Indicating, *eba* is If-Indicating, *tame* is Cause/Purpose-Indicating, *toki* is Time-Indicating and *nara* is If / Special Case-Indicating).

3 Discovering Commonsensical Causations

By using the noun keyword together with "if" forms we automatically retrieve the causal knowledge about the inputted noun from the WWW. For example, when *water* is inputted and all the forms give the same results : counting *mizu wo nomu to / mizu wo nondara / mizu wo nomeba kimochi ii* we can be quite certain that "drinking water" causes "feeling nice".

Table 1. Examples for "if bear a child"

Particle	Effect	Usualness
eba	one/something is cured	7
to	woman changes	4
tara	leave the woman	3
to	fears decrease	3
to	put on weight	3
eba	body-line gets a bit out of order	2
to	woman gets determined	2
to	woman gets stronger	2
tara	quit one's job	2
tara	(home) becomes difficult to live	2
tara	cut off from the work	1
tara	work becomes more difficult	1
eba	population will survive	1
eba	the more one has them the life gets difficult	1
eba	the more one has them one gets younger	1

3.1 Our System

Previous Module. In the beginning of our research, we decided to work with nouns as keywords for collecting minimal Schankian scripts. For this purpose we extracted relation-oriented sentences for creating dictionaries as verb dictionary, noun dictionary and n-gram dictionaries using WWW corpus (1,907,086 sentences) retrieved with Larbin robot. The verbs and nouns dictionaries consist of 79,460 verbs and 134,189 nouns retrieved with help of ChaSen [8]. For creating scripts automatically, our system had to search for the relationships between verbs and nouns and also between verb pairs. In that step, we used the verbs and nouns which had the highest occurrence (which we call "Usualness"

after Rzepka et al. [9]), as they are used by human every day and are often used in our everyday lives, for example *television, movie, food*. Also this time we experimented mostly on the daily-usage nouns.

Architecture of Current Module. Basically, the latest module for discovering commonsensical causations can be summarized into the following processing steps:

a) The user inputs a noun as a keyword;
b) The system uses our web-based corpus for frequency check to retrieve 3 most frequent verbs following the keyword noun;
c) The most frequent particle between noun keyword and 3 most frequent verbs is discovered;
d) Forms of 3 most frequent verbs are transformed into three "if" forms;
e) By using Yahoo Japan resources, the system checks if the noun-particle unit occurs with the new verb forms unit;
f) If yes - sentences which include the conditional clause are saved;
g) With help from ChaSen analyzer the system gets the most frequent commonsensical causations as following:

$$\mathbf{Ms} = \mathbf{N} + \mathbf{P}_{\max} + \mathbf{V}_{\max} + \mathbf{I} + \mathbf{S}_{\max}$$

N:*Noun keyword;*
$P_{\max}$: *the most frequent particle joining noun and verb;*
$V_{\max}$: *most frequent verb occurring after the N;*
I: *"if" form;*
$S_{\max}$: *the most frequent commonsensical causation appearing after the condition;*

4   Retrieval Experiment

As the results are very big amounts of data, for our experiment we decided to retrieve only three most frequent commonsensical causations inputting only six nouns used in everyday life (*child, cigarette, water, room, food, bath, money, mobile phone, car, light and subway*) as shown in this example:
KEY: child – if(*to/tara/eba*) –
- to bear: if/when a child is born/baby – then – ...(see Table 1)
- to raise: if/when one raises a child/baby – then – ...
- to have: if/when one has a child/baby – then – ...

In average, our system was able to discover 130 casual knowledge units (usually about 20 from retrieved 150 units was excluded because of doubling the meaning or having errors) for one verb, therefore after limiting verbs to three it was 3x130=390 units for one noun which were retrieved in approximately 110 minutes. In total it gives 6x390=2,340 units achieved. By only increasing number of noun keywords to 60 and most frequent verb limit to 10 it is easy to
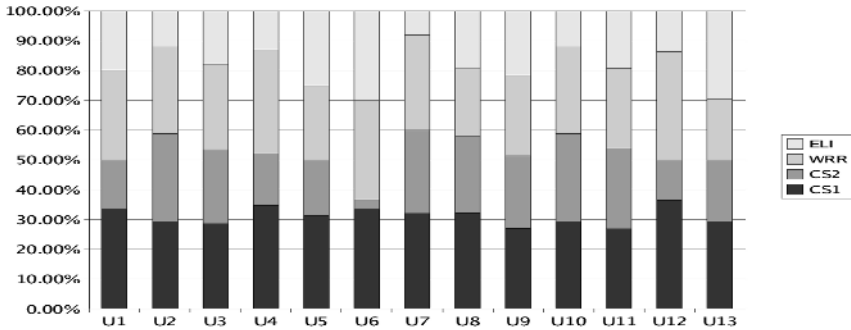
Fig. 1. Example: Degree of naturalness of utterances about a "meal" made by four systems

calculate that the system would achieve approximately 10x130x60=78,000 units which are almost two times more than Inui et al. achieved in their work [7]. We claim that in our case the units are more common as the Internet gives us a possibility to find information about usual, everyday life situations which are never described by newspapers used by Inui. For example a noun *water* gives a machine knowledge that is obvious for a human as "drinking water make body colder" or "filling with water can cause overflow".

5   Applications Experiment

In order to see user's perception of the basic commonsense knowledge included in a utterance, we performed a set of experiments basically using three kinds of utterances following input with one *keyword* from the previously mentioned set:

– ELIZA's output [ELI] (input sentence structure changing to achieve different outputs)
– WWW random retrieval output [WRR] (a shortest of 10 sentences retrieved by using *keyword* and query pattern "did you know that?")
– WWW commonsense retrieval output "high" [CS1] (sentences using common knowledge of highest usualness (most frequent mining results)
– WWW commonsense retrieval output "low" [CS2] (sentences using common knowledge of the lowest usualness (least frequent mining results).

Typical ELIZA [10] answer is "why do you want to talk about smoking" if the *keyword* is "smoking". For the same *keyword* WRR retrieved a sentence "did you know that people wearing contact lenses have well protected eyes when somebody is smoking?". An example of CS1 is "you will get fat when you quit smoking" and CS2 is "smoking may cause mouth, throat, esophagus, bladder, kidney, and pancreas cancers". We selected 10 most common noun keywords of different kinds (water, cigarettes, subway, voice, snow, room, clock, child, eye, meal) not

avoiding ones often used in Japanese idioms (voice, eye) to see if it influences the text-mining results. 13 referees were evaluating every set of four utterances in two categories – "naturalness degree" and "will of continuing a conversation degree" giving marks from 1 to 10 in both cases. The system comparison results proved that ELIZA does not eager users for continuing the chat but is still useful to keep the utterance naturalness. However, we proved that using commonsense even of the highest usualness is more natural than famous classic system (ELI 46%, CS1 54%). We also confirmed that query-based web-mining (WRR) results have slightly better user's acceptance than less common causal knowledge (CS2) which we find useful for creating a method for automatic category-based query formation depending of user's input.
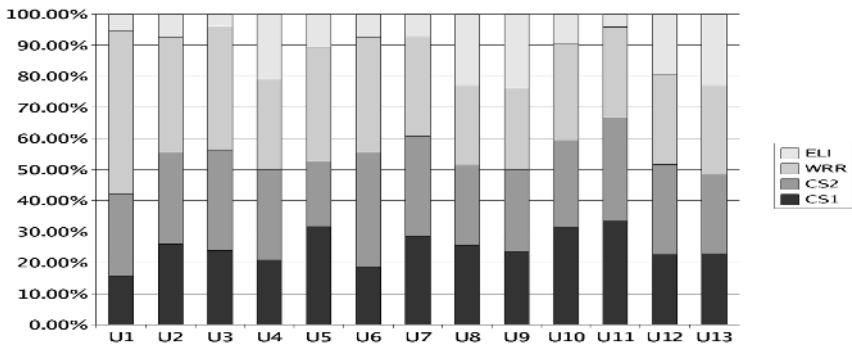

Fig. 2. Example: Degree of "continuation will" (the keyword used by four systems is "meal")

6   Conclusions and Future Work

In our first experiment, we showed how easily commonsensical causations can be discovered in enormous, mostly chaotic, data resources as WWW. There is remaining problem of time consumption but it is mostly due to the netiquette which does not allow for very fast retrieval within the search engine results. However, the commonsense processing in our future plans is supposed to work with an algorithm reducing causations by the context which will simplify query formation by increasing numbers of query keywords and making the search incomparably faster. It should also help to get rid of causation units' ambiguity, as the Internet brings also often contradictory statements like "drinking water makes you healthier" and "drinking water makes you sick". We do not have to assume that one of these claims is wrong - by discovering the contextual information we will become able to distinguish in which cases above mentioned statements are correct and in which, by contradiction, are not. In the application experiment we proved that this retrieved data can make a Human-Computer Interfaces sound more natural and interesting if we use opposite weights of commonsense

expressions. In this paper we have shown that three of several Japanese "if" forms which are *tara*, *to* and *eba* are useful for retrieving causal knowledge for commonsense processing and showed an example of such processing while creating utterances more natural than classic fully automatic methods as ELIZA which remains popular even if such approach requires laborious rules creation. We achieved higher naturalness without almost any labor and it is obvious that users prefer keep talking to systems based on the WWW that to these limited to their internal databases.

References

1. Lenat, D.: Common Sense Knowledge Database CYC, (1995) http://www.opencyc.org/, http://www.cyc.com/.
2. Singh, P.: The public acquisition of commonsense knowledge, Proceedings of AAAI Spring Symposium on Acquiring (and Using) Linguistic (and World) knowledge for Information Access. Palo Alto, CA: AAAI,(2002).
3. Skowron, M., Araki, K.: Voluntary Contributions of Unaware Internet Users? On Automatic Knowledge Retrieval from the WWW, to appear in AAAI 2005 Spring Symposium Report (Knowledge Collection from Volunteer Contributors (KCVC05), Stanford, California, March 21-23, 2005).
4. Skowron, M., Araki, K.: Automatic Knowledge Retrieval from the Web, to appear in Springer Verlag series "Advances in Soft Computing", (Intelligent Information Systems 2005. New Trends in Intelligent Information Processing and Web Mining Gdansk, Poland, June 13-16, (2005).
5. Ge, Y., Rzepka, R., Araki, K.: Automatic Scripts Retrieval and Its Possibilities for Social Sciences Support Applications To appear in Springer Verlag series "Advances in Soft Computing", (Intelligent Information Systems 2005. New Trends in Intelligent Information Processing and Web Mining Gdansk, Poland, June 13-16, (2005).
6. Sato, H., Kasahara, K., Matsuzawa, K.: Retrieval of simplified causal knowledge in text and its application. In Proc. of The IEICE, Thought and language (1998).
7. Inui,T., Inui,K., Matsumoto,Y.: What Kinds and Amounts of Causal Knowledge Can Be Acquired from Text by Using Connective Markers as Clues? DS 2003, LNAI 2843, (2003) 180–193.
8. Asahara, M., Matsumoto, Y., Extended Models and Tools for High Performance Part-of-Speech Tagger, COLING 2000,July (2000) 21–27.
9. Rzepka, R., Araki, K, Tochinai, K. Is It Out There? The Perspectives of Emotional Information Retrieval from the Internet Resources, Proceedings of the IASTED Artificial Intelligence and Applications Conference, ACTA Press, Malaga, 2002, pp. 22-27.
10. Weizenbaum,J.: ELIZA: A Computer Program for the Study of Natural Language Communication between Man and Machine, Communications of the ACM,9(1): pp. 36-45, 1966.

APForecast: An Adaptive Forecasting Method for Data Streams

Yong-li Wang[1,2], Hong-bing Xu[1], Yi-sheng Dong[1],
Xue-jun Liu[1], and Jiang-bo Qian[1]

[1] Department of Computer Science and Engineering,Southeast University
SiPaiLou No.2, Nanjing, 210096, China
`wyl_seu@126.com`, {`hbxu,ysdong,xuliu,jbqian`}`@seu.edu.cn`
[2] Department of Common Computer Teaching, Jiamusi University
Jiamusi 154007, P.R. China

Abstract. This research investigates continuous forecasting queries with alterable forecasting-step over data streams. A novel Adaptive Precision forecasting method to forecasting single attribute value of item in a single steam (stream-value), called AForecast, is proposed. The concepts of *dual slide windows* and *forecasting-steps conduction operator* are introduced. AForecast determines multiple forecasting-step based on the change ratio of stream-value and forecasts random-variant stream-value using relative precise prediction of deterministic components of data streams. Based on the theory of interpolating wavelet and optimal linear kalman filtering, this method can approximately generate optimal forecasting precision. Experiment results on actual power load data prove that this method can provide online accurate prediction on stream-value.

1 Introduction

In many modern applications, such as industry controlling and networking, as well as in new and emerging applications, like sensor networks and pervasive computing, data are commonly viewed as an infinite, possibly ordered data sequences rather than a finite data set stored on disk. These infinite (potentially) ordered sequences of data are referred to as *data streams*[1]. The characteristics of data streams processing, such as online continuous computing, one-pass scanning and limited time-space resources etc, put forward new challenges on traditional forecasting methods. For example, a fundamental objective of a power-system operating and controlling scheme is to maintain a match between the system's overall real-power load and generation. The Automatic Generation Control (AGC) loop addresses this objective by using system load and electrical frequency samples to periodically update the set-point power for key "swing" generators with a control sample rate ranging from 1 to 10 min. To improve performance, emerging AGC strategies employ a look-ahead control algorithm that requires real-time estimates of the system's future load out to several samples using a one to ten minute sample period (a traditional typical horizon of 30 to 120 min). An alterable forecasting-step strategy can provid a flexible control tracks that adapt to the time-variant real-power load.

The artificial Intelligence methods are suitable for forecasting periodic forecast. Its characteristic is higher precision and more expensive computing complexity. The extrapolate method's run-velocity is rapid but hard to forecast random variant non-linear component. Its forecast precision is low. If these two methods are applied to online forecasting separately, the forecasting results will not be satisfactory.

We believe that the mixed metheod that combines the precision advantage of complex forecasting methods and the velocity advantage of linear extrapolating forecasting methods is a good choice. We propose a forecasting method over data streams that can balance between speed and precision using an adaptive strategy, named Adaptive Forecasting model (AForecast briefly), in this paper. It use less forecasting points at the period when the stream is relatively smooth, and use more forecasting points at the period when the change of stream-value is relatively intensive. Experiment prove it can get a higher average forecasting precision with a low computation complexity.

2   Related Works

At present, the existing theories and methods about trend analyzing on data streams mainly focus on the analysis of similarity, abnormity and the difference of patterns [2]. But there are very few literature in forecasting stream-value. A linear regression analyzing and interpolating method used to forecast the instantaneous stream-value with w future window size is proposed in [3]. Data streams have also been treated as time series, and ideas from control theory are borrowed for the purposes of resource management. Kalman Filter is selected as a general and adaptive filtering solution for conserving resources in [4]. In power load forecasting domain, three practical techniques, Fuzzy Logic (FL), Neural Networks (NN), and Auto-regressive model (AR) for 30min load forecasting have been compared and discussed in [5]. It concludes that FL and NN can be good candidates for predicting the very short-term load trends on-line.

However, none of these methods mentioned above combines with the relatively stable forecasting information (forecasting-step≥1hr). Moreover, to the best of our knowledge, almost all methods forecast the future streams-value in a fixed forecasting-step manner. The fixed forecasting-step methods performs bad in cases of data streams with the characteristics of random change period.

3   Continuous Forecasting Query Model and Definition

A *stream* is an ordered sequence of items that resembles tuples in relation databases. In a data stream context, *stream-value* denotes a single attribute value of item in a single steam; *forecasting-step* denotes the interval between the actual measured stream-value and the forecasting stream-value. In research of power systems, *stream-value* denotes the *load-value*.

In actual applications, a discrete data stream $s(t)$ is composed of deterministic components $s_d(t)$ and stochastic components $s_r(t)$. The deterministic

components of data streams are primarily a result of the cause-effect dependence on measurable inputs such as time of day, day of week, temperature, etc. The stochastic components are primarily driven by the random variations in the acquiring environment. $s_d(t)$ is more easily to forecast accurately. Consequently, forecasting for stream-value is left to forecast $s_r(t)$.

In order to implement the continuous forecasting for the two components of data streams, we introduce the definition of *dual slide window*. A window of length I, called *I interval windows*, is used as the base unit for updating the deterministic components. A window of alterable length, called *alterable interval windows*, is used as the base unit for updating the stochastic components. The architecture of AForecast model is illustrated in Fig. 1.
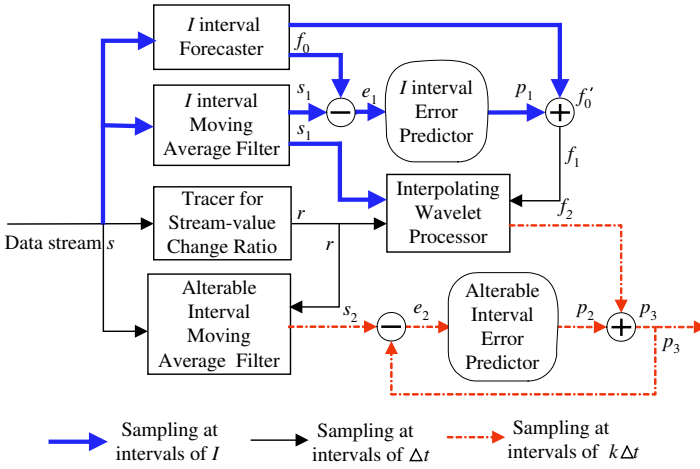

Fig. 1. Architecture of AForecast model

The AForecast model comprises five components:

1) **Moving Average Filter** is used to filter out the noise in the original data stream. It can be classified into *I interval* and *alterable interval*. It computes the average of stream-value in a slide window and regards the average as the new streaming measured value.

2) **I Interval Forecaster** is used to forecast the deterministic component of the data stream. There is a I intervals (I>k)between two deterministic stream-value forecast.

3) **Tracer for Stream-value Change Ratio** is used to map the different forecasting-steps for data streams. We define stream-value change ratio $\delta(t)=(s(t)-s(t-\Delta t))/\Delta t$ as a quantification of stream-value changes, where $\delta(t) \in [0,(\max(s(t))-\min(s(t)))/\Delta t]$.

4) **Interpolating Wavelet Processor** is used to generate more refined forecasting stream-value by binary scale interpolating wavelet method with multiple interpolating-resolution. $2^r \in [0, \lceil I/\Delta t \rceil]$ (r denotes the interpolating res-

olution) and as a result, the maximum binary scale interpolating resolution $r_{max}=\lceil \log_2 I/\Delta t \rceil$. To map the relation between changes of stream-value and interpolating resolution, we define forecasting-steps conduction operator D_t as $D_t(\delta)=\delta(t) \cdot r_{max}/(\max(s(t))\text{-}\min(s(t)))$.

5) **Error Predictor** is composed of I interval error predictor that is used to estimate the deterministic component of stream-value in a slow-moving manner and alterable interval error predictor that is used to estimate the stochastic component of stream-value in a quick-moving manner. Its objective is to optimally estimate the future value of random error.

The processing procedure is given as follows. Firstly, we use I interval forecaster to provide two forecasting stream-values f_0, f_0' at intervals of I. When current timestamps surpass I, we use e_1 ($e_1=s_1\text{-}f_0$, where s_1 denotes the I interval average stream-value) as the input of I interval error predictor and the estimate of error p_1 will be generated, As a result, the improved I interval forecasting stream-value f_1 ($f_1=p_1+f_0$') is obtained. Then we use interpolating wavelet processor to generate the interpolated values of forecasting value at intervals of alterable steps between s_1 and f_1 under the conduction of r generated by Tracer for stream-value change ratio. We use $e_2(e_2=s_2\text{-}f_2$', where s_2 denotes the alterable interval average stream-value, f_2' denotes the prior forecasting stream-value p_3) as the input of alterable interval error predictor and generate the estimate of alterable steps error p_2. Finally we construct the final forecasting stream-value $p_3(p_3=p_2+f_2)$.

4   Adaptive Interpolating Algorithm

We use the interpolating wavelet with dynamic forecasting-step conduction operator according to the variety ratio of stream-value instead of equal-interval spline interpolating. This method bears the merits of adaptability precision for forecasting stream-value. In order to satisfy the requirements of online processing and only one-pass scanning computation, we design an improved kalman-filtering algorithm with amnesia factor to estimate the forecasting error. For limitation of space, we omit the implementing details and only describe the key interpolating algorithm.

Definition 1. **Interpolating wavelet**. For $f \neg \in V_j$, we define a projection operator in V_j, which interpolates in stream-value samples $f(2^j k)$ in binary scale manner. The interpolating equation[7] is shown as follows:

$$P_{V_j} = \sum_{k=-\infty}^{+\infty} f(2_j k)\phi_j(t - 2^j k) .\tag{1}$$

Where P_{V_j} is P-1 order interpolating polynomial, ϕ is a interpolating function which produces a new interpolated value at different sampling intervals, $\phi_j(t\text{-}2^j k)(k \in Z)$ is a set of Riesz Base of ϕ's generating space, and j depends on forecasting-steps conduction operator $D_t(\delta)$.When the interpolating resolution shrinks doubly, we can get a much more refined interpolated-value of

projection $P_{V_j} f(t)$ in a adding "detail" manner, and this detail remedy the difference between $P_{V_j} f(2^j(k+1/2))$ and the middle sampled-point $f(2^j(k+1/2))$. $\psi_{j,k} = \phi_{j-1,2k+1}$ is the *interpolating wavelet* with resolution conduction.

Multiresolution interpolating algorithm is shown in Fig. 2.

Multiresolution interpolating Algorithm : Multires_Interpolating (*leftIpoint*, *rightIpoint*, *r*)
Input: left-end sampling-value *leftIpoint*, right-end sampling-value *rightIponit*, interpolating resolution
Output: wavelet interpolated-value of resolution *r*

(1) $l \leftarrow r$; /*initialize the loop variable*/

(2) $msapn \leftarrow \log_2((t_{rightIpoint} - t_{leftIpoint})/\Delta t)$;

(3) $t_{span} \leftarrow 2^{\lceil mspan \rceil} \Delta t$; /*compute interpolated span*/

(4) $P_{V_i} s(2^l(n+t_{span}/2)) \leftarrow \sum_{k=-m}^{+m} s(2^l k)\phi(n-k+t_{span}/2)$; /*compute the interpolated-value between sampling-value, corresponding to request the value of $\sum_{k=-m}^{+m} s(2^l k)h_i[n-k]$; invoke sub procedure to compute filter h_i */

(5) $s(2^l n) \leftarrow P_{V_i} s(2^l(n+t_{span}/2))$; /*make the current interpolated value act as the right end of the next interpolating course */

(6) $t_{span} \leftarrow t_{span}/2$; /*shrink interpolated span*/

(7) **if** $l > 0$, **then** $l \leftarrow l-1$, **goto** (1)

Fig. 2. Mutiresolution Interpolating Wavelet Algorithm

In order to effectively reduce the time cost for online forecasting, we use searching-list method to implementing interpolating expensive wavelet algorithm, i.e. store every interpolating coefficient in a linear list in advance. As a result we can locate the right coefficient based on index. The time complexity to generate every transitional interpolated-point is $O(1)$. The time complexity of computing forecasting-step based on the $\delta(t)$ is $O(1)$. The maximum level of forecasting-step (i.e. interpolating resolution) is $\log_2(I/\Delta t)$. The resulting time complexity of generating every interpolated-value is $O(\log_2(I/\Delta t))$.

5   Experiment Evaluation

In order to validate the performance of AForecast, we use actual load data of a power system in Nanjing. AForecast read a tuple at a fixed interval (10 millisecond) from the data set and send it to the model to simulate AGC system. *I interval forecaster* is implemented by weather-sensitive BP Neural Network. I is set to 1 hour. We also implement 15min level forecasting method (called 15mForecast) based on linear extrapolating and 5min level forecasting method (called 5mForecast) based on mixed forecasting technique [6], and carry out contrasting experiment on the average forecasting-precision and the average run-time. The experiments are conducted on a 2.66GHz Pentium machines with 256MB main memory and a 80GB hard disk. We implement the main part of

the algorithm in Visual C++ 6.0 on windows 2000 server and use the toolbox function of MATLAB to implement the other part of contrasting experiment algorithm.

Experiment 1: testing the overall average forecasting-precision of AForecast. We use the 36 hours load data during 2002/3/15-2002/3/16 in experiment. It takes 46.3ms to generate a small interval forecasting load-value averagely. The *average forecasting-precision* $\overline{p}$ ($\overline{p}$ =1- $\sqrt{1/n \cdot \sum_{i=1}^{n} p^2}$, $p = (v_i^f - v_i^a)/v_i^a$, where v_i^f is forecasting stream-value and v_i^a is actual measured stream-value) is 95.496 percent.
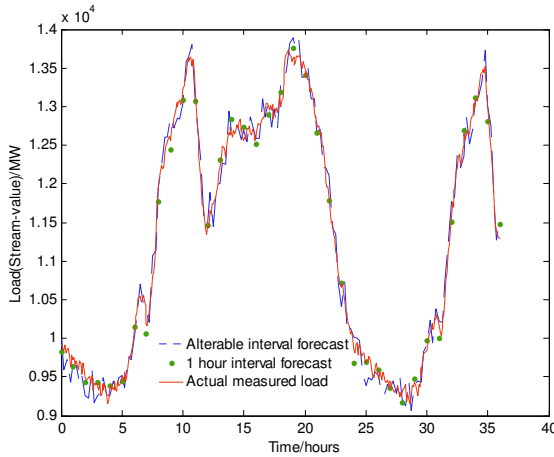
Fig. 3. 36 hours curve and discrete point of forecasting load and actual measured load

Fig. 3 shows an actual measured load curve, an alterable forecasting-step forecasting load curve and the discrete points for 1 hour forecasting load. Even as our expectation, forecasting-precision dose not obviously alter at the period of load-value intensely variation, because alterable forecasting-step forecaster can adapt to the variety of stream-value. The improved action of introduced Multiresolution interpolating strategy is validated. We can observe that the highest forecasting-precision occurs at a position after a new I interval (1 hour) forecasting stream-value generated, because I interval (1 hour) forecasting stream-value has relatively weak effect on small interval forecasting stream-value at this period. The forecasting-precision reduce gradually until I interval (1 hour) forecasting stream-value is updated. The shorter the distance between the position of actual measured stream-value, the higher the forecasting-precision is.

Experiment 2: testing the average forecasting-precision of AForecast with different forecasting-steps. we compared with the forecasting-precision of the other two algorithms at 5min and 15min level. the average forecasting-precision of AForecast algorithm at different forecasting-steps are computed adaptively except for the specified position 300s and 900s. We observe the 5mins level *average*

forecasting-precision of AForecast is lower than 5mForecast's but higher than 15mForecast's. At 15min level the *average forecasting-precision* of AForecast is higher than 15mForecast's.

Based on experiment results we can conclude that: 1) from 36 hours forecasting result curve we observe that our method can adaptive well no matter at the inflexion or at the period when stream-value quickly fluctuates.2) the general *average forecasting-precision* of AForecast is the highest among three algorithms.

6   Conclusions

We investigate the problem of stream-value forecasting based on typical application of data streams domain in this paper. The proposed AForecast method integrates the accurate merit of non-linear forecasting method and the rapid merit of linear forecasting method. Its key idea is that adaptively setting up multiple forecasting-step based on the variety cases of stream-value at different periods in a one-pass scanning manner. Compared with conventional forecasting methods, our method has two advantages: much higher average precision and more selected forecasting-steps. Experiments prove that the proposed method is a feasible solution to online forecasting for time-varying data streams.

References

1. Babcock, B., Babu, S., Datar, M., Motwani, R. and Widom, J.: Models and Issues in Data Streams. In Proc. ACM Symp. on Principles of Database Systems, (2002) 1–16.
2. Nasraoui, O., Rojas, C. and Cardona, C.: Single Pass Mining of Evolving Trends in Web Data with Explicit Retrieval Similarity Measures. In Proceedings of "International Web Dynamics Workshop", International World Wide Web Conference, New York, NY, May.(2004).
3. Faloutos, C.: Stream and Sensor data mining. Tutorials in 9th International Conference on Extending DataBase Technology (EDBT 2004), Heraklion, Greece, (2004) 25–27
4. Ankur, J., Edward Y. C. and Wang,Y. F.: Adaptive Stream Resource Management Using Kalman Filters. SIGMOD Conference (2004) 11–22
5. Liu, K., Subbarayan, S., Shoults, R. R., Manry, M. T., Kwan, C., Lewis, F. L. and Naccarino,J.: Comparison of Very Short-Term Load Forecasting Techniques. IEEE Transactions on Power Systerm, (1996),Vol. 11(2)
6. Trudnowski, J. D., McReynolds, W. L. and Johnson, M. J.: Real-time very short-term load prediction for power-system automatic generation control. IEEE Transactions on Control Systems Technology, Mar (2001),Vol. 9(2), 254–260
7. Mallat, S.: A Wavelet Tour of Signal Processing, Second Edition. (1999) by Academic Press, 221–226

Finding Closed Itemsets in Data Streams

Hai Wang[1], Wenyuan Li[2], Zengzhi Li[1], and Lin Fan[1]

[1] School of Electronics & Information Engineering, Xi'an Jiaotong University
Xi'an Shaanxi 710049, China
`hwang@mailst.xjtu.edu.cn, lzz@mail.xjtu.edu.cn, doby_007@163.net`
[2] Nanyang Technological University, Centre for Advanced Information Systems
Nanyang Avenue, N4-B3C-14, Singapore 639798
`liwy@pmail.ntu.edu.sg`

Abstract. Closed itemset mining is a difficult problem especially when we consider the task in the context of a data stream. Compared to mining from a static transaction data set, the streaming case has far more information to track and far greater complexity to manage. In this paper, we propose a complete solution based on CLOSET+ algorithm to closed itemset mining in data streams. In data streams, bounded memory and one-pass constraint are expected. In our solution, these constraints are both taken into account.

1 Introduction

Frequent pattern mining is the most important and demanding task in many applications. Although it has been widely studied [3, 4, 6, 8, 10–15] and used, it is challenging to extend it to data streams. Compared to mining from a static data set, the streaming case has more complex and more difficulties to overcome.

In recent years, many emerging applications, such as network traffic analysis, Web click stream mining, sensor network data analysis, and dynamic tracing of stock fluctuation, call for study of a new kind of data, called stream data. Designing algorithms for data streams is a challenging task: (a) there is a sequential one-pass constraint on the access of the data; (b) and it must work under bounded (i.e., fixed) memory with respect to the data stream.

Also, the continuity of data streams motivate time-sensitive data mining queries that many existing algorithms do not adequately support. For example, an analyst may want to finding out how a particular frequent pattern evolves over the lifetime of the stream. So, there is an increasing interest to revisit data mining problems in the context of this new model and application.

In this paper, we study the problem of finding closed itemset[9] in a data stream. Frequent itemsets mining has been extensively studied in the last years. Several variations to the original Apriori algorithm, as well as completely differet approaches have been proposed. The Apriori downward closure property makes it possible to effectively mine sparse datasets, also for very low min_supp thresholds. Sparse datasets in fact contain transactions with week correlations, and the downward closure property allows to strongly reduce the search space and

the dataset itself as the execution progresses. On the other hand, dense datasets contain strongly related transactions, and are much harder to mine since only a few itemsets can be in this case pruned, and the number of frequent itemsets grows very quickly while decreasing of the minimum support threshold. As a consequence, the mining task becomes rapidly intractable by the traditional mining algorithms, which try to extract all the frequent itemsets.

Closed itemsets mining is a solution to this problem. Closed itemsets are a small subset of the frequent itemsets, but they represent exactly the same knowledge in a more succinct way. From the set of closed itemset it is in fact straightforward to derive both the identities and supports of all frequent itemsets. Mining the closed itemset is thus semantically equivalent to mine all frequent itemsets, but with the great advantage that closed itemset are orders of magnitude fewer than frequent ones. Using closed itemset we implicitly benefit from data correlations which allow to strongly reduce problem complexity.

In this paper, we proposed a method that uses CLOSET+[2] algorithm in the data stream environment to find closed itemset. We actively maintain the closed itemset under a time windows framework. And the closed itemset are stored using a tree structure similar to FP-tree[8], and updated incrementally with incoming transactions. We call this structure closed itemset tree.

The contribution of this paper is a solution of mining closed itemset in data stream environment. The remaining of the paper is organized as follows. Section 2 presents the problem definitions and provides a basic analysis of the problem; Sections 3 explains how to overcome the problem encountered in shift the CLOSET+ problem. Because of in CLOSET+ , to generate the global FP-tree need to scan the transaction data base twice, that is not fit to the data stream environment. And also, how to store the mining result is another challenge. So the shift is necessary. Section 4 outlined the algorithm and section 5 concludes the study.

2   Problem Definition

Our task is to find all the closed itemsets in a data stream. Stream data takes the form of continuous, potentially infinite, as opposed to finite, statically stored data sets. So firstly, we need to shift old concepts related to closed itemsets to the new data stream environment.

To study frequent pattern mining in data streams, we first examine the same problem in a transaction database. Let $\mathbf{I} = \{i_1, i_2, \ldots, i_m\}$ be a finite set of items and D a set of transactions where each transaction $t \in D$ is a list of distinct items $t = \{x_0, x_1, \ldots, x_T, x_i \in I\}$. An ordered sequence of n distinct items $X = \{i_0, i_1, \ldots, i_n\} | X \subseteq \mathbf{I}$ is called itemset of length n, or n-itemset. The number of transactions in D including an itemset I is defined as the support of I (or supp(I)). Given a threshold min_supp, an itemset is said to be frequent if its support is greater than (or equal to) min_supp, infrequent otherwise.

Let $\mathcal{D}$ and I, $\mathcal{D} \in D$ and $I \in \mathbf{I}$, be subsets of all the transactions and items appearing in D, respectively. The concept of closed itemset is based on the two following functions f and g:

$$f(\mathcal{D}) = \{i \in \mathbf{I} | \forall s \in \mathcal{D}, i \in s\}$$
$$g(I) = \{s \in D | \forall i \in I, i \in s\}$$

function f returns the set of itemsets included in all the transactions belonging to $\mathcal{D}$, while function g returns the set of transaction supporting a given itemset I.

Definition 1. *An itemset I is said to be closed if and only if*

$$c(I) = f(g(I)) = f \circ g(I) = I$$

where the composite function $c = f \circ g(I)$ is called Galois operator or closure operator.

The closure operator defines a set of equivalence classes over the lattice of frequent itemsets: two itemsets belong to the same equivalence class iff they have the same closure, i.e. they are supported by the same set of transactions. We can also show that an itemset I is closed if no superset of I with the same support exists. Therefore mining the maximal elements of all the equivalence classes corresponds to mining all the closed itemset.

Now change our view to the data stream environment. Assume that one can only see the set of transactions in a limited size of window at one moment. In data streams, the set of transactions is the coming transactions during the specified time period T. So all other concepts we discussed is based on the specified time period t. We redefine the closed itemsets in data streams as the closed itemsets over time period t.

CLOSET+ follows the popular divide-and-conquer paradigm and the depth-first search strategy which has been verified a winner for mining long patterns by several efficient frequent parrern mining algorithms. It uses FP-tree as the compression technique. A depth-first search and horizontal format-based method like CLOSET+ will compute the local frequent items of a certain prefix by building and scanning its projected database.

Besides adopting the item merging and sub-itemset pruning methods, an item skipping technique is used in CLOSET+ to further prune search space and speed up mining.

CLOSET+ is a highly scalable and both runtime and space efficient algorithm for dense and sparse datasets, on different data distributions and support thresholds. So it is suitable for data stream environment. The basic idea behind our solution is based on how to shift the CLOSET+ algorithm to data stream environment. The construction of the global FP-tree is the key point of the shift. To achieve this goal, we must overcome constraint of the bounded memory and the one pass requirement. How to store the mining result is also another important issue to pursue.

3   Working with Bounded Memory

Up to this point, we have shown the basic idea of our solution. In this sub-section and the next, we discuss how we make adjustments to satisfy the data stream constraints.

We begin with the issue of bounded memory. Without doubt, any attempts to process an unbounded data stream is likely to exhaust the limited computing resources before producing any results. To overcome this, we process the stream in a sliding window fashion.
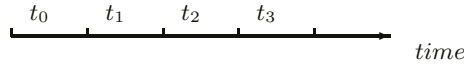

Fig. 1. Time Axis

As shown in Fig.1, the time window are divide averagely. The result of closed itemset for each window will be maintained. Using this scenario, we can answer the following queries:(1)what is the closed itemset over period t_m?(2)where are the periods when (a,b) is closed? That is, one can (1) mine closed itemset in the current window, (2) mine evolution of closed itemset based on the changes of their occurrences in the sequence of windows.

For each time window, we can record window-based count for each closed itemset. We use a compact tree representation of closed itemset, called closed itemset tree. Fig.2 shows an example. Each node in the closed itemset tree represents a closed itemset (from root to the node) and its frequency is recorded in the node. This tree shares the similar structure with FP-tree. The difference is that it strores closed itemset rather than streaming data. Actually, we also can use the same FP-tree construction method in [8] to build this tree by taking the set of closed itemset as input.

Usually closed itemset do not change dramatically over time. Therefore, the tree structure for different time windows will likely have considerable overlap. If we can embed the time property into each node in the closed itemset tree, space can be saved. Thus we propose to use only one closed itemset tree, where at each node, the frequency for each time window is maintained. The structure has been used in [1] to store the frequent itemset mining result.

Intuitively, we can not maintain all the result closed itemsets for every time window because of memory boundary. To deal with this problem, there are several choices. The simplest way is to delete the latest window directly. Or we can use some kind of strategy to delete some branch of the closed itemset tree. For example, we may delete the closed itemsets appeared in the last window but not in the current window. This sounds not a good idea too. Because the stream data often appears the periodicity. For example, if the periodicity is double size of the window size. We will definitely lost useful information if we use the deletion strategy just mentioned before. Our design is to use some decaying function to track the frequency degree of every mined closed itemsets.

Definition 2. *Frequency degree of the specified closed itemsets in a data stream: The sum of the supports of that closed itemset in the specified continuous windows.*

Obviously, frequency degree at some level reflects the importance of the closed itemset. It is hard to get the true value of the frequency degree of the specified

Each node in the closed itemset tree represents a closed itemset (from root to the node) and its frequency is recorded in the node.
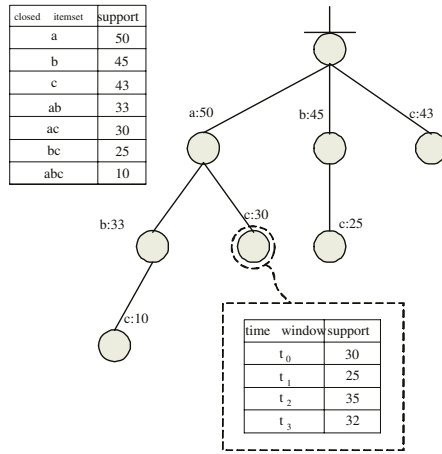

closed itemset	support
a	50
b	45
c	43
ab	33
ac	30
bc	25
abc	10

time window	support
t_0	30
t_1	25
t_2	35
t_3	32

Fig. 2. Closed Itemset Tree

closed itemset. But we can simulate it approximatively. By attach a variable fd to each seen closed itemset, we can approximatively track the frequency degree of that closed itemset. When we first met a closed itemset, set its fd variable to a specified value plus that support of that closed itemset in the current window. And as the time passed, use the decaying function to drop that value down. If that value touched a specified threshold, we delete that branch. If before the degree reach that threshold, that closed itemset appears again, we reset its value to value plus its support in the current window.

Let's give an example to explain the strategy above. Suppose the decaying function is fd=fd-1 as the time window sliding one and the threshold is 0. We consider the first 10 time windows in the data stream. And also suppose there is a itemset X and in the 1,3 time window it is closed with support 3,1 separately. When we first met the closed itemset X in the data stream and set its fd to 3. As the time window sliding, the fd do the self-degression operation until the we met it again in the third time window. Then, we reset its fd to $2 + 1 = 3$. Then repeat the process. After the seventh window sliding, the fd of X touches 0. We delete the branch of X.

4   Accessing Data in One-Sequential Pass

In the CLOSET+ algorithm, the construction of FP-tree needs scan the transaction data stream twice. Firstly, scan the transaction stream to find the set of frequent items and sort them in the frequent descending order to get the f_list. Then scan the transaction data stream again to build the FP-tree. In data stream environment, we only can pass the data stream once. So we need to reduce that

two pass algorithm into one pass.

The idea is simple: we begin by assuming a default frequent item list and give a initial ordering to them. The transtraction incoming one by one. As we process each of them, we update the corresponding items' frequency in the list and insert the transaction into the FP-tree. Upon reaching the end of window, we update the ordering of the frequent item list, and use this new ordering in the next window.

This model does work well in most cases. But for some cases especially that item frequencies does change a lot from window to window, we should modify the strategy a little to adapt the situation. Our method is to use a tiny taste window to correct the f_list obtained in the previous window. Firstly, we should give the size of the taste window and then set a buffer to store the data in that window. When a new time window coming, we should not insert the transactions to the FP-tree directly. Rather, we store the transaction data to the taste window buffer and at the same time scan the transaction data to correct the f_list obtained in the previous window until reach the end of the taste window. Then we construct the FP-tree. This time, we firstly scan the data in the taste window buffer and then the incoming stream data.

The chosen of the taste window size is the key point of this model. The taste window must not impact the time and space complexity of the whole algorithm. And also, it should reflect the data changes as big as it could.

This model can handle both cases of data changing smoothly and tempestuously. It is proved to be effective and reduces the construction of FP-tree to a single pass.

5   Algorithm

In this section, we give the complete algorithm for mining closed itemset in data stream.

Algorithm 1: (incremental update of the closed itemsets tree structure with incoming stream data)

Input: (1) An closed itemsets tree, (2) a min_support threshold, σ, (3) an empirical item frequency list, f_list_e, and (4) an incoming transaction stream T_s, and the transactions actually are arriving one at a time from a stream), (5) the time window size T_n and the taste time window size T_{tst}

Output: The updated closed itemsets tree

Method:

1. Initialize the closed itemset tree to empty, set the f_list with the empirical value
2. Do the following to construct the FP-tree for the current time window.
    (a) set the time accumulate variable t to 0 and if the temporary f_list_t is empty, set the f_list with f_list_e, else set the f_list with the f_list_t and make the f_list_t empty.
    (b) use the taste time window strategy to correct the f_list and store the transaction data in the taste window to buffer.

   (c) insert each transaction in taste window buffer to the FP-tree and then the incoming data until t reach T_s. At the same time, record the item frequency in another variable f_list_t.

3. Do the closed mining as follows:

   (a) Using CLOSET+ algorithm mining closed itemsets for the current time window and save the result to the closed itemsets tree with time property embedded.

   (b) update the frequencies of each closed itemsets and do the branch deletion operation.

   (c) If not reach the end of the stream, goto step (2).

6   Conclusions

In this paper we have investigated the problem of finding closed itemset in a data stream environment and gave an solution to this problem.

We incrementally maintain the closed itemset of the data stream in an effective data structure closed itemset tree. Based on this, we take the memory bounded and one pass constraints into account and developed a one pass algorithm based on CLOSET+ to find the closed itemsets in the data stream environment. This model is examined by our analysis. We have showed that mode are suitable for the closed itemset mining in data streams.

References

1. C. Giannella, J. Han, J. Pei, X. Yan, and P.S. Yu, Mining Frequent Patterns in Data Streams at Multiple Time Granularities, in H. Kargupta, A. Joshi, K. Sivakumar, and Y. Yesha (eds.), Next Generation Data Mining, AAAI/MIT, 2003
2. J. Pei, J. Han, and J. Wang. Closet+: Searching for the best strategies for mining frequent closed itemsets. In SIGKDD '03, August 2003.
3. J. Liu, Y. Pan, K. Wang, and J. Han. Mining frequent item sets by opportunistic projection. In SIGKDD'02, July 2002.
4. M. Zaki and C. Hsiao. CHARM: An efficient algorithm for closed itemset mining. In SDM'02, April 2002.
5. D. Burdick, M. Calimlim, and J. Gehrke. MAFIA: A maximal frequent itemset algorithm for transactional databases. In ICDE'01, April 2001.
6. E. Han, G. Karypis, V. Kumar. Scalable Parallel Data Mining for Association Rules. In TKDE 12(2), 2000.
7. J. Pei, J. Han, and R. Mao. CLOSET: An efficient algorithm for mining frequent closed itemsets. In DMKD'00, May 2000.
8. J. Han, J. Pei, Y. Yin. Mining frequent patterns without candidate generation. In SIGMOD'00, May 2000.
9. N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal. Discovering frequent closed itemsets for association rules. In ICDT'99, Jan. 1999.
10. R.J. Bayardo. Efficiently Mining long patterns from databases. SIGMOD'98, June 1998.
11. S. Brin, R. Motwani, J.D. Ullman, S. Tsur. Dynamic Itemset Counting and Implication Rules for Market Basket Data. In SIGMOD'97, May 1997.

12. D. Gunopulos, H. Mannila, and S.Saluja. Discovering All Most Specific Sentences by Randomized Algorithms. In ICDT'97, Jan. 1997.
13. H. Toivonen. Sampling Large Databases for Association Rules. In VLDB'96, Sept. 1996.
14. J. Park, M. Chen, P.S. Yu. An Effective Hash Based Algorithm for Mining Association Rules. In SIGMOD'95, May 1995.
15. R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases. In SIGMOD'93, May 1993.

An Efficient Schema Matching Algorithm

Wei Cheng, Heshui Lin, and Yufang Sun

Institute of Software, Chinese Academy of Sciences, Beijing 100080, China
{chengwei,heshui02,yufang}@iscas.cn

Abstract. Schema matching, the problem of finding semantic correspondences between elements of source and warehouse schemas, plays a key role in data warehousing. Currently, the mappings are largely determined manually by domain experts, thus a time-consuming process. In this paper, based on a multistrategy schema matching framework, we develop a linguistic matching algorithm using semantic distances between words to compute their semantic similarity, and propose a structural matching algorithm based on semantic similarity propagation. After describe our approach, we present experimental results on several real-world domains, and show that the algorithm discovers semantic mappings with a high degree of accuracy.

1 Introduction

Schema matching, the problem of finding semantic correspondences between elements of two schemas [3], plays a key role in data warehousing. In data warehouse, data from source format are required to be transformed into the format of data warehouse data. So discovering semantic correspondences between schemas from different sources is one of the critical steps toward data warehousing. Schema matching is also a key component of model management [1]. Currently, schema matching is largely performed manually by domain experts, thus it's a time-consuming and labor-intensive process. Moreover, the manual effort required is typically linear in the number of schemas to be matched and the complexity of the schema structures. Hence (semi-)automated and less labor-intensive schema matching approaches are needed.

A lot of previous work on automated schema matching has been done by researchers [3,4,5,10,11,13,15]. Matching approaches used in these prototypes include name-based matching, instance-based matching, and structure matching, etc. In this paper, we propose a schema matching algorithm by integrating both linguistic and structural matching. In the linguistic matching phase, we developed an algorithm based on semantic distances between words to compute their semantic similarity. In the structural matching phase, we proposed a structural-based matching algorithm GSMA to compute semantic similarity based on similarity propagation.

The rest of this paper is organized as follows. Some related work is discussed in Section 2. Section 3 provides an overview of our matching approach and details of the linguistic and structural matching algorithm. Section 4 presents experimental results on several real-world domains. Section 5 summarizes the contributions and suggests some future research work.

2   Related Work

A survey of existing techniques for automatic schema matching is given in [16], it proposed orthogonal taxonomy of schema matching: schema-level vs. instance-level, element-level vs. structure-level, and language-based vs. constraint-based matching approaches. Here, four schema match prototypes are reviewed, namely, Cupid [11], LSD [3], SemInt [10] and SKAT [13].

The LSD system uses machine learning techniques to perform schema matching [3]. It consists of four components: base learners, meta-learner, prediction combiner, and constraint handler. LSD trains four kinds of base learners to assign tags of elements of a mediated schema to data instance of a source schema, the predictions of base learners are weighted and combined by the meta-learner. Structural matching is not utilized in LSD.

The SemInt system is an instance-based matching prototype [10]. It derived constraint-based and content-based matching criteria from instance data values. Then a match signature is determined. Signatures can be clustered according to their Euclidean distance. A neural network is trained on the cluster centers to output an attribute category, and thus to determine a list of match candidates. SemInt does not make use of linguistic matching or structural matching in the matching process.

Cupid combines linguistic matching with structure matching to get a composite matching result. It first computes similarities between elements of two schemas using linguistic matching based on their names, data types and domain constraints, etc. Then, it uses a *TreeMatch* algorithm to compute similarities of schema elements based on their similarity of the contexts or vicinities. Finally, a weighted similarity of the two kinds of similarities is computed, thus a mapping is determined [11].

The SKAT prototype is a rule-based schema matching system [13]. In SKAT, match or mismatch relationships of schema elements are expressed by rules defined in first-order logic, and methods are defined to derive new matches. The initial application-specific match and mismatch relationships are manually provided by users. Based on that, new generated matches are accepted or rejected.

3   Our Matching Approach

In our prototype, we use a multistrategy approach to combine outputs of different matchers, which include linguistic matching based on linguistic characteristics and structural matching based on schema structure. Importantly, our prototype is a platform to combine different algorithms and approaches in a flexible way. The architecture of our prototype is shown in Figure 1. It operates in four phases.

In the first phase, called preprocess phase, schemas are converted into directed graphs using RDF model [14], with nodes representing elements of a schema, edges representing relationships between nodes. The prototype exploits the whole matching space, creates all theoretic match pairs and removes constraint-conflict match pairs. The second phase, called linguistic matching, matches elements of two schemas based on their semantic similarity. We use an algorithm based on the WordNet thesaurus to compute the semantic distance between two objects, then their semantic similarity can be obtained via a transform. The third phase is the structural matching of schema elements. A structural matching algorithm called GSMA is used to com-

pute similarities between nodes of graphs. The algorithm is based on the assumption that elements of two distinct models are similar when their adjacent elements are similar. The fourth phase is the combination phase, in which matching results obtained from previous two steps are weighted and combined. Then a matching is determined by choosing pairs of schema elements with the highest similarity value.
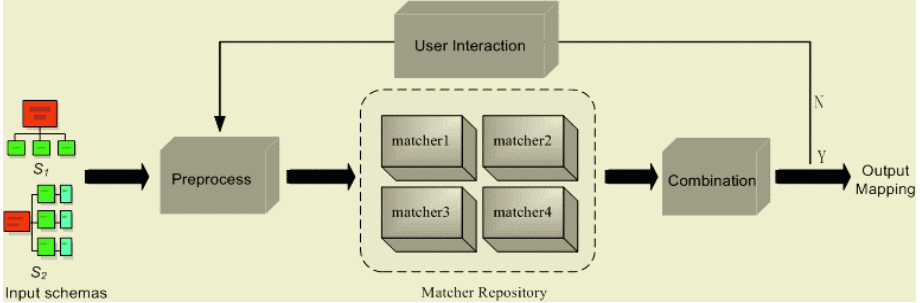

Fig. 1. Framework of the schema matching system

3.1 Linguistic Matching

The linguistic matching is primarily based on schema element names using the name matcher. Commonly, names of schema elements, especially regular named elements, are very useful for schema matching. Much previous research work on this field has been done [2,9]. Below outlines the linguistic matching process in three steps.

1. An algorithm based on the WordNet thesaurus is used to compute semantic distances between nodes in two graphs that represent elements in two schemas. Words in the thesaurus are organized as a tree/subtree structure, each node in the tree corresponds to a word, and nodes are connected by edges [7,8]. We define the step length of the route between two nodes of the tree as their semantic distance d. Given two nodes w_1, w_2 of the words tree, we get:

$$d = step\_legnth(w_1, w_2),\qquad(1)$$

   where $step\_length(\ )$ is the function to compute the length of steps between two nodes.

2. Semantic similarity sim between two nodes w_1 and w_2 then can be obtained from their semantic distance d via a transform:

$$sim(w_1, w_2) = \frac{\alpha}{d(w_1, w_2) + \alpha},\qquad(2)$$

   where α is an adjustable parameter.

3. All mapping pairs are compared and clustered. Only highly similar mapping pairs can be selected as matching candidates. As to some domain-specific synonyms such as "Dept" for "Department" and "Pno" for "Personal Number", obviously, the above algorithm cannot provide a good solution. To get accurate match results of such cases, we build a domain-specific thesaurus which includes a set of domain-specific synonym tables, common abbreviation tables and acronym tables.

3.2 Structural Matching

In our prototype, we developed a GSMA (Generic Structural Matching Algorithm) algorithm based on similarity propagation. The similarity propagation idea is borrowed from the Similarity Flooding algorithm [12]. The structural matching algorithm is based on following intuitions: i) Two elements in two object models are similar if they are derived or inherited by similar elements. ii) If two elements in two object models are similar, the similarity of their adjacent elements increases.

In the GSMA algorithm, given two graphs A and B, the initial similarity between node pair (a, b) is defined as below:

$$Sim_0 = \begin{cases} 0 / initialMap & (a \neq b) \\ 1 & (a = b) \end{cases}$$
(3)

The similarity Sim_k ($k \geq 1$) of node pair (a, b) after k iterations is defined as:

$$Sim_k = wSim_{k-1} + \frac{w_I}{|I(a)||I(b)|} \sum_{i=1}^{|I(a)||I(b)|} \sum_{j=1} Sim_{k-1}(I_i(a),I_j(b)) + \frac{w_O}{|O(a)||O(b)|} \sum_{i=1}^{|O(a)||O(b)|} \sum_{j=1} Sim_{k-1}(O_i(a),O_j(b)),$$
(4)

where $I(n)$ represents the set of all in_neighbor nodes of node n and $O(n)$ represents the set of all out_neighbor nodes of node n. And w, w_I, w_O are the weight values for the similarity value of k-1 iteration Sim_{k-1}, the mean of similarity values of in neighbors of (a, b) and the mean of similarity values of out neighbors of (a, b), which come from previous k-1 iterations, respectively. And $w + w_I + w_O = 1$.

Below is the GSMA algorithm:

Input: Directed graph A, B and initial similarity matrix between nodes of A and B
Output: Stable similarity matrix between nodes of A and B after propagations.
Step 1: Initialize the similarity matrix following equation (1).
Step 2: Begin the loop until similarities get stable and convergent, the convergence rule is:

$$| Sim_k(a,b) - Sim_{k-1}(a,b) | < \varepsilon .$$
(5)

(1) Computing the similarity between pair of nodes (a, b):
   ① Computing the mean of similarities between in_neighbor nodes of (a, b):

$$Sim_k I(a,b) = \frac{1}{|I(a)||I(b)|} \sum_{i=1}^{|I(a)||I(b)|} \sum_{j=1} Sim_{k-1}(I_i(a),I_j(b))$$
(6)

   ② Computing the mean of similarities between out_neighbor nodes of (a, b):

$$Sim_k O(a,b) = \frac{1}{|O(a)||O(b)|} \sum_{i=1}^{|O(a)||O(b)|} \sum_{j=1} Sim_{k-1}(O_i(a),O_j(b))$$
(7)

   ③ Computing the similarity between (a, b):

$$Sim_k(a,b) = w \times Sim_{k-1}(a,b) + w_I Sim_{k-1}I(a,b) + w_O Sim_{k-1}O(a,b)$$
(8)

(2) Revising the similarity of (a, b):

① Computing the standard deviation of the similarity vectors between node a and any node in B:

$$\overline{Sim} = \frac{1}{|B|} \sum_{j=1}^{|B|} Sim_k(a, B_j) \,, \tag{9}$$

$$StDev = \sqrt{\frac{1}{|B|-1} \sum_{j=1}^{|B|} (Sim_k(a, B_j) - \overline{Sim})^2} \tag{10}$$

② Revising the similarity of (a, b):

$$Sim_k(a,b) = Sim_k(a,b) * (1 + (Sim_k(a,b) - \overline{Sim})/StDev) \tag{11}$$

Step 3: Normalize the similarity matrix:

$$Sim_k(a,b) = \frac{Sim_k(a,b)}{Max(Sim_k(a, B_j))} \,. \tag{12}$$

Step 4: Output the similarity matrix $Sim(A,B)$ of A and B.

In every iteration, each row of the similarity matrix is revised by the standard deviation of that row. The purpose of the revising operation is to differentiate and cluster similarity values. Over k iterations, when the Euclidean length of the residual vector $\Delta(Sim_k - Sim_{k-1})$ of two similarity values in two sequential iterations is less than a threshold ε, the iteration is terminated, i.e. similarities of all node pairs are stabilized.

4   Experimental Evaluation

We have implemented the linguistic and structural matching algorithm in our prototype system and have evaluated them on several real-world domains [6]. The experimental environment is a PIII 667 PC with 256M memory and Redhat Linux 9.0 OS, the develop tool is Anjuta 1.2.1. To evaluate the quality of the matching operations, we use the same match measures used in [4]. Manually determined real matches for a match task are compared with matches returned by automatic match processing in the sets shown in Figure 2. A, B, C, D corresponds to false negatives, true positives, false positives and true negatives, respectively.
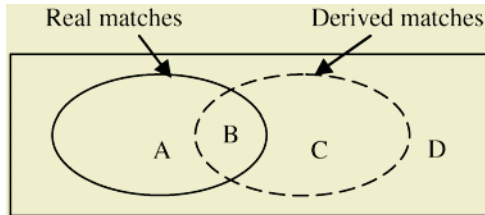
Fig. 2. Match categories

Based on these match categories, three quality measures originated from the IR (information retrieval) field are computed: Precision = B/(B+C), reflects the reliability of the matching result; Recall = B/(A+B), specifies the share of real matches that is found; F-Measure = 2/(1/Recall+1/Precision), the harmonic mean of Precision and Recall, a combined measure for match quality.

To evaluate the efficiency of the matching algorithm, we defined 8 match tasks, each matching two different schemas. Experimental results are shown in Figure 3 and 4.
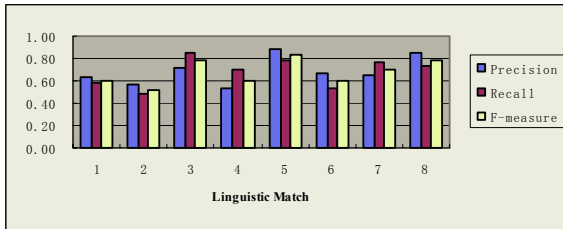
Fig. 3. Measure values of match tasks using linguistic match
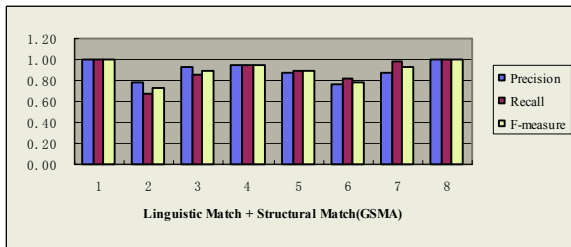
Fig. 4. Measure values of match tasks using linguistic&structural match with GSMA algorithm

From Figure 3 and 4, we know that when use the linguistic matcher, the system get an average *F-measure* value of 68%. When the structural matcher with the GSMA algorithm is added, the average *F-measure* value is increased to 90%. This shows the GSMA algorithm can highly improve the matching accuracy. Experimental results also show that the revising operation by standard deviation makes the GSMA algorithm have strong ability to differentiate and cluster nodes, i.e. it make similarity values between nodes that are actually similar far more greater than similarity values between nodes that are actually unsimilar. This result is just what we expected.

Compared with other structural matching algorithms, i.e. the Similarity Flooding algorithm and the *Treematch* algorithm of Cupid, GSMA algorithm has two advantages:

1. The Similarity Flooding algorithm works well for directed labeled graphs, it degrades highly when graph is unlabeled or labeling is uniform [12]. While the GSMA algorithm does not rely on edge labeling.
2. Cupid uses the *Treematch* algorithm to do structural matching [11]. But the outputs depend heavily on the order in which schemas are compared to each other. And as to schemas with similar structure but different semantic correspondences,

Cupid could not always differentiate them correctly. In such cases, the GSMA algorithm always gives satisfying results.

5   Conclusion and Future Work

In this paper, we described an efficient schema matching algorithm for the well-known problem of schema matching. Experimental results show that our algorithm works well on real-world domains. Although our experiments show promising results, we do not believe our approach is a robust solution, since it needs the inclusion of other matchers and approaches. And the complex schema matching problem, i.e. *n:m* (*n, m* > 1) matches, should be studied deeply. Our long-term goal is to make our schema matching prototype be a generic schema matching system, which is more powerful and efficient.

References

1. Bernstein, P.A., A. halevy, R.A.Pottinger: A Vision for Management of Complex Models, SIGMOD 2000.
2. Bechhofer S and Goble CA, Delivering Terminological Services, Journal of AI*IA Notizie 12(1) pp: 27-32, 1999.
3. Doan, A., P. Domingos, A. Halevy: Reconciling Schemas of Disparate Data Sources: A Machine-Learning approach. SIGMOD 2001.
4. Do, H. and E.Rahm: COMA – A System for flexible combination of schema matching approaches, VLDB 2002.
5. Embley,D.W. et al. Multifaceted Exploitation of Metadata for attribute Match Discovery in information Integration. WIIW 2001.
6. Http://anhai.cs.uiuc.edu/archive/summary.type.html
7. G.A. Miller. WordNet: A lexical database for English. Communications of the ACM, 38(11):39–41, November 1995.
8. Http://www.cogsci.princeton.edu/~wn/
9. L. A. Carr, W. Hall, S. Bechhofer and C. A. Goble, "Conceptual Linking: Ontology-based Open Hypermedia", Proceedings of the Tenth International World Wide Web Conference, Hong Kong, pp.334-342, May 2001.
10. Li W, Clifton: SemInt: A Tool for Identifying Attribute Correspondences in Heterogeneous database Using Neural Network, Data & Knowledge Engineering, 2001.
11. Madhavant,J., P.A.Bernstein, E.Rahm: Generic Schema Matching with Cupid. VLDB 2001.
12. Melnik S, Garcia-Molina H, Rahm E: Similarity Flooding: A versatile graph matching Algorithm, ICDE 2002.
13. Mitra P, Wiederhold G, Jannink J: Semiautomatic integration of knowledge sources, FUSION 99.
14. O. Lassila and R. Swick. Resource Description Framework (RDF) Model and Syntax Specification. http://www.w3.org/TR/REC-rdf-syntax/, 1998.
15. Palopoli L, Terracina G, Ursino D: The system DIKE: towards the semi-automatic synthesis of cooperative information systems and data warehouse. ADBIS-DASFAA Conf 2000.
16. Rahm, E., P.A. Bernstein. A survey of approaches to automatic schema matching. The VLDB Journal, 10(4):334 ~ 350, 2001.

A Divisive Ordering Algorithm
for Mapping Categorical Data to Numeric Data

Huang-Cheng Kuo

Department of Computer Science and Information Engineering
National Chiayi University, Chiayi City, Taiwan 600
`hckuo@mail.ncyu.edu.tw`

Abstract. The amount of computing time for K Nearest Neighbor Search is linear to the size of the dataset if the dataset is not indexed. This is not endurable for on-line applications with time constraints when the dataset is large. However, if there are categorical attributes in the dataset, an index cannot be built on the dataset. One possible solution to index such datasets is to convert categorical attributes into numeric attributes. Categories are ordered and then are mapped to numeric values. In this paper, we propose a new heuristic ordering algorithm to compare with two previously proposed algorithms that borrow the idea from minimal spanning trees. The new algorithm divisively builds a binary tree by recursively partitioning the categories. Then, we in-order traverse the tree and get an ordering of the categories. After mapping and indexing, we can efficiently retrieve a small portion of the dataset and perform K nearest neighbor search on the portion at the cost of a little bit of accuracy. Experiments show the divisive ordering algorithm performs better than the other two algorithms.

Keywords: Mapping, Divisive Ordering, Approximate Nearest Neighbor Search, Indexing

1 Introduction

K Nearest Neighbor Search (KNN) is a frequently applied technique in information systems. Also, it is the search phase for Memory-Based Reasoning (MBR) which is a useful data mining technique [BL97]. One of the major weaknesses of MBR is its scalability when search for the k nearest neighbors of the new object [BL97]. Similarity between the new object and each object in the dataset has to be computed. MBR response time may not be endurable for a very large dataset in online applications. A possible approach is to limit the size of dataset by methods such as instance selection [YX02, Zh92]. Another approach is to use indexing data structure for storing training examples.

   Objects in most MBR applications have multiple input attributes. A lot of studies on k nearest neighbor (KNN) search use tree-based indexing. Multi-dimensional data structures, such as R-Trees [Gu84] or its variants, are suitable for indexing. Since such indexing methods usually do not work well for *exact* KNN search if the dimensionality of the data is high [IM98]. A lot of *approximate* k-nearest neighbor search methods have been proposed to avoid the curse of dimensionality, such as using statistical density model to prioritize clusters for data to be scanned for post-processing [BFG99].

An R-tree-like multidimensional data structure is usually used for indexing objects. However, multidimensional indexing data structures do not handle categorical values well. Though a categorical attribute can be converted into a set of Boolean attributes according to its distinct categorical values, this will increase the dimensionality and decrease the performance. Also, this will lose the semantics of similarity between categories of a categorical attribute.

Therefore, we map the categorical values of a categorical attribute into numeric values [KLH04]. The mapped numerical values of a similar pair of categories should be close. The similarity between any pair of categories of an attribute is either given by domain experts or computed by an algorithm.

The study of finding a linear order of gene expression data has been explored [Bar03, LKM03]. Euclidean metric is assumed for distance between two gene expression profiles. The goal of ordering is to minimize the sum of distances between gene expression profiles adjacent in the linear ordering. However, a given dendrogram of the genes is assumed. In this paper, we assume only a distance matrix for the nominal objects.

In section 2, we define a measure metric for a category ordering. The categories of an attribute are ordered, such that similar categories are ordered together. In section 3, three ordering algorithms are presented. Mapping categories to numeric values for a certain ordering is also discussed. We demonstrate the performance of the algorithms in section 4. Section 5 is a conclusion.

2   Measure Metrics for Category Ordering and Mapping

For the purpose of ordering a set of categories, we consider a complete undirected graph, where a vertex is a category and the weight of an edge is the distance between the two vertices of the edge. The smaller the edge is, the more similar the two categories are. We define the measurement for an ordering with respect to a given undirected weighted complete graph.

Definition: Ordering Problem
Given an undirected weighted complete graph of n vertices, a simple path of length $n − 1$ is an ordering of the vertices. The edges are the distances between pairs of vertices. The ordering problem is to find a path, called *ordering path*, of maximal value according to a certain scoring function.

We propose two types of scoring functions, namely relationship scoring and pairwise difference scoring, for the ordering problem. The pairwise ordering scoring function can not directly be applied to an ordering path before the categories are mapped to numeric values. It evaluates the mapping result that depends heavily on the ordering path.

Relationship Scoring
We consider distance relationships among a 3-tuple of three categories. If the distance of a pair of categories is largest among of all three pairs, the pair of categories should be away from each other on the ordering path. If such a relationship is satisfied, we say that the distance (or similarity) relationship is preserved for the 3-tuple.

Definition: Reasonable Ordering Score

Let $P = <v_1, v_2, \ldots, v_n>$ be an ordering path of an undirected weighted complete graph G, where $|V(G)| = n$. $Dist(vi, vj)$ is the distance between vertices vi, vj. On an ordering path there $(n - 2)$ 3-tuples $<v_{i-1}, v_i, v_{i+1}>$ where $1 < i < n$. A 3-tuple $<v_{i-1}, v_i, v_{i+1}>$ is reasonable if and only if $dist(v_{i-1}, v_{i+1}) \geqq dist(v_{i-1}, v_i)$ and $dist(v_{i-1}, v_{i+1}) \geqq dist(v_i, v_{i+1})$. The percentage of reasonable 3-tuples on an ordering path P is called *reasonable ordering score* of the ordering path P.

Pairwise Difference Scoring

Our purpose is not only to order the categories, but also to assign numeric values to categories. The distance between two categories v_i and v_j is $dist(v_i, v_j)$. Let $mapping(v_i)$ be the mapping value for a category v_i. We hope that $dist(v_i, v_j)$ is close to the absolute value of $mapping(v_i) - mapping(v_j)$ as much as possible.

Therefore, we use the metric *root mean squared error* (*rmse*) for measuring the quality of *mapping* result.

$$rmse = \sqrt{\frac{\sum_{v_i \neq v_j}(dist(v_i, v_j) - dist_m(v_i, v_j))^2}{n*(n-1)/2}} \text{ , where}$$

$$dist_m(v_i, v_j) = |mapping(v_i) - mapping(v_j)|$$

3 Ordering Algorithms and Mapping

Given an undirected weighted complete graph, an ordering path is a simple path that visits all vertices. The first two algorithms borrow the idea of finding a minimum spanning tree from a weighted graph [KLH04]. We propose the third algorithm that recursively partitions the categories.

3.1 Prim-Style Ordering Algorithm

The Prim-style ordering algorithm imitates the Prim's minimum spanning tree algorithm. The edge selection condition is different. In an ordering problem, we are constructing a simple path. Initially, no edges and no vertices are included on the ordering path. In the first step, a minimal edge and the two vertices that the edge is incident on are put on the ordering path. Then, the algorithm iteratively chooses a minimal edge that is incident on a vertex u on the ordering path and a vertex v not on the ordering path. However, the (u, v) edge is not added to the ordering path. Instead, an edge incident on v and an end vertex of the ordering path is added.

3.2 Kruskal-Style Ordering Algorithm

Like the Kruskal's minimal spanning tree algorithm, the Kruskal-style ordering algorithm adds a minimal edge as long as the edge does not create a cycle. Since the algorithm drops edges that create a cycle, edges with large weights might be added in later iterations. Therefore, the algorithm does not guarantee the minimal total distances of the edges on the ordering path.

3.3   Divisive Ordering Algorithm

This algorithm builds a binary tree from a given undirected weighted complete graph. It first picks a vertex as the root. Then it splits the rest of the vertices into two groups of vertices and puts a group on each sub-tree. The process is repeated for the vertices in each sub-tree. An ordering path is constructed by in-order traversing the binary tree.

The central vertex of a set C of vertices may be defined in some ways. In this paper, we define the central vertex v as below.

$$v = \min_{v_i \in C} \{ \sum_{v_j \in C, j \neq i} dist(v_i, v_j) \}$$

We use agglomerative hierarchical clustering algorithm for grouping a set of vertices into two groups. Average linkage [Voor86] is adopted for defining the distance between two clusters during the clustering process.

The central vertex is the root of the binary tree, and the two groups of vertices are the vertices in the two sub-trees. The partition procedure is recursively applied to the sub-trees.

3.4   Mapping Function

Let $P = <v_1, v_2, \ldots, v_n>$ be an ordering path. Let $dist(v_i, v_j)$ be the distance between categories v_i and v_j. The following function maps v_1 to 0 and maps v_n to 1. All the other categories are mapped to values such that $mapping(v_{i+1}) - mapping(v_i)$ is proportional to $dist(v_i, v_{i+1})$ for $i = 1$ to $n - 1$.

$$mapping(v_i) = \sum_{j=0}^{i-1} dist(v_j, v_{j+1}) / \sum_{j=0}^{n-1} dist(v_j, v_{j+1})$$

If there are no constraints on the response time, records of the whole dataset are retrieved for distance calculating, the true K nearest neighbors will be correctly answered. Time for distance calculating is linear to the size of the dataset. However, in many applications, especially on-line applications, there are constraints on response time.

4   Experiments

In the first subsection, we compare the algorithms with generated graphs. A generated graph represents the distances among the distinct values of a categorical attribute. In the second subsection, a real-life dataset is used to compare KNN result by the three algorithms.

a. Mapping Scores of Synthetic Graphs
In our experiments, we adopt the strategy to plot the categories in a 2-D space. Such that, similar categories are plotted close to each other. Then, distance between a pair of categories is the Euclidean distance between the two points in the 2-D space.

Rectangles with length/width ratio 10 are used to simulate the distances between categories. 50 scores are averaged for each given number of categories and each or-

dering algorithm. The left-hand side chart of figure 1 shows the reasonable ordering scores. The Kruskal-style ordering algorithm performs worst. The Prim-style ordering algorithm performs better than divisive ordering algorithm when the number of categories is small. But, the divisive ordering algorithm outscores Prim-style ordering algorithm when the number of categories is large, roughly speaking, larger than 100. Also, the reasonable ordering scores for the divisive ordering algorithm are quite stable for all numbers of categories. The scores are around 0.9.
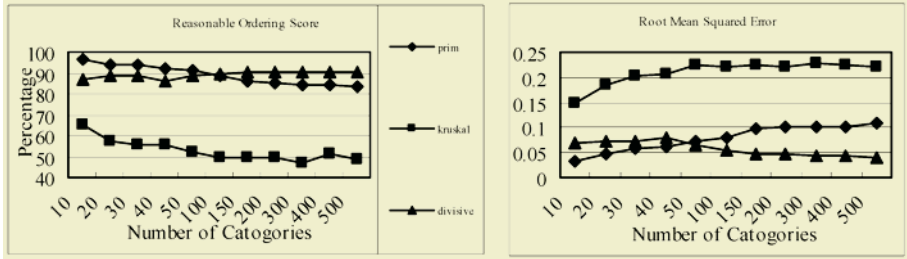
Fig. 1. Comparison of the algorithms on synthetic data

(a) Retrieve K * 2 objects

(b) Retrieve K * 3 objects

(c) Retrieve K * 4 objects

(d) Retrieve K * 5 objects
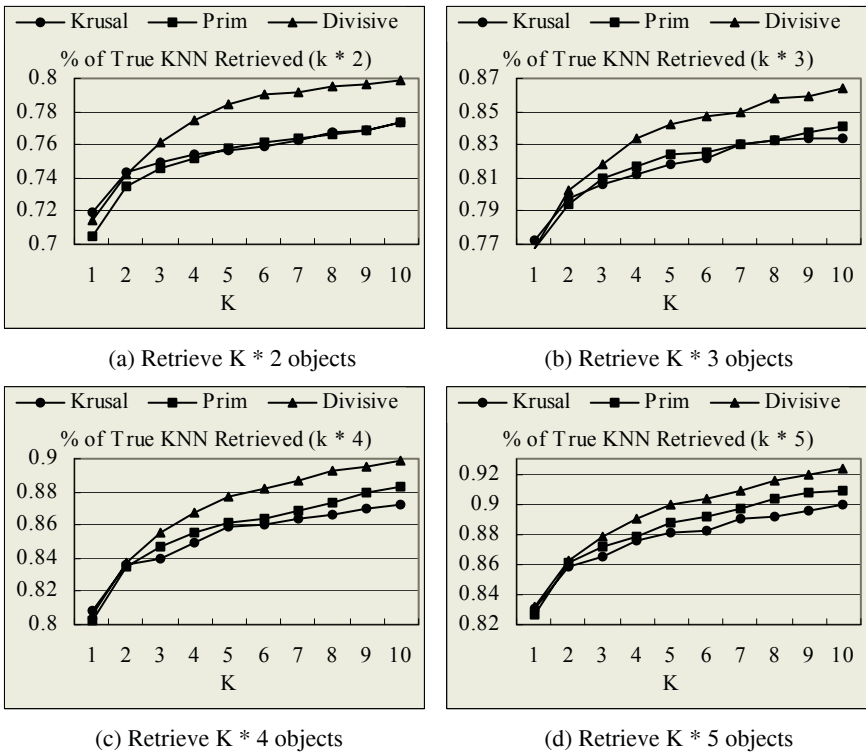
Fig. 2. Accuracy of Ordering Algorithms

b. Approximate KNN Search with Mapping

We use the "Census-Income" dataset from the University of California, Irvine (UCI) KDD Archive [HB99]. The similarity definition from CACTUS [GGR99] with simplification [KLH04] for each nominal attribute is adopted.

K times 2, 3, 4, and 5, respectively, records are filtered from the indexing structure. Those records are checked for *approximate* KNNs of the query record based.

We can observe that under any condition, (1) as K increases the percentage of true KNN retrieved increases, and (2) as the number of filtered records increases, the percentage of true KNN retrieved increases.

The divisive ordering algorithm outperforms the other two algorithms. This observation supports that better ordering algorithm results in better filtering quality.

5   Conclusions

We presented three ordering algorithms for mapping categories to numbers, such that, objects with categorical attributes can be mapped to numeric-only objects, and can be stored in a multidimensional indexing data structure. Since the distance relationships among categories of each attribute are preserved as much as possible, the indexing data structure can be served as a filtering mechanism for k nearest neighbor search.

For the experiments with generated sets of distance matrices with different numbers of categories, the result showed that ordering quality depends on the nature of the distance relationship among the categories.

The time complexity for the constructing all the complete graphs of the categorical attributes from a dataset with m records and n categorical attributes is $O(m^2 * n^2)$. This pre-processing phase can be done off-line. During the search phase, only a small portion of the large dataset is retrieved for distance computing. However, the pre-processing phase has to be re-examined from time to time when there are records added to the datasets. Sampling techniques can be used to retrieve a small, but enough, number of records for constructing the complete graphs of distances.

For the future work, we will the ordering path change when a new category is added. If the dataset is huge, will the data structure be re-organized according to the new mapping function? If the data structure is re-organized in a later time, when will the re-organization be performed? Response time constraints should be considered no matter which strategy is adopted to handle this incremental issue.

Multidimensional scaling (MDS) [ARL01, CC01] projects objects into a low-dimension space, usually 2D space, for visualization. Pairwise dissimilarity interrelationships are preserved as many as possible. It is usually used for small datasets, due to its time complexity. Comparison to MDS is undergoing in terms of execution time and goodness of non-metric MDS.

Acknowledgement

References

[ARL01] Dimitris K. Agrafiotis, Dmitrii N. Rassokhin, Victor S. Lobanov, "Multidimensional Scaling and Visualization of Large Molecular Similarity Tables," Journal of Computational Chemistry, Vol. 22, No. 5, 2001, pp. 488-500.

[Bar03] Ziv Bar-Joseph, et. al., "K-ary Clustering with Optimal Leaf Ordering for Gene Expression Data," Bioinformatics, Vol. 19, No. 9, 2003, pp. 1070–1078.

[BFG99] Kristin P. Bennett, Usama Fayyad, Dan Geiger, "Density-Based Indexing for Approximate Nearest-Neighbor Queries," ACM KDD, 1999, pp. 233-243.

[BL97] Michael J.A. Berry, Gordon Linoff, "Memory-Based Reasoning," chapter 9 in *Data Mining Techniques: for Marketing, Sales, and Customer Support*, 1997, pp.157-186.

[CC01] Trevor F. Cox and Michael A.A. Cox, *Multidimensional Scaling*, 2nd ed, Chapman & Hall, 2001.

[GGR99] V. Ganti, J. Gehrke, and R. Ramakrishnan, "CACTUS-Clustering Categorical Data Using Summaries," ACM KDD, 1999, pp. 73-83.

[GKR98] David Gilbson, John Kleinberg, Prabhakar Raghavan, "Clustering Categorical Data: An Approach Based on Dynamical Systems," VLDB Conference, 1998, pp. 311 – 322.

[Gu84] A. Guttman, "R-Trees: A Dynamic Index Structure for Spatial Search," ACM SIGMOD, 1984, pp. 47-57.

[HB99] S. Hettich and S. D. Bay, The UCI KDD Archive [http://kdd.ics.uci.edu]. Irvine, CA: University of California, Department of Information and Computer Science, 1999.

[IM98] P. Indyk and R. Motwani, "Approximate Nearest Neighbors: Toward Removing the Curse of Dimensionality," ACM Symposium on Theory of Computing, 1998, pp. 604-613.

[KLH04] Huang-Cheng Kuo, Yi-Sen Lin, Jen-Peng Huang, "Distance Preserving Mapping from Categories to Numbers for Indexing," Springer *Lecture Notes in Artificial Intelligence*, Vol. 3214, 2004, pp. 1245-1251.

[LKM03] Seung-Kyu Lee, Yong-Hyuk Kim, Byung Ro Moon, "Finding the Optimal Gene Order in Displaying Microarray Data," Genetic and Evolutionary Computation Conference, 2003, pp. 2215-2226.

[SM58] R. R. Sokal and C. D. Michener, "A Statistical Method for Evaluating Systematic Relationships," University of Kansas Science Bulletin, 38, 1958, pp. 1409 – 1438.

[Voor86] E. M. Voorhees, "Implementing agglomerative hierarchical clustering algorithms for use in document retrieval," Information Processing and Management, Vol. 22, 1986, pp. 465--476.

[Zh92] J. Zhang, "Selecting Typical Instances in Instance-Based Learning", in Proceedings of the Ninth International Conference on Machine Learning, 1992, pp. 470-479.

[YX02] Kai Yu, Xiaowei Xu, Jianhua Tao, Martin Ester, and Hans-Peter Kriegel, "Instance Selection Techniques for Memory-Based Collaborative Filtering," in Proceedings of the 2nd SIAM International Conference on Data Mining, 2002.

A Mutual Influence Algorithm
for Multiple Concurrent Negotiations
– A Game Theoretical Analysis⋆

Ka-man Lam and Ho-fung Leung

Dept. of Computer Science & Engineering
The Chinese University of Hong Kong
Shatin, N. T., Hong Kong, China

Abstract. Buyers always want to obtain goods at the lowest price. To achieve this, a buyer agent can have multiple concurrent negotiations with all the sellers. It is obvious that when the buyer obtains a new low offer from one of the sellers during negotiation, the buyer should have more bargaining power in negotiating with other sellers, as other sellers should offer a even lower price in order to make a deal. In this way, the concurrent negotiations mutually influence one another. In this paper, we present an algorithm to enable mutual influence among multiple concurrent negotiations. We show that the outcome of this algorithm is in Nash equilibrium when we model and analyze the negotiation as a game.

1 Introduction

One of the key properties defining intelligent agent is social ability, which is the capability of an agent to interact with other agents or humans [8, 9]. In the context of multiagent system, interaction often means communication and cooperation. There are two types of cooperation. One is that agents have common interest and they simply cooperate with each other to reach their common goal. Another is that agents can benefit mutually through cooperation, but they actually have conflicting interests. In these cases, agents need to negotiate in order to resolve conflicts and reach mutually beneficial agreements.

There are various applications for negotiation [2, 7], which include data allocation in information servers, resource allocation, task distribution, coalition formation, etc. In recent years, another application area, electronic commerce has received much concern [1, 3]. In this field, multiagent systems are built for various kinds of applications, including auction, bargaining, and negotiation. For negotiation applications, software agents are built to represent users in negotiating for goods or services. There are two types of agents, the buyers and the sellers. Obviously, buyers want to minimize the prices at which they obtain the goods or services. On the other hand, sellers want to sell the goods or services at

---

the highest possible prices. In order to make an agreement on the price, buyers and sellers carry out negotiations. In this paper, we concentrate on negotiations in this domain, using the alternating offers protocol [6], and show how a buyer can lower the price by mutually influencing multiple concurrent negotiations.

Negotiations exist in many different forms, of which, the simplest form is one-to-one negotiation. In this form of negotiation, one buyer negotiates with one seller at a time. The advantages of this form of negotiation are that it is easy to implement and the overhead on the system is low. However, for a particular good or service, very often there is more than one provider. If this is the case, one-to-one negotiation may not enable the buyer to minimize the price at which it obtains the good or service. One may do so by having one-to-one negotiations sequentially with all the sellers, and then go back to the seller who offers the lowest price. However, such process is lengthy and inefficient.

Another form of negotiation is one-to-many negotiation [3]. In one-to-many negotiations, sometimes in the form of reverse auctions, a buyer seeks the good or service from a number of providers. However, there are still limitations. In most reverse auctions, buyers make no counter-offers. As a result, buyers are negotiating in a passive manner. Another limitation is that sellers are offering their own price, but the prices do not mutually influence one another.

To benefit from mutual influence, a buyer can carry multiple concurrent negotiations, which is to have multiple one-to-one negotiations concurrently with all the sellers. Intuitively, if the buyer obtains a good offer from one of the sellers, it should have more bargaining power when negotiating with other sellers. Then other sellers should give better offers in order to make a deal. In this way, the prices of the sellers mutually influence one another and the buyer can minimize the price.

To enable mutual influence, Nguyen and Jennings [4, 5] propose how to coordinate multiple concurrent negotiations. However, performance of their model depends very much on the accuracy of the buyer's information about the probability distribution of sellers' types. In this paper, we introduce an algorithm, which enables mutual influence in multiple concurrent negotiations. We show that the result of this algorithm actually is in Nash equilibrium [6] when we analyze the negotiation as an extensive game, that is, each player is playing the best response to each other players' strategies.

The rest of the paper is organized as follows. In the next section, we begin the discussion by a motivating example. Then we formalize the algorithm in section 3. Section 4 models and analyzes the negotiation as a game and finds the Nash equilibrium. Finally, section 5 concludes the paper and possible future work is introduced.

2   A Motivating Example

Consider a buyer agent b, which wants to buy a good. Its reservation price, RP^b is \$10, which means it will not buy the good at a price higher than \$10. There are two seller agents, s_1 and s_2. Their reservation prices, RP_1^s and RP_2^s, are \$7 and \$5 respectively, which means the sellers will not sell the good at a price

lower than \$7 and \$5 respectively. Following the convention, we assume that if no agreement can be reached, then no transaction will occur, and all participating agents will effectively get a utility of zero. On the contrary, if an agreement is made between the buyer and seller s_i at a price of \$$p$, then the payoff of the buyer will be $RP^b - p$ and the payoff of the seller will be $p - RP_i^s$.

The buyer starts multiple concurrent negotiations by asking the two sellers for the good. Suppose sellers s_1 and s_2 reply with proposals of \$9 and \$8 for the good respectively. To enable mutual influence, the buyer sends a message to seller s_1, telling it seller s_2 gives an offer of \$8. Here, we assume there is no communication between the two sellers. After knowing that seller s_2 offers the buyer \$8, the best response for seller s_1 is to make a counter proposal by offering a price lower than \$8, say \$7 to the buyer, otherwise it will get zero payoff as the buyer will make an agreement with seller s_2. To enable mutual influence again, the buyer sends a message to seller s_2, telling it that seller s_1 gives an offer of \$7. Then, by knowing seller s_1 offers the buyer \$7, the best response for seller s_2 is to make a counter proposal by offering a price lower than \$7, say \$6 to the buyer. Similarly, the buyer again sends a message to seller s_1, telling it that seller s_2 gives an offer of \$6. However, since \$6 is lower than s_1's reservation price, seller s_1 cannot make a lower offer, otherwise it will get a negative payoff. As seller s_1 cannot make a lower offer, the buyer makes an agreement with seller s_2 at a price of \$6.

Now suppose before the buyer makes an agreement with seller s_2, there is another seller, s_3, whose reservation price is \$4. Then the buyer can send a message to this seller, telling it that someone offers the good at a price of \$6. By mutual influence, this seller makes a counter offer of \$5. The buyer can again tell seller s_2 that seller s_3 offers \$5. Then seller s_2 replies with a counter offer of \$5, but not a price lower than \$5 as this is its reservation price. By mutual influence again, seller s_3 will offer a price lower than \$5, at which the buyer will make an agreement with this seller as this is the lowest price that the buyer can get from these negotiations.

From this example, we can see that buyer can negotiate actively by enabling mutual influence, with which buyer can minimize the price in obtaining the good. In the next section, we formalize a mutual influence algorithm and show how low the price the buyer can minimize.

3   Mutual Influence Algorithm

Suppose there is a buyer b, with reservation price RP^b and there are n sellers s_1 to s_n, with reservation prices RP_1^s to RP_n^s. The mutual influence algorithm can be formalized as follows.

Step 1. Buyer starts negotiations by asking sellers s_1 to s_n for the good, and waits for proposals.

Step 2. Suppose a proposal CO_i^k, from seller s_i, is the minimal among the received offers in round k. Buyer sends messages to seller s_j, where j in $[1, n]$ and $j \neq i$, announcing there is an offer CO_i^k from one of the sellers, and waits for counter offers.

Step 3. Repeat step 2 until $CO_j^m \geq CO_i^{m-1}$.

Step 4. Accept the counter offer CO_i^{m-1} from seller s_i if $CO_i^{m-1} < RP^b$, otherwise no agreement can be made.

In *step 1*, the buyer starts negotiations by asking the sellers for the good. We assume, for simplicity, every seller must reply with a counter offer, even if the counter offer remains the same as the previous offer. In *step 2*, knowing someone offers CO_i^k to the buyer, sellers with offers greater than CO_i^k propose a counter offer greater than their respective reservation prices, but smaller than CO_i^k. Otherwise, the buyer will accept the offer CO_i^k and these sellers will get zero payoff. *Step 2* is repeated until the minimum counter offer the buyer received in round m, from seller s_j, is greater than or equal to that received in round $m-1$, from seller s_i. This means that the price cannot be further lowered by mutual influence. So, in *step 4*, buyer accepts the counter offer CO_i^{m-1}, if this value is smaller than its reservation price. However, if this is not the case, which means that no seller can offer a price lower than the buyer's reservation price, then no agreement can be made.

As a result, the lowest price that the buyer can obtain lies within the lowest and the second lowest reservation prices of all the participating sellers. If the second lowest reservation price of the participating sellers is higher than the buyer's reservation price, then the lowest price that the buyer can obtain lies within the lowest reservation price of the participating sellers and the buyer's reservation price. However, if all the sellers' reservation prices are greater than the buyer's reservation price, then no agreement can be made.

This mutual influence algorithm can be applied if there is no communication between sellers. In this case, since every seller prefers an agreement, sellers are competing with each other and price can be lowered. However, if sellers are allowed to communicate with each other, they can collude not to lower the price. If this is the case, no mutual influence can be made. Another assumption is that we assume buyer always tells the truth and sellers always believe buyer's messages. Obviously, the buyer can fake messages when enabling mutual influence, and sellers have to decide whether or not to believe the buyer's messages. However, this much complicates the situation, and is left to future work.

4   The Negotiation Game and the Nash Equilibrium

4.1   Negotiation Without Mutual Influence Algorithm in One-Buyer-Two-Seller Games

A one-buyer-two-sellers negotiation game without mutual influence algorithm can be modeled as an extensive game with imperfect information [6], as shown in Fig. 1. This game is played by three players: the buyer, seller s_1, and seller s_2. First, seller s_1 chooses its action, which is the selling price, p_1. Then, seller s_2 chooses the selling price, p_2. Since a selling price can be any positive number, there are infinitely many actions for the sellers. Imperfect information means players are imperfectly informed about some or all the choices that have already
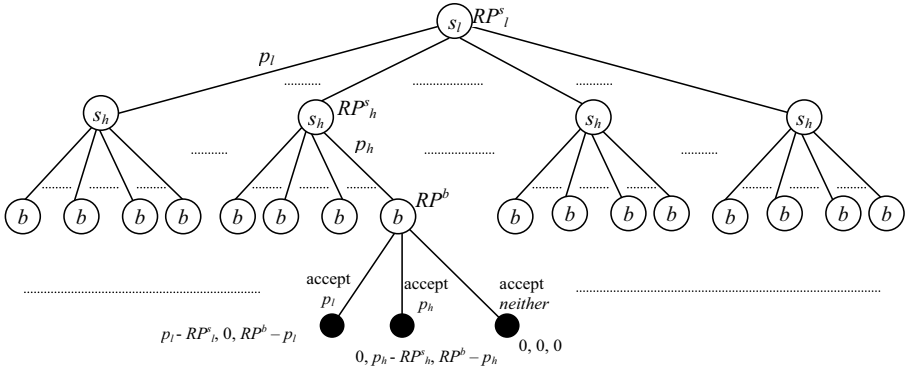
Fig. 1. A one-buyer-two-sellers negotiation game without mutual influence algorithm.

been made. In this example, the sellers do not know what selling price the other seller has chosen. At last, the buyer chooses whether to buy the good from seller s_1, from seller s_2, or accept neither of the proposals. If the buyer chooses to buy the good from seller s_1, then seller s_1 gets a payoff equals to the selling price minus its reservation price, which is $p_1 - RP_1^s$, and seller s_2 gets zero payoff, while the buyer gets a payoff equals to its reservation price minus the selling price, which is $RP^b - p_1$. Similarly, if the buyer chooses to buy the good from seller s_2, then seller s_1 gets zero payoff, seller s_2 a payoff of $p_2 - RP_2^s$, and the buyer gets a payoff of $RP^b - p_2$. If the buyer accepts neither of the proposals, all of the players get zero payoffs. In negotiations without mutual influence algorithm, the buyer then makes counter proposals and the game is repeated. Since negotiation without mutual influence algorithm is an extensive game with imperfect information, there is no pure strategy Nash equilibrium.

4.2 Negotiation with Mutual Influence Algorithm in One-Buyer-Two-Seller Games

With mutual influence algorithm, instead of making counter proposals, the buyer tells seller s_h is the selling price, p_l, that seller s_l has chosen, where s_h the seller proposing a higher selling price, and s_l is the seller proposing a lower selling price. Then the following rounds of negotiations become extensive games with perfect information, as shown in Fig. 2.

When every player in the game is using the best response to every other players' strategies, the game is in Nash equilibrium. Suppose seller s_h chooses a selling price p_h such that $RP_h^s < p_h < p_l$, then the buyer will accept p_h because $RP^b - p_h > RP^b - p_l > 0$. Then, seller s_l will get zero payoff. This is not a Nash equilibrium because seller s_l can increase its payoff by choosing p_l' such that $RP_l^s < p_l' < p_h$. Then, the buyer will accept p_l' because $RP^b - p_l' > RP^b - p_h > 0$. In this case, seller s_h will get zero payoff. If $p_l' < RP_h^s$, where $RP_h^s > RP_l^s$, Nash equilibrium is reached. It is because for seller s_h to increase its payoff, it must
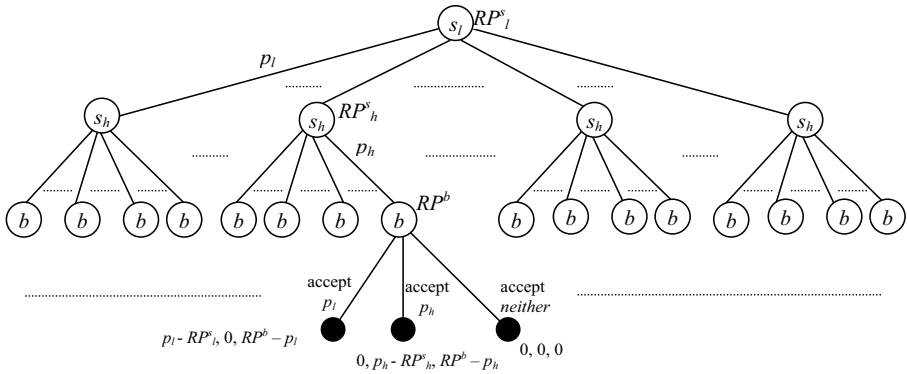
Fig. 2. A one-buyer-two-sellers negotiation game with mutual influence algorithm.

choose p'_h such that $RP^s_h < p'_h < p'_l$, but since $p'_l < RP^s_h$, there is no p'_h which can increase seller s_h's payoff. On the other hand, if $RP^s_l > RP^s_h$, there exists p'_h such that $RP^s_h < p'_h < RP^s_l$, then the buyer accepts p'_h, this is also a Nash equilibrium.

So, the Nash equilibrium is that the buyer accepts fp_1, from seller s_1, such that $RP^s_1 < fp_1 < RP^s_2$ if $RP^s_2 > RP^s_1$ and accepts fp_2, from seller s_2, such that $RP^s_2 < fp_2 < RP^s_1$ if $RP^s_1 > RP^s_2$.

4.3   Negotiation with Mutual Influence Algorithm in One-Buyer-n-Seller Games

For a game with one buyer and n sellers, it can be considered as $n - 1$ one-buyer-two-sellers games. Suppose seller s_i is offering the lowest selling price at the beginning, then the $n - 1$ one-buyer-two-sellers games consist of the buyer, seller s_i and seller s_j, where j in $[1, n]$ and $j \neq i$. Suppose $RP^s_1 < RP^s_2 < \ldots < RP^s_i < \ldots < RP^s_n$, the Nash equilibrium for the $n - 1$ games will be $RP^s_1 < fp_1 < RP^s_i$, $RP^s_2 < fp_2 < RP^s_i$, $\ldots$, $RP^s_{i-1} < fp_{i-1} < RP^s_i$, $RP^s_i < fp_{i+1} < RP^s_{i+1}$, $\ldots$, $RP^s_i < fp_n < RP^s_n$. Then there exists $fp'_1, fp'_2, \ldots, fp'_n$, such that $RP^s_1 < fp'_1 < RP^s_2$, $RP^s_2 < fp'_2 < RP^s_3$, $\ldots$, $RP^s_{n-1} < fp'_n < RP^s_n$, so $RP^s_1 < fp'_1 < RP^s_2 < fp'_2 < \ldots < fp'_n$, the buyer accepts fp'_1 in Nash equilibrium.

From the above analysis, we find that the result from mutual influence algorithm actually is the Nash equilibrium of the negotiation game.

Theorem 1. *By mutual influence algorithm, the lowest price that the buyer can obtain lies within the lowest and the second lowest reservation prices of all the participating sellers.*

Theorem 2. *The result from mutual influence algorithm is a Nash equilibrium.*

5   Conclusions and Future Work

A good price with one seller should enable buyer to have more bargaining power in negotiating with other sellers. This paper presents a mutual influence algorithm in doing so. By simply sending messages to the sellers, telling them the minimum offer the buyer receives so far, the buyer can enable mutual influence among the sellers. On knowing someone is offering the buyer a lower price, sellers are put into competition. The best response is to reply the buyer with a lower offer, otherwise the buyer will accept the lower offer and these sellers will get zero payoff. One important contribution of this algorithm is that its result: the final price that the buyer can obtain lies within the lowest and the second lowest reservation prices among the sellers, is the Nash equilibrium when we model the negotiation as a game.

One important assumption in this algorithm is that sellers are assumed to believe all the messages from the buyers and the buyers never lies. In fact, buyer can always tell lies in messages, telling sellers that another seller is offering a lower price. Then, the sellers may face the problem of having to choose whether to believe the message, and deciding what counter-offers to make. The issues about trust and honesty will be dealt with in future work. Besides, a mutual influence algorithm for multi-issues can also be developed in future work.

References

1. He, M., Jennings, N. and Leung, H. F.: On Agent Mediated Electronic Commerce. IEEE Trans. Knowledge and Data Engineering (2003) 15(4):985-1003
2. Kraus, S.: Strategic Negotiation in Multi-Agent Environments. MIT Press. (2001)
3. Lomuscio, A., Wooldridge, M., Jennings, N. R.: A classification scheme for negotiation in electronic commerce. International Joint Group Decision and Negotiation (2003) 12(1):31-56
4. Nguyen, T. D., Jennings, N. R.: A heuristic model for concurrent bi-lateral negotiations in incomplete information settings. In Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (2003) 1467-1469
5. Nguyen, T. D., Jennings, N. R.: Coordinating multiple concurrent negotiations. In Proceedings of the Third International Conference on Autonomous Agents and Multi-Agent Systems (2004) 1064-1071
6. Osborne, M. J., Rubinstein, A.: A Course in Game Theory. MIT Press.(1994)
7. Rosenschein, J. S., Zlotkin, G.: Rules of Encounter. MIT Press. (1994)
8. Weiss, G.. eds.: Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. The MIT Press. (1999)
9. Wooldridge, M.: An Introduction to MultiAgent Systems. John Wiley & Sons Ltd. (2001)

Vulnerability Evaluation Tools of Matching Algorithm and Integrity Verification in Fingerprint Recognition

Ho-Jun Na[1], Deok-Hyun Yoon[2], Chang-Soo Kim[1,*], and Hyun-Suk Hwang[3]

[1] Pukyong National University, Interdisciplinary Program of Information Security, Korea
nahj@mail1.pknu.ac.kr, cskim@pknu.ac.kr
[2] Hyundai Information Technology, Korea
dhyoon@hit.co.kr
[3] Pukyong National University, Institute of Engineering Research, Korea
hhs@mail1.pknu.ac.kr

Abstract. Personal identification technology is a critical field of information security that has raised various equipments and communications in the modern society. Most studies have concentrated on fingerprint input sensors, recognition, and performance evaluation. However, the vulnerability evaluation for integrity verification of fingerprint recognition systems has partly been studied so far. In this paper, we design and implement two prototype evaluation tools for the vulnerability evaluation of a matching algorithm and integrity evaluation in the fingerprint recognition. We show the results of integrity violation through the process of normal transfer and variant transfer for extracted minutiae data.

1 Introduction

Information needed to maintain security as well as general businesses has been shared on the network environment as computer networks develop. The necessity of authentication of an entity tends upwards in using these application services. Many information security technologies have developed to solve this issue. One of the technologies is biometrics. Biometrics are applied in various industries such as access control systems, security management, Electronic Commerce, missing child finder [1].

However, the most of the research in biometrics have studied algorithms and performance evaluation. Therefore, a performance and security evaluation standard of information security products like the intrusion interception system, the intrusion detection system, and the biometrics system must be also presented. The performance evaluation must be executed by the standard [2].

In this study, we designed and implemented a tool which can verify the integrity, which is necessarily provided in case Biometrics data is transmitted through the network in X9.84 of NIST of USA.

2 Biometrics System

Biometrics is defined as assuring a human with physical and behavior characteristics based on "Are you really you?" or "What do you do, and how do you do it?" [3]. This

---

* Dr. Chang-Soo Kim is the corresponding author.

has an advantage of providing high stability because it is without the inconvenience of needing to be carried, without the riskiness of loss and theft, and the impossibility of estimation different from knowledge-based or token-based technology [4]. The sub components for general model of the biometrics system are as follows [5].

(1) Data collection: Data is collected by input devices and sensors.
(2) Signal processing: Signal processing converts collected data into the data format needed for matching. It calls the processed biometrics data the "template".
(3) Matching: Matching is the comparison of the template taken from signal processing with the templates stored in the storage database.
(4) Storage: The storage system stores the templates originating from the signal processing.

3   Evaluation Method of Biometrics Recognition Security

The evaluation results of occasionally operated manually and subjective decisions are an obstacle when comparing with results of other researchers and developing technology of competition. The evaluation should be executed by an exact standard of a trusted organization to compare with other systems objectively. There are two methods of performance and security evaluation in the biometrics recognition systems [6].

The security evaluation must execute the security function and the assurance evaluation through internationally common criteria. Security of the inside and security requirement items about stability for an attack from outside of the system is required.

We presented two methods of security evaluation and implemented the evaluation tools. Security evaluation for the minutiae matching algorithm executes evaluation through minutiae variation in a fingerprint image variation tool. In addition, the evaluation system evaluates whether integrity function is provided or not when the filtered minutiae from a client in a fingerprint recognition system based on the network are transferred to the server.

4   Vulnerability Evaluation of Matching Algorithm

Vulnerability is defined as the characteristics of a system causing performance decrease or loss to get the results needed illegally in an artificial environment by BWG of England [7]. Vulnerability for security in a decision system means the vulnerability of the threshold value.

The minutiae matching algorithm is the most important part of fingerprint recognition systems, as it decides the final user. The algorithm that matches best is divided into two categories, the minutiae-based algorithm and the correlation-based algorithm [8][9].

The minutiae-based algorithm has two phases, sorting registered templates and inputted minutiae and then counting the minutiae agreed upon after matching.

Next, the correlation-based algorithm matches the ridges of the fingerprint with the valley pattern to sort the ridges. We can apply the evaluation system to the minutiae-based matching algorithm, for which the extraction process is the same.
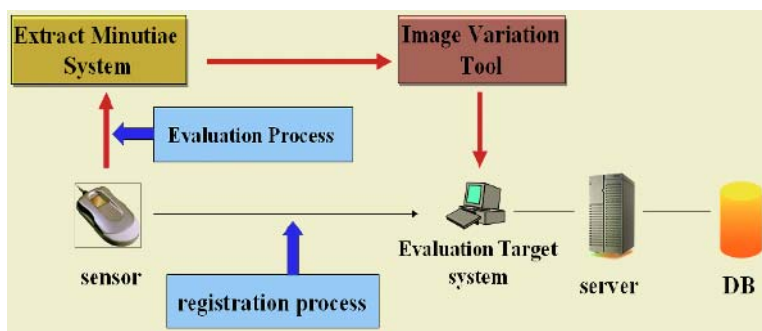
Fig. 1. Matching Algorithm Evaluation Process

4.1 Evaluation Process

The process of the vulnerability evaluation of the matching algorithm in fingerprint recognition systems is shown in Fig.1. The evaluation is divided into the registration and evaluation process.

First, the registration process is the same as the process for general fingerprint recognition systems, which is to store user's fingerprints in the format of templates in a database.

Next, in the evaluation process the minutiae are extracted from the inputted image using the minutiae extraction system. If the information of the minutiae is extracted, the image variation is executed about the minutiae part of the image using an image variation tool. The varied image is inputted into the evaluation system. Matching of the minutiae extracted varied image with the stored template is executed in the evaluation target system. Finally, it is checked that the evaluation system has responded precisely about the varied minutiae.

4.2 Fingerprint Image Variation

Fingerprint image variation based on the minutiae extraction system is executed. The fingerprint image can be varied at the 4 parts of the minutiae extraction process as in Fig.2.

First, variation in the input point and preprocess can be executed at the image without preprocessing of the minutiae extracted from the image attained through the fingerprint input sensor. Second, the variation is executed at the fingerprint image through direction extraction, gray-level enhancement, and binary process. Finally, the variation is executed at the fingerprint image through the session process, which makes the image a width of 1 pixel. We executed the variation of the session image in this study.

4.3 Minutiae Variation Algorithm

The procedure of the variation is as follows: (1) inputting number of needed variation at the dialog (2) selecting variation coordinates using a random function (3) connecting to the point that meets the next ridge as the angle of extracted minutiae (4) ending point disappears and new bifurcation point is created.
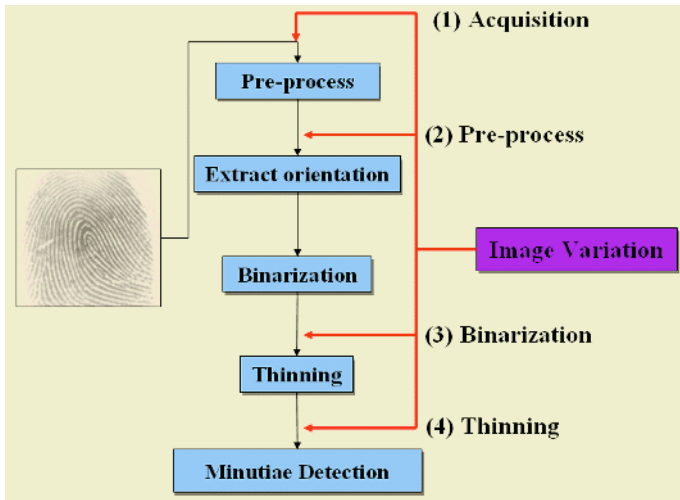
Fig. 2. Points of Image Variation

The selected ending point is changed to the bifurcation of other coordinates as shown in Fig.3. The angle of the ending point selected randomly is 180 degrees if (a). The starting coordinates of variation are the coordinates of the ending point extracted, and the variation is processed (b) if the fingerprint ridge is expanded to where it meets with the next ridge at an angle.
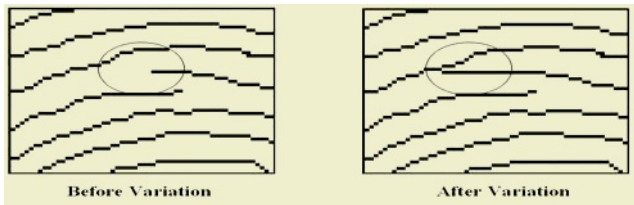
Fig. 3. Minutiae Variation

5   The Integrity Verification of Fingerprint Data

The proposed fingerprint data integrity verification system is largely composed of a client and server module. After executing the variation of the fingerprint data and counting the hash value in the client module, the system encrypts the data and transfers it to the server module. The server module receives the fingerprint data encrypted, and computes the hash value of the fingerprint data and compares it with the hash value after computing the fingerprint data and hash value. The performance process of the integrity verification module is shown in Fig. 4.

We implement a tool which can automatically vary the parts of minutiae records with our developed algorithm and a validation tool which can violate data integrity to verify the integrity function in the data communication. We gather the fingerprint data from FVC2004[10] and each 10 of the left and right hand of our laboratory 5
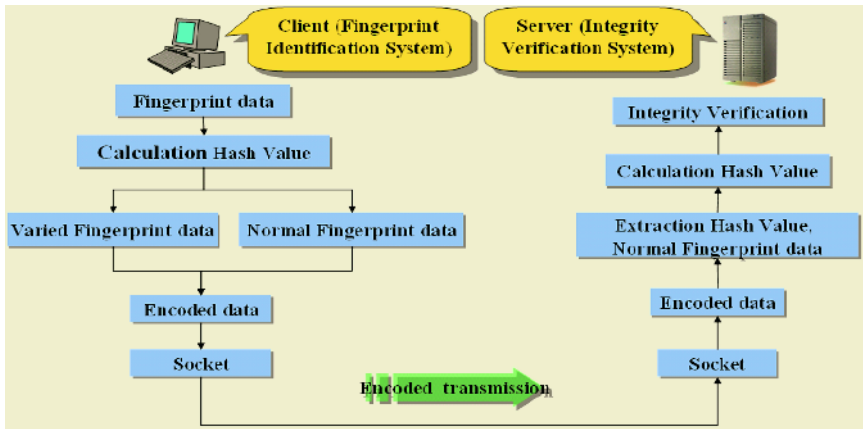
Fig. 4. Process of Variation in Client and Integrity Verification in Server

students to extract minutiae. We experimented on our fingerprint system with two methods. The first method of the evaluation can test whether the fingerprint variation happens or not through the process changing ending point to bifurcation point like the example shown in Fig. 3. The selected fingerprints used FVC2004 which is unchangeable to input data, and the fingerprints of students which are able to get dynamic input data. Therefore, we can test the variation availability by applying the evaluation process to our designed fingerprint systems. The result shows that the minutiae variation is able to be searched in dynamic fingerprint and static fingerprint like FVC2004. Second, we develop a tool which can check up the violation of the integrity when the extracted minutiae data is transmitted. Also, we can assure the integrity violation using our fingerprint system in case the variation is compellingly executed on the way of the transmission. Therefore, this research shows the results about whether the variation happens or not for the fingerprint recognition systems which provide encryption and integrity function.

We use the hash function of MD5 and SHA-1 for the integrity and use SEED and Triple-DES algorithm for the security.

The vulnerability points of security are largely divided into three parts, sensor device/client, data transmission, and security evaluation system. Table 1 shows the research results of security evaluation about the fingerprint recognition system. The security evaluation of the Knag at al.[11] is executed at the sensor device/client, and our method is executed at both the sensor device/client and the data transmission. However, the research of the security evaluation for server system has not yet been found in fingerprint system.

Table 1. Evaluation Points of Security Vulnerability

Security Vulnerability Points	Kang et al.[11]	Our Method
Sensor Device/Client	O	O
Data Transmission	X	O
Security Evaluation(Server) System	X	X

Fig. 5. Normal Data Integrity Verification

Fig. 6. Varied Data Integrity Verification

6   Experiment Results

We experimented on the integrity evaluation of the fingerprint data, which is to verify that the systems provide the function of integrity through the process of normal transfer and variant transfer. Fig.5 shows the results for the transferring of normal fingerprint data.

The results of normal transfer through the received hash value and the computed hash value are shown above. The results of transferring the varied fingerprint data are shown in Fig.6. We know that the fingerprint is varied because the received hash value by the variation of the second field and the computed hash value do not agree.

7   Conclusion

In recent years, a variety of biometrics recognition system products have been released over the world since its market was introduced. In particular, the biometrics recognition systems, which can assure rapid and exact identification, have concentrated on the interest since the 911 terrorism. Therefore, each country's products approach the standard to get the profits of the biometrics industry.

In addition, they have decided upon the security criteria of biometrics recognition systems and have announced that organizations must use the products along with the criteria. In future, it is expected that the biometrics recognition systems like other security products will be evaluated and granted a grade. This objective and exact evaluation have an advantage which is to speed up the development of the industry. However, its disadvantage can also come into existence. Reliable organizations must decide the criteria, and the results of the evaluation should also be carefully managed.

We designed and implemented the prototype of a tool which can evaluate an issue in setting up the threshold value and can verify the integrity function. The setting of the threshold value is a decision criterion of the matching algorithm of security vulnerability presented by BWG of England. Also, the integrity function is necessarily provided in case the biometrics recognition data is transferred through the network in X9.84 of NIST of USA.

References

1. Elsevier Advanced Technology, "The Biometric Industry Report : Market and Technology Forecast to 2003", pp.3-11, (2001)
2. U.S DOD, BMO.: Biometric System protection Profile For Medium Robustness Environments. (2001)
3. Lin Hong,: Automatic Personal Identification Using Fingerprints. A Dissertation of Computer Science, Michigan State University, (1998)
4. Yager Neil and Amin adnan : Fingerprint verification based on minutiae feature: a review", Pattern Analysis and Applications, Vol.7, No.1, pp94-113, (2004)
5. Anil Jain, Lin Hong, and Ruud Bolle: On-Line Fingerprint Verification. IEEE Transactions on Pattern Analysis Machine Intelligence, Vol. 19, No. 4, pp.302-314, (1997)
6. D.H.,Yoon, C.S.Kim, K.H.Rhee, G.P.Shim, J.S.Kim, B.G.No: A study on the security vulnerability evaluation for fingerprint recognition systems, EALPIIT2002, pp. 373-378, (2002)
7. UK Government Biometrics Working Group: Principles of Biometric Security System Vulnerability Assessment. (2001)
8. Tico. M. and Kuosmanen P.: Fingerprint matching using an orientation-based minutia descriptor. Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol.25 No 8, pp1009-1014, (2003)
9. Gu J., Zhou J., and Zhang D.: A Combination model for orientation field Fingerprint. Pattern Recognition, Vol. 37, No. 3, pp543-553, (2004)
10. http://bias.csr.unibo.it/fvc2004: (2004)
11. Kang, H., Lee, B., Kim, H., Shin, D.: A Study on Performance Evaluation of the Liveness Detection for Various Fingerprint Sensor Modules, LNCS 2774, pp.1245-1253, (2003)

Algorithms for CTL System Modification

Yulin Ding and Yan Zhang

School of Computing & Information Technology
University of Western Sydney
Kingswood, N.S.W. 1797, Australia
{yding,yan}@cit.uws.edu.au

Abstract. Model updating, as a new concept to be employed as a standard and universal method for system modification, has been started in [4] and further developed in this paper. This paper introduces algorithms for Computation Tree Logic (CTL) model update. The algorithms correct errors in a CTL Kripke model to make this model satisfy its various required properties. These update algorithms are designed through pseudo-code, which details the logic of the algorithms. The microwave oven example in [3] has been correctly treated by the pseudo-code, which demonstrates the feasibility of these algorithms. The algorithms and their pseudo-code are the foundations for later implementation of a CTL model updater and integration of a CTL model checker and this updater.

1 Introduction

As one of the most promising formal methods, automated verification has played an important role in computer science development. Currently, model checkers with SMV or Promela [5] series as their specification languages are widely available for research, experiment, and partial industry usage. Nowadays SMV [3], NuSMV, Cadence SMV and SPIN [5] are well accepted as the state of the art model checkers. Recently, error diagnosis and repair methods have begun to employ a formal methods approach. Buccafurri and his colleagues [2] used abductive model revision techniques, repairing errors in concurrent programs. It is a new approach towards automated diagnosis and integrated repair on model checking. Also, the update of the knowledge base is enhanced by the modality update by Baral and Zhang [1]. They discussed knowledge update and its minimal change, based on modal logic $S5$. Furthermore, Ding and Zhang [4] employed a formal approach called Linear Temporal Logic (LTL) model update for system modification, which is the first step of the theoretical integration of model checking and knowledge update. The LTL model update modifies the existing LTL model of an abstracted system to automatically correct the errors occurring within this model.

We intend to apply model update to develop a practical Computation Tree Logic (CTL) model updater, which represents a universal tool to repair errors automatically. In order to extend the work of Ding and Zhang [4] from LTL model update to CTL model update, we introduce algorithms of CTL model update as the main contribution of this paper. As preparation for implementation of

CTL model update, we focus on the pseudo-code for CTL model update, which is compatible with SAT, the pseudo-code for model checking as presented in [3, 6]. This compatibility makes our algorithms more standard and portable for later integration of the CTL model updater into the CTL model checker.

2 Syntax and Semantics of the CTL Kripke Model

Definition 1. *[3] Let AP be a set of atomic propositions. A Kripke model M over AP is a three tuple $M = (S, R, L)$ where*

1. S is a finite set of states and a state is a set of atoms.
2. $R \subseteq S \times S$ is a transition relation.
3. $L : S \to 2^{AP}$ is a function that assigns each state with a set of atomic propositions.

Definition 2. *[6] Computation tree logic (CTL) has the following syntax given in Backus naur form:*

$$\phi ::= \top | \perp |p|(\neg\phi)|(\phi \wedge \phi)|(\phi \vee \phi)|\phi \to \phi|AX\phi|EX\phi$$
$$|AG\phi|EG\phi|AF\phi|EF\phi|A[\phi \cup \phi]|E[\phi \cup \phi]$$

where p is any propositional atom.

Definition 3. *[6] Let $M = (S, R, L)$ be a Kripke model for CTL. Given any s in S, we define whether a CTL formula ϕ holds in state s. We denote this by $M, s \models \phi$. Naturally, the definition of the satisfaction relation $\models$ is done by structural induction on all CTL formulas:*

1. *$M, s \models \top$ and $M, s \not\models \perp$ for all $s \in S$.*
2. *$M, s \models p$ iff $p \in L(s)$.*
3. *$M, s \models \neg\phi$ iff $M, s \not\models \phi$.*
4. *$M, s \models \phi_1 \wedge \phi_2$ iff $M, s \models \phi_1$ and $M, s \models \phi_2$.*
5. *$M, s \models \phi_1 \vee \phi_2$ iff $M, s \models \phi_1$ and $M, s \models \phi_2$.*
6. *$M, s \models \phi_1 \to \phi_2$ iff $M, s \not\models \phi_1$, or $M, s \models \phi_2$.*
7. *$M, s \models AX\phi$ iff for all s_1 such that $s \to s_1$ we have $M, s_1 \models \phi$.*
8. *$M, s \models EX\phi$ iff for some s_1 such that $s \to s_1$ we have $M, s_1 \models \phi$.*
9. *$M, s \models AG\phi$ holds iff for all paths $s_0 \to s_1 \to s_2 \to \cdots$, where s_0 equals s, and all s_i along the path, we have $M, s_i \models \phi$.*
10. *$M, s \models EG\phi$ holds iff there is a path $s_0 \to s_1 \to s_2 \to \cdots$, where s_0 equals s, and for all s_i along the path, we have $M, s_i \models \phi$.*
11. *$M, s \models AF\phi$ holds iff for all paths $s_0 \to s_1 \to s_2 \to \cdots$, where s_0 equals s, there is some s_i such that $M, s_i \models \phi$.*
12. *$M, s \models EF\phi$ holds iff there is a path $s_0 \to s_1 \to s_2 \to \cdots$, where $s_i = s$, and for some s_i along the path, we have $M, s_i \models \phi$.*
13. *$M, s \models A[\phi_1 \cup \phi_2]$ holds iff for all paths $s_0 \to s_1 \to s_2 \to \cdots$, where s_0 equals s, that path satisfies $\phi_1 \cup \phi_2$, i.e. there is some s_i along the path, such that $M, s_i \models \phi_2$, and, for each $j < i$, we have $M, s_j \models \phi_1$.*
14. *$M, s \models E[\phi_1 \cup \phi_2]$ holds iff there is a path $s_0 \to s_1 \to s_2 \to \cdots$, where s_0 equals s, that path satisfies $\phi_1 \cup \phi_2$, i.e. there is some s_i along the path, such that $M, s_i \models \phi_2$, and, for each $j < i$, we have $M, s_j \models \phi_1$.*

A CTL formula is evaluated on a Kripke model M. A path in M from a state s is an infinite sequence of states $\pi \stackrel{def}{=} [s_0, s_1, \cdots, s_{i-1}, s_i, s_{i+1}, \cdots]$ such that $s_0 = s$ and the relation (s_i, s_{i+1}) holds for all $i \geq 0$. We donate $(pre(s_i), s_i) \in R$, where $s_{i-1} = pre(s_i)$, and $(s_i, succ(s)) \in R$, where $s_{i+1} = succ(s_i)$.

3   Algorithms for CTL Model Update

Definition 4. (**CTL Model Update**) *Given a CTL Kripke model* $M = (S, R, L)$ *and a CTL formula* ϕ. $\mathcal{M} = (M, s_0) \not\models \phi$, *where* $s_0 \in S$. *An update of* $\mathcal{M}$ *with* ϕ, *is a new CTL Kripke model* $\mathcal{M}' = (M', s_0')$, *where* $M' = (S', R', L')$ *and* $s_0' \in S'$, *such that* $\mathcal{M}' \models \phi$.

function CTL*UPDATE($\mathcal{M},\phi$) /* $\mathcal{M} \not\models \phi$. Update $\mathcal{M}$ to satisfy ϕ. */
INPUT $M = (S, R, L)$, $\mathcal{M} = (M, s_0)$, where $s_0 \in S$ and $\mathcal{M} \not\models \phi$;
OUTPUT $M' = (S', R', L')$, $\mathcal{M}' = (M', s_0')$, where $s_0' \in S'$ and $\mathcal{M}' \models \phi'$;
{ case
    ϕ is $\bot$: return(M);
    ϕ is atomic p : return {UPDATE$_p(\mathcal{M},p)$};
    ϕ is $\neg\phi_1$: return{UPDATE$_\neg(\mathcal{M},\phi_1)$};
    ϕ is $\phi_1 \vee \phi_2$: return{CTL*UPDATE($\mathcal{M},\phi_1$) or CTL*UPDATE($\mathcal{M},\phi_2$)};
    ϕ is $\phi_1 \wedge \phi_2$: return UPDATE$_\wedge(\mathcal{M},\phi_1,\phi_2)$;
    ϕ is $EX\phi_1$: return UPDATE$_{EX}(\mathcal{M},\phi_1)$;
    ϕ is $AX\phi_1$: return UPDATE$_{AX}(\mathcal{M},\phi_1)$;
    ϕ is $EF\phi_1$: return UPDATE$_{EF}(\mathcal{M},\phi_1)$;
    ϕ is $AF\phi_1$: return UPDATE$_{AF}(\mathcal{M},\phi_1)$;
    ϕ is $EG\phi_1$: return UPDATE$_{EG}(\mathcal{M},\phi_1)$;
    ϕ is $AG\phi_1$: return UPDATE$_{AG}(\mathcal{M},\phi_1)$;
    ϕ is $E(\phi_1 \cup \phi_2)$: return UPDATE$_{EU}(\mathcal{M},\phi_1,\phi_2)$;
    ϕ is $A(\phi_1 \cup \phi_2)$: return UPDATE$_{AU}(\mathcal{M},\phi_1,\phi_2)$;
}
function UPDATE$_p(\mathcal{M},p)$ /* $\mathcal{M} \not\models p$. Update s_0 to satisfy p. */
{ 1. $s_0' := s_0 \cup \{p\}$;
  2. $S' := S - \{s_0\} \cup \{s_0'\}$;
  3. $R' := R - \{(s_0, s_i) \mid \forall s_i = succ(s_0)\} \cup \{(s_0', s_i) \mid \forall s_i = succ(s_0)\} -$
              $\{(s_j, s_0) \mid \forall s_j = pre(s_0)\} \cup \{(s_j, s_0') \mid \forall s_j = pre(s_0)\}$;
  4. $L': S' \rightarrow 2^{AP}$, where $\forall s \in S'$, if $s \in S$, then $L'(s) = L(s)$;
    else $s = s_0'$, and $L(s_0') :=$ the set of atoms occurring in s_0' as defined above;
  5. $\mathcal{M}' := (M', s_0')$, where $M' = (S', R', L')$ and $s_0' \in S'$;
  6. return $\{\mathcal{M}'\}$;
}
function UPDATE$_\neg(\mathcal{M},\phi)$ /* $\mathcal{M} \not\models \phi$. Update $\mathcal{M}$ to satisfy ϕ. */
{ case
    ϕ is $\neg p$:   1. $s_0' = s_0 - \{p\}$;
                 2.— 6. are the same as those in the function UPDATE$_p(\mathcal{M},p)$;
    ϕ is $\neg(\phi_1 \vee \phi_2) = \neg\phi_1 \wedge \neg\phi_2$: return{UPDATE$_\wedge(\mathcal{M},\neg\phi_1,\neg\phi_2)$};

ϕ is $\neg(\phi_1 \wedge \phi_2) = \neg\phi_1 \vee \neg\phi_2$: return $\{\text{UPDATE}_\neg(\mathcal{M},\neg\phi_1)$
    or $\text{UPDATE}_\neg(\mathcal{M},\neg\phi_2)\}$;

ϕ is $\neg EX(\phi_1) = AX(\neg\phi_1)$: return$\{\text{UPDATE}_{AX}(\mathcal{M},\neg\phi_1)$ $\}$;

ϕ is $\neg AX(\phi_1) = EX(\neg\phi_1)$: return$\{\text{UPDATE}_{EX}(\mathcal{M},\neg\phi_1)\}$;

ϕ is $\neg EF(\phi_1) = AG(\neg\phi_1)$: return$\{\text{UPDATE}_{AG}(\mathcal{M},\neg\phi_1)\}$;

ϕ is $\neg AF(\phi_1) = EG(\neg\phi_1)$: return$\{\text{UPDATE}_{EG}(\mathcal{M},\neg\phi_1)\}$;

ϕ is $\neg EG(\phi_1) = AF(\neg\phi_1)$: return$\{\text{UPDATE}_{AF}(\mathcal{M},\neg\phi_1)\}$;

ϕ is $\neg AG(\phi_1) = EF(\neg\phi_1)$: return$\{\text{UPDATE}_{EF}(\mathcal{M},\neg\phi_1)\}$;

ϕ is $\neg E(\phi_1 \cup \phi_2)$: return$\{\text{UPDATE}_{\neg EU}(\mathcal{M},\phi_1,\phi_2)\}$;

ϕ is $\neg A(\phi_1 \cup \phi_2) = E[\neg\phi_2 \cup (\neg\phi_1 \wedge \neg\phi_2)] \vee EG\neg\phi_2$:
    return$\{\text{UPDATE}_{EU}(\mathcal{M},\neg\phi_2,\neg\phi_1 \wedge \neg\phi_2)$ or $\text{UPDATE}_{EG}(\mathcal{M},\neg\phi_2)$ $\}$;
}

function $\text{UPDATE}_\wedge(\mathcal{M},\phi_1,\phi_2)$/*$\mathcal{M}\not\models \phi_2 \wedge \phi_2$. Update $\mathcal{M}$ to satisfy $\phi_1 \wedge \phi_2$.*/
{ 1. $\mathcal{M}'$ =CTL*UPDATE(CTL*UPDATE($\mathcal{M},\phi_1$),ϕ_2);
  2. if $M' \models \phi_1 \wedge \phi_2$, then return $\{\mathcal{M}'\}$; else $\mathcal{M}'$ =$\text{UPDATE}_\wedge(\mathcal{M}',\phi_1,\phi_2)$;
}

function $\text{UPDATE}_{EX}(\mathcal{M},\phi)$ /* $\mathcal{M}\not\models EX\phi$. Update $\mathcal{M}$ to satisfy $EX\phi$. */
{ 1. Select a state $s_1 = succ(s_0)$, such that $\mathcal{M}_1=(M,s_1) \not\models \phi$;
  2. return $\{\text{CTL*UPDATE}(\mathcal{M}_1,\phi)\}$;
}

function $\text{UPDATE}_{AX}(\mathcal{M},\phi)$ /*$\mathcal{M}\not\models AX\phi$. Update $\mathcal{M}$ to satisfy $AX\phi$. */
{ 1. Select a state $s_1 = succ(s_0)$, such that $\mathcal{M}_1 = (M,s_1) \not\models \phi$;
  2. $\mathcal{M}':=$ CTL*UPDATE($\mathcal{M}_1,\phi$);
  3. if $\mathcal{M}'\models AX\phi$, then return $\{\mathcal{M}'$ $\}$; else $\text{UPDATE}_{AX}(\mathcal{M}',\phi)$;
}

function $\text{UPDATE}_{EF}(\mathcal{M},\phi)$ /*$\mathcal{M}\not\models EF\phi$. Update $\mathcal{M}$ to satisfy $EF\phi$. */
{ 1. Select a state s_i on a path $\pi = [s_0,s_1,\cdots,s_i,\cdots]$;
    2. return $\{\text{CTL*UPDATE}(\mathcal{M}_i,\phi)\}$; /* where $\mathcal{M}_i= (M,s_i)$ */
}

function $\text{UPDATE}_{AF}(\mathcal{M},\phi)$ /* $\mathcal{M}\not\models AF\phi$. Update $\mathcal{M}$ to satisfy $AF\phi$. */
{ 1. Select a state s_i on a path $\pi = [s_0,s_1,\cdots,s_i,\cdots]$,
    such that $\mathcal{M}_i = (M,s_i) \not\models \phi$;
  2. $\mathcal{M}':=$CTL*UPDATE($\mathcal{M}_i,\phi$);
  3. if $\mathcal{M}'\models AF\phi$, then return $\{\mathcal{M}'$ $\}$; else $\text{UPDATE}_{AF}(\mathcal{M}',\phi)$;
}

function $\text{UPDATE}_{EG}(\mathcal{M},\phi)$ /*$\mathcal{M}\not\models EG\phi$. Update $\mathcal{M}$ to satisfy $EG\phi$. */
{ 1. Select a path $\pi = [s_0,s_1,\cdots,s_i,\cdots]$ in $\mathcal{M}$;
  2. Select a state $s_i \in \pi$, such that $\mathcal{M}_i= (M,s_i) \not\models \phi$;
  3. $\mathcal{M}':=$CTL*UPDATE($\mathcal{M}_i,\phi$);
  4. if $\mathcal{M}'\models EG\phi$, then return $\{\mathcal{M}'$ $\}$; else $\text{UPDATE}_{EG}(\mathcal{M}',\phi)$;
}

function $\text{UPDATE}_{AG}(\mathcal{M},\phi)$ /* $\mathcal{M}\not\models AG\phi$. Update $\mathcal{M}$ to satisfy $AG\phi$. */
{ 1. Select a path $\pi = [s_0,s_1,\cdots,s_i,\cdots]$, where $s_i \in \pi$ and $\mathcal{M}_i=(M,s_i) \not\models \phi$;
  2. $\mathcal{M}':=$CTL*UPDATE($\mathcal{M}_i,\phi$);
  3. if $\mathcal{M}'\models AG\phi$, then return $\{\mathcal{M}'$ $\}$; else $\text{UPDATE}_{AG}(\mathcal{M}',\phi)$;
}

function $\text{UPDATE}_{EU}(\mathcal{M},\phi_1,\phi_2)$
/* $\mathcal{M} \not\models E(\phi_1 \cup \phi_2)$, update $\mathcal{M}$ to satisfy $E(\phi_1 \cup \phi_2)$. */
{ 1. Select any path $\pi = [s_0, s_1, \cdots, s_j, \cdots, s_i, \cdots]$;
  2. If($\forall s_i \in \pi, \mathcal{M}_i \not\models \phi_2$), $\mathcal{M}':=\text{CTL*UPDATE}(\mathcal{M}_i,\phi_2)$;
      /* where $\mathcal{M}_i = (M, s_i)$. */
  3.If $\mathcal{M}' \models E(\phi_1 \cup \phi_2)$, return$\{\mathcal{M}'\}$;
  4. else Let s_i be the earliest state in π, such that $\mathcal{M}_i \models \phi_2$;
          Select s_j ($j < i$), such that $(M, s_j) \not\models \phi_1$;
          $\mathcal{M}':=\text{CTL*UPDATE}(\mathcal{M}'_j,\phi_1)$; /* where $\mathcal{M}'_j = (M', s_j)$. */
  5. If $\mathcal{M}' \models E(\phi_1 \cup \phi_2)$, return$\{\mathcal{M}'\}$; else $\text{UPDATE}_{EU}(\mathcal{M}',\phi_1,\phi_2)$;
}
function $\text{UPDATE}_{AU}(\mathcal{M},\phi_1,\phi_2)$
/* $\mathcal{M} = (M, s_0) \not\models A(\phi_1 \cup \phi_2)$. Update $\mathcal{M}$ to satisfy $A(\phi_1 \cup \phi_2)$. */
{ 1. Select a path $\pi = [s_0, s_1, \cdots, s_j, \cdots, s_i, \cdots]$,
      where $\mathcal{M}_j = (M, s_j) \not\models \phi_1$ or $\mathcal{M}_i = (M, s_i) \not\models \phi_2$;
  2. If($\mathcal{M}_i \not\models \phi_2$),$\mathcal{M}':=\text{CTL*UPDATE}(\mathcal{M}_i,\phi_2)$;
  3.If $\mathcal{M}' \models A(\phi_1 \cup \phi_2)$, return$\{\mathcal{M}'\}$;
  4. else $\mathcal{M}':=\text{CTL*UPDATE}(\mathcal{M}'_j,\phi_1)$; /* where $\mathcal{M}'_j = (M', s_j)$. */
  5. If $\mathcal{M}' \models A(\phi_1 \cup \phi_2)$, return$\{\mathcal{M}'\}$; else $\text{UPDATE}_{AU}(\mathcal{M}',\phi_1,\phi_2)$;
}
function $\text{UPDATE}_{\neg EU}(\mathcal{M},\phi_1,\phi_2)$
/*$\mathcal{M} = (M, s_0) \not\models \neg E(\phi_1 \cup \phi_2)$. Update $\mathcal{M}$ to satisfy $\neg E(\phi_1 \cup \phi_2)$. */
{ 1. Select a path $\pi = [s_0, s_1, \cdots, s_j, \cdots, s_i, \cdots]$,
      where $\mathcal{M}_j = (M, s_j) \models \phi_1$ and $\mathcal{M}_i = (M, s_i) \models \phi_2$;
  2. $\mathcal{M}':=\text{CTL*UPDATE}(\mathcal{M}_i,\neg\phi_2)$, or $\mathcal{M}':=\text{CTL*UPDATE}(\mathcal{M}_j,\neg\phi_1)$;
      /* A nondeterministic choice. */
  3. If $\mathcal{M}' \models \neg E(\phi_1 \cup \phi_2)$, return$\{\mathcal{M}'\}$ else $\text{UPDATE}_{\neg EU}(\mathcal{M}',\phi_1,\phi_2)$;
}

The pseudo-code begins with the main function CTL*UPDATE($\mathcal{M},\phi$), where all sub-functions call either the main function itself, or other sub-functions. These sub-functions then call either themselves or further sub-functions. Eventually the lowest level sub-functions always call the main function, which forms recursive calls. No matter how many recursive calls are executed, the final functions to perform the update are either the main function for the case of "ϕ is atomic p" by returning $\text{UPDATE}_p(\mathcal{M},\phi)$ or the function $\text{UPDATE}_\neg(\mathcal{M},\neg\phi)$ for the case of "$\neg\phi$ is $\neg p$".

The update is eventually finalized in the state. The function $\text{UPDATE}_p(\mathcal{M},p)$ is a lowest level call in the whole set of algorithms and contains the update of the smallest element,the propositional atom. In this function, the updated state s'_0 and three tuples of $\mathcal{M}'$:- S',R' and L' are assigned. These new assignments form a new updated model $\mathcal{M}'$, which is returned by the function $\text{UPDATE}_p(\mathcal{M},p)$.

The function $\text{UPDATE}_{AX}(\mathcal{M},\phi)$ is an intermediate function and called by the main function. In this function, every process of an execution (including 3 steps) only updates one successor of s_0 by recursively calling the main function. The final results should be that all successors of s_0 satisfy ϕ. After each call of

this function, it checks whether the updated model $\mathcal{M}'$ satisfies $AX\phi$ or not. If it does not, the function calls itself again recursively. Otherwise it returns the new updated model $\mathcal{M}'$.

4   An Example

We present a case study to illustrate features of our CTL model update approach. For a microwave oven scenario [3], we assume that the processes of a microwave oven include both a normal heat process and a faulty process. For the normal process, no error occurs. The oven is closed to heat food. For the faulty process, the oven will not heat food after it is started. The purpose of model checking is to identify that a faulty process exists. The purpose of model updating, on the other hand, is to correct the original model which contains the faulty process. Fig. 1 gives the Kripke structure for this Microwave oven. s_1, s_2, $\cdots$, s_7 are states of this system. The path $s_1 \to s_2 \to s_5 \to s_3$ is the faulty process. The path $s_1 \to s_3 \to s_6 \to s_7 \to s_4$ is a normal heat process.
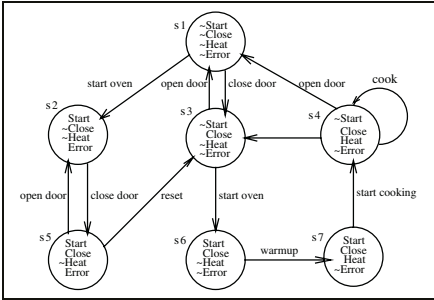
Fig. 1. The Original CTL Kripke Structure of a Microwave Oven.
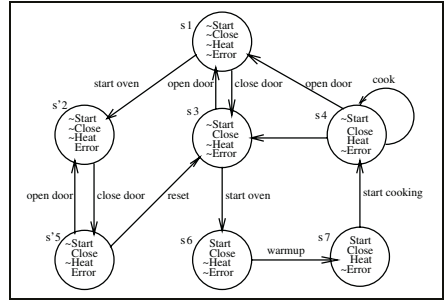
Fig. 2. The Updated CTL Kripke Structure of the Microwave Oven.

The formal definition of the Kripke structure of the microwave oven is given: $M = (S, R, L)$, where $S = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7\}$, $R = \{(s_1, s_2), (s_2, s_5), (s_5, s_2), (s_5, s_3), (s_3, s_1), (s_1, s_3), (s_3, s_6), (s_6, s_7), (s_7, s_4), (s_4, s_4), (s_4, s_3)(s_4, s_1)\}$, $AP = \{Start, Close, Heat, Error\}$, L assigns states s_1, s_2, s_3, s_4, s_5, s_6 and s_7 in M with $\{\neg Start, \neg Close, \neg Heat, \neg Error\}$, $\{Start, \neg Close, \neg Heat, Error\}$, $\{\neg Start, Close, \neg Heat, \neg Error\}$, $\{\neg Start, Close, Heat, \neg Error\}$, $\{Start, Close, \neg Heat, Error\}$, $\{Start, Close, \neg Heat, \neg Error\}$, $\{Start, Close, Heat, \neg Error\}$ respectively.

The CTL formula $\neg EF(Start \wedge EG \neg Heat)$ is checked. From the model checking, The loop path π containing states $\{s_1, s_2, s_5, s_3\}$ satisfies $EG \neg Heat$ with starting state s_1. The states on this loop which satisfy $Start$ are states s_2 and s_5. Thus, states satisfying $EF(Start \wedge EG \neg Heat)$ are s_2 and s_5. We need the Kripke structure not to have any state which satisfies $EF(Start \wedge EG \neg Heat)$. The result of the model checking is that the microwave oven system described by the Kripke structure does not satisfy the given specification. Thus, we need to update the system to satisfy the specification. We use our algorithms in the

previous section to update this Kripke structure. This also serves to check the correctness of our algorithms.

Step 1: in the function CTL*UPDATE($\mathcal{M}$,ϕ), the function UPDATE$_\neg$($\mathcal{M}$,$\neg\phi$) is called. In our case, $\neg\phi$ is $\neg EF(Start \wedge EG\neg Heat)$.

Step 2: in the function UPDATE$_\neg$($\mathcal{M}$,$\neg\phi$), our case is "$\neg\phi$ is $\neg EF(\phi_1) = AG(\neg\phi_1)$", where ϕ_1 is $(Start \wedge EG\neg Heat)$, it returns UPDATE$_{AG}$($\mathcal{M}$,$\neg(Start \wedge EG\neg Heat)$).

Step 3: the function UPDATE$_{AG}$($\mathcal{M}$,$\neg(Start \wedge EG\neg Heat)$) calls the main function CTL*UPDATE($\mathcal{M}$,$\neg Start \vee \neg EG\neg Heat$). This function will return CTL*UPDATE($\mathcal{M}$,$\neg Start$) or CTL*UPDATE($\mathcal{M}$,$\neg EG\neg Heat$) for the case $\phi_1 \vee \phi_2$. For simplicity, we choose the former one.

We need to update the atomic proposition $Start$ in states s_2 and s_5 of path π with $\neg Start$ instead, then no states on path π have the specification $EF(Start \wedge EG\neg Heat)$. That is $\mathcal{M}' = (M', s_1) \models \neg EF(Start \wedge EG\neg Heat)$.

The new updated Kripke structure is as Fig. 2. The results are from the CTL update algorithms described in this paper and minimal change rules which will be described in our forthcoming paper.

5   Conclusions

This paper has presented core algorithms for CTL model update. These algorithms are described using pseudo-code, which will be implemented later as a major part of CTL model update. The algorithms are the extended work of Ding and Zhang from paper [4] and has provided solutions for updating the CTL Kripke model to satisfy its major semantics such as AG and AF. The well known microwave oven example [3] has been correctly updated by the pseudo-code. The pseudo-code is compatible with the SAT pseudo-code for CTL model checking as presented in [3, 6]. This compatibility is the foundation for later integration of the model checker and the model updater. The algorithms in this paper have provided a promising approach for CTL system modification.

References

1. Baral, C. and Zhang, Y. (2005). Knowledge updates: semantics and complexity issues. To appear Artificial Intelligence (AIJ), 2005.
2. Buccafurri, F. et al.(1999). Enhancing model checking in verification by AI techniques. Artificial Intelligence 112(1999) 57-104.
3. Clarke, E. Jr. et al.(1999), Model checking, The MIT press, Cambridge, Massachusetts, London, England.
4. Ding, Y. and Zhang, Y. (2005). A logic approach for LTL system modification. In proceedings of the 15th International Symposium on Methodologies for Intelligent Systems. 25-28 May, 2005. Saratoga Springs, New York.
5. Holzmann, G. (2003). The SPIN model checker: primer and reference manual. Addison-Wesley Professional.
6. Huth, M. and Ryan, M. (2000), Logic in computer science: modelling and reasoning about systems. University Press, Canbridge.

A Robust Approach for Improving Computational Efficiency of Order-Picking Problems

Yu-Min Chiang[1], Shih-Hsin Chen[2], and Kuo-Chang Wu[1]

[1] Department of Industrial Engineering and Management, I-Shou University
1, Section 1, Hsueh-Cheng Rd., Ta-Hsu Hsiang,
Kaohsiung County, Taiwan 840, R.O.C
`ymchiang@isu.edu.tw, m9320007@stmail.isu.edu.tw`
[2] Department of Industrial Engineering and Management, Yuan Ze University
No. 135, Yuan-Tung Rd., Chung-Li, Taoyuan, Taiwan, R.O.C
`s939506@mail.yzu.edu.tw`

Abstract. The new market forces have affected the operation of distribution centers (DCs) extremely. The DCs demand an increased productivity and a low cost in their daily operation. This paper focuses on the order-picking operation, which is the core operation of DCs, to improve the material-handling process and enhance the competitiveness. The study applied an integrated framework that combined two algorithms, branch-and-bound and tabu search, to the order-picking problem and developed a neighborhood selection criterion based on constraint propagation and the prune-and-search algorithm to improve computational efficiency. The proposed method forbids bad node when searching a solution by branch-and-bound algorithm or tabu search. Besides, in order to make the algorithm more robust and reasonable, the fuzzy modeling was employed. The empirical result shows the approach decreases the computational effort effectively on the both algorithms. The study may be of interest for the researchers in the optimization of DCs' operations and the DCs' managers.

1 Introduction

Distribution centers (DCs) play an important role in modern supply chain environment. It coordinates the manufacturers and customers in the supply chain. However, the material-handling process takes a large proportion of cost, which damaging the survival of DCs [1]. Because the competition is keen in the logistics industry, the study focuses on the order picking operation, which is the core operation of DCs, to improve the material-handling process and enhance their competitiveness.

There are several researches concerning the order-picking problem of DCs [1-5]. Generally, the problems in these models include order-picking within single aisle, weight constraints to zone picking area, number of stock location, maximum load of a picking vehicle, order type, picking strategy, and whole planning of picking area. Previous researchers may use dynamic programming, zero-one integer programming, genetic algorithm, tabu search, and simulated annealing to solve the order-picking problem.

As for the method to improve the computational efficiency, Rana used the weight constraint of picking vehicle to zone picking area [1], thus the problem size could be

reduced. Goetschalckx and Ratliff proposed no skip rule to exclude impossible route to increase the efficiency [2]. Although these works can improve the computational efficiency of DCs, it is still not enough for obtaining order picking information in a short time. Therefore, the study developed a framework for order-picking problem which coordinates exact algorithm and heuristics and furthermore proposes a novel approach, which is called neighborhood selection criterion, to improve the computational efficiency. The algorithm avoids algorithms, such as branch-and-bound algorithm or tabu search [6] employed in the research, to select bad nodes before branching or moving a point. Consequently, the approach is useful to reduce the computational effort. Besides, in order to make the algorithm more robust and reasonable, the fuzzy inference system was employed.

The organization of the paper is as followed. Section 2 is the model of the order-picking problem. The methodology is presented in section 3, which is mainly introduces the neighborhood selection criterion the research proposed and the fuzzy inference system. Section 4 is the empirical result and discussion. Section 5 draws the conclusion and guides the future research.

2   Model of Order-Picking Problem

2.1   Assumption of Physical Environment

The research supposes that the width of the aisle should be taken into consideration, and each stock location has its own coordinates, including x, y and z. The I/O or original point O is at (0, 0). The shape of the warehouse is shown as in Fig. 1. There are totally 10 x 10 racks in 2-dimensional space and every rack has three levels. The width of the aisle is 2 meters, and the length of each location is 3, width is 1 and height is 1 meter respectively. As far as the capacity of picking vehicle is concerned, its weight limit is set to 60 kilograms. Besides, the order type is batch picking which collects all orders during a specific time period.
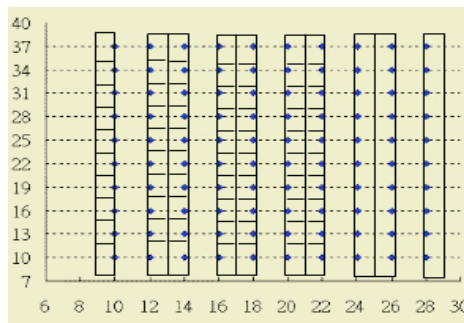
Fig. 1. The shape of the warehouse and the coordinates of each stock location

2.2   Mathematical Model of Order-Picking Problem

The distance between stock locations is a rectilinear distance, not an Euclidian one.

Suppose D_{ij} means the distance from location $i(x_i, y_i, z_i)$ to the location $j(x_j, y_j, z_j)$. A_i and A_j denote the aisle number of the two individual locations respectively. Thus, the rectilinear distance D_{ij} is:

$$D_{ij} = |x_i - x_j| + |y_i - y_j| + (z_i + z_j) \qquad if\ A_i = A_j\ , \text{otherwise}$$

$$D_{ij} = |x_i - x_j| + min\{|2Y - y_i - y_j|, |y_i + y_j|\} + (z_i + z_j) \tag{1}$$

Where Y is the total length of an aisle.

The mathematical model of the order-picking is as following:

$$\text{minimize} \quad Z = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} D_{ij} X_{ijk} \tag{2}$$

Subject to

$$\sum_{j=1}^{n} \sum_{k=1}^{n} X_{ijk} = 1,\ i = 0,1,2,\ldots, n \tag{3}$$

$$\sum_{i=1}^{n} \sum_{k=1}^{n} X_{ijk} = 1,\ j = 0,1,2,\ldots, n \tag{4}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} X_{ijk} = 1,\ k = 0,1,2,\ldots, n \tag{5}$$

$$\sum_{i=1, i \neq j}^{n} X_{ijk} = \sum_{r=1, r \neq j}^{n} X_{jr(k+1)} \tag{6}$$

$$\sum_{j}^{n} X_{0j0} + \sum_{j}^{n} X_{(n+1)j0} = 1 \tag{7}$$

Where X_{ijk} means the picking sequence is from location i to location j of the sequence k and its value is either 0 or 1.

3   Methodology

The integrated framework of the research adopts the branch-and-bound algorithm and tabu search. The presupposition is to use cluster algorithm to group items that can be picked by one vehicle [7]. The cluster will contain large size and small size items to be picked, among them the small one is solved by branch-and-bound algorithm and the large one is solved by tabu search. In order to improve computational efficiency, the research proposed a neighborhood selection criterion based on constraint propagation and the prune-and-search algorithm. The proposed method forbids bad node when searching a solution by branch-and-bound algorithm or tabu search. Besides, in order to make the algorithm more robust and reasonable, the fuzzy modeling was employed.

3.1 Neighborhood Selection Criterion

If the branch-and-bound algorithm is able to prune some data points in advance, it is definitely speed up the computation because it no longer has to enumerate impossible points. As for tabu search, it has to move all possible points by local search method. However, if we also set a constraint for the selection of movement, it won't calculate these moves and saves computational time. Consequently, the study attempts to define a neighborhood selection criterion based on constraint propagation and prune-and-search algorithm.

The research lets 65% points be allowed to move and others 35% will be pruned. These 65% values play the role as critical value for later use. The procedures of the neighborhood selection criterion for the branch-and-bound and tabu search are as following:

Step1: Calculate the symmetric distance matrix for each location, including the I/O or the starting point.

Step2: Use the prune-and-search algorithm to find out the distance at the middle position of each location to others in $O(n)$ time-complexity.

Step3: When expanding the possible nodes of branch-and-bound or doing the swap movement of tabu search, it is going to apply the criterion to validate the branching or movement. If it is branch-and-bound algorithm, it goes to step3.1; if the swap movement is backward swap, go to step 3.2; otherwise, go to step 3.3.

  Step3.1: Suppose the current node is i which connects to the point $i+1$. The method checks whether the node $i+1$ locates at the adjacent area to the node i. If yes, execute the branch; otherwise, prune the node we want to branch. Go to step 4.

  Step3.2: Suppose the new position of the moved point is at position i. Check whether the following point, $i+1$, is as the adjacent area to the moved point at position i. If yes, approve the movement; otherwise, discard the movement. Go to step 4.

  Step3.3: Suppose the new position of the moved point is at position i. Check whether the previous point, $i-1$, is as the adjacent area to the moved point at position i. If yes, approve the movement; otherwise, discard the movement. Go to step 4.

Step4: If there are still have nodes needed to branch or move, go to step3. Otherwise, exit the neighborhood selection criterion.

3.2 Fuzzy Modeling Approach

The rule of the neighborhood selection criterion is that the distance between the two nodes is no more than the critical value. However, it may not always robust to sets specific value as critical value. Therefore, the research proposed a fuzzy modeling approach of the distance definition of whether the point is far away. There are inputs, rules, and output in the fuzzy inference system. Figure 2 shows the relationship between the three components.

Input:

The inputs defined here are the rectilinear distance between two nodes and the distance variance. The membership function of inputs is in the pi-shape. Besides, there are three terms, named "Close", "Median", and "FarAway", in the input variable "Distance", and two terms, "Low" and "High", in the input variable "Variance" respectively. There are four parameters of a term to set the pi-shape membership function. Table 1 shows the parameters of the terms of input variable "Distance".
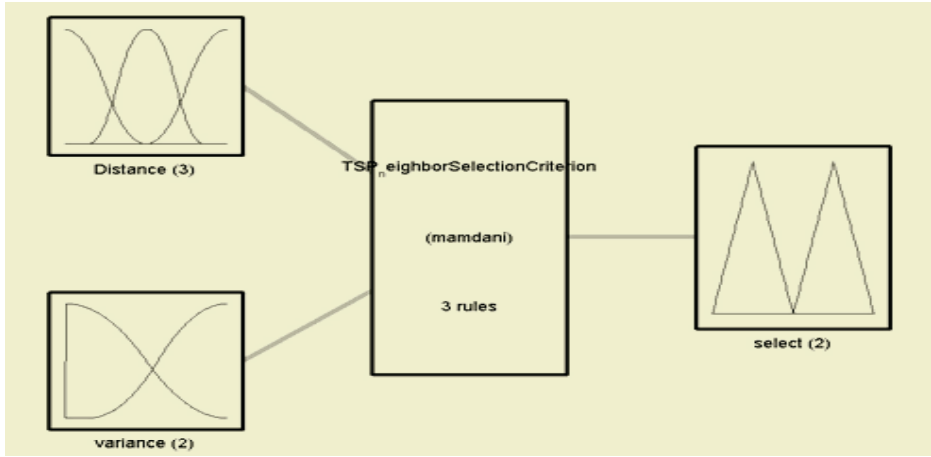
Fig. 2. The fuzzy modeling system

Table 1. The parameters of terms of input variable "Distance"

Term	Parameter of [a, b ,c ,d] of each term
Close	[lower_bound, lower_bound, lower_bound, middle_point];
Median	[middle_point/2, middle_point, middle_point , middle_point*1.5]
FarAway	[middle_point, middle_point, middle_point, upper_bound]

Rules:

The most important component of the fuzzy inference system is the fuzzy rules. It determines interaction between the input variables and the output variable. Three rules are defined in the research: (1) If the distance between i and j is close, then select it. (2) If the distance between i and j is far away, then don't select it. (3) If the distance between i and j is median and the variance is high, then select it. The implementation of the fuzzy model is by the mamdani method and sugeno method.

Output:

The output variable is to determine whether we should select the node. The boundary is between 0 ~ 1. Its attribute includes *selected* or *not selected*. Suppose there are two node i and j; if the value of output variable is less than 0.5, it is belong to the "not selection" term and $crticalValue_{ij} = 0$ which means the connection between i and j is forbidden. Otherwise, it accepts the connection and $crticalValue_{ij} = 1$. The membership function of the output variable is in triangular shape.

4 Empirical Result and Discussion

To verify the approach, the research conducts experiments to show the efficiency improving. The programs were written in Java and run on PIII 1.4 G (IBM X220 Dual CPU). Because the branch-and-bound algorithm is the kernel algorithm of ILOG JSolver, the experiment applied the software component to assist the research. Besides, the study also employees a tabu search framework, which is named *OpenTS* [8]. Furthermore, We utilized the fuzzy toolbox of Matlab and use Matlab to call the Java programs of branch-and-bound algorithm and tabu search we wrote rather than implemented them by Matlab language.

The order items are 10 which are randomly generated. Table 2 shows the coordinates of these items. The coordinate of I/O is at (0, 0). The order picker enters from the I/O into the warehouse. After picking the 10 items, the picker goes back to I/O. Although there were only 10 data points, if user wants to evaluate all possible combinations, there are 10! (3628800) possible solutions. The following sections present the implementation performance of branch-and-bound algorithm and tabu search respectively. The two sections show the effect on the computational effort and the solution quality. Then, the research explains the result in the discussion section.

Table 2. Location of the 10 order items

Number	1	2	3	4	5	6	7	8	9	10
X	26	26	14	28	14	22	20	24	10	26
Y	19	25	22	22	28	13	31	25	13	34
Z	1	0	2	0	1	0	0	2	2	1

4.1 The Performance of Branch-and-Bound Algorithm

The original number of nodes of branch-and-bound algorithm was 967670. The objective value of the picking distance is 222 meter. It needs 10.9 seconds to finish all the iteration. The picking sequence is {3, 5, 7, 1, 4, 2, 10, 8, 6, 9}.

Then, the experiment goes to evaluate the performance of the branch-and-bound algorithm when applying the neighborhood selection criterion. The number of choice points becomes 283817 and the computational time is 8.76 seconds. Most important of all, the objective value is still the same. By the above experimental result, we can find out the data nodes needed to be examined that was only 29% compared to original node size and the computational time was decreased 20% in the case. The reason why the saving time is inconsistent with the reduced nodes is that the method requires extra time when it adds constraints to check the neighborhood.

4.2 The Performance of Tabu Search

The section demonstrates the application of the neighborhood selection criterion into tabu search. If the tabu search doesn't apply the approach, it should move all possible moves. The original tabu moves was 45 here and it needed 0.047 seconds in average. However, when the tabu search employees the method, the possible moves of each

iteration reduced to 36.72 in average and the computational time was 0.015 seconds. The corresponding improved rate is 17.8% and 32%, respectively. The neighborhood selection criterion approach is able to eliminate unnecessary moves so that increase the computation efficiency

4.3   The Performance of the Fuzzy Approach on Tabu Search

Based on previous experiment result, the work continues the experiment which integrated the fuzzy modeling. The procedures of the examination is the same as Section 4.2. The time-saved is recorded by current cputime() in Matlab. After the implementation of the fuzzy inference system which is mandani type, the objective value is 222 which is still optimal. The number of nodes is 32.68317 in average. Besides, the computational time is 0.2030.

The research found the number of nodes of the fuzzy approach was less than previous one. However, the computational time was ten times higher than before. The reason might be caused by the poor performance of the Matlab operation.

5   Conclusion and Suggestions

The research proposed the neighborhood selection criterion to prune unnecessary nodes. Besides, the fuzzy inference system is also applied in the searching procedure. The empirical result proved the approach is useful on reducing the time complexity and improving the computational efficiency for both the exact algorithm and heuristic. The idea of the algorithm is based on the distance between two locations. When the distance is too far away, the approach rejects the branching or moving action when applying branch-and-bound algorithm or tabu search respectively. In the fuzzy approach, it was proved that it also reduced the number of moves of tabu search. Although the performance is not as good as we only use the tabu search in Java, we may revise the Matlab program into Java ones so that it can improve the performance. The research may be of interest for the researchers in the optimization of DCs' operations and the DCs' managers.

There are some suggestions to make the research be better in the future. First, there are only 10 test order data in the research, the later research can enlarge the problem size and check the solution quality. Second, although we use the fuzzy inference system to assist the criterion, there are some algorithms can be applied. For example, the ANFIS (Adaptive neural network of fuzzy inference system) that adopts the sugeno fuzzy inference system can replace the current FIS model.

References

1. Rana, K., Order-Picking in Narrow-Aisle Warehouse, International Journal of Physical Distribution & Logistics Management, 20(2) (1991) 9-15
2. Liu, C.M., Clustering technique for stock location and order-picking in a distribution center, Computers & Ops Res, 26 (1999) 989-1002

3. Wang, L.H., An integrated storage assignment and order picking sequencing model for distribution centers, master thesis, National Taiwan University of Science and Technology, Taiwan (1998)
4. Chen, L.I., Planning and Evaluation for Manual Picking Operations in Distribution Centers, master thesis, Institute of Industrial Engineering and Management, Yuan-Ze University (2001)
5. Chiang, Y. M. and S. H. Chen, Solving Order Picking Problem by Cluster Analysis and Tabu Search, 2002 National Conference of Management on Technology, HsinChu, Taiwan (2003)
6. Glover, F., Tabu Search II, ORSA Journal on Computing, 2(1) (1990) 4-32
7. Chen, S. H., master thesis, I-Shou University, Taiwan, (2004)
8. OpenTS: http://www-124.ibm.com/developerworks/oss/coin/OpenTS/index.html

An Efficient MDS Algorithm for the Analysis of Massive Document Collections

Yoshitatsu Matsuda and Kazunori Yamaguchi

Kazunori Yamaguchi Laboratory, Department of General Systems Studies,
Graduate School of Arts and Sciences, The University of Tokyo,
3-8-1, Komaba, Meguro-ku, Tokyo, 153-8902, Japan
{matsuda,yamaguch}@graco.c.u-tokyo.ac.jp
http://www.graco.u-tokyo.ac.jp/~matsuda

Abstract. In order to solve multidimensional scaling (MDS) efficiently, we proposed an algorithm, which apply stochastic gradient algorithm to minimizing well-known MDS criteria [1]. In this paper, the efficient MDS algorithm is applied to the text mining and compared with the SOM [2]. The results verified the validity of our algorithm in the analysis of a massive document collection. Our algorithm could find out some interesting structures from about 100000 articles in Usenet (NetNews).

1 Introduction

Multidimensional scaling (MDS), one of the methods in multivariate data analysis, is used widely for reducing the dimension of the data space [3]. It finds a configuration of objects in a low-dimensional space, which preserves the intrinsic disparities among all the objects in a high-dimensional space as "faithfully" as possible. Generally, MDS is carried out by minimizing a criterion (a cost function) evaluating the "faithfulness." The criterion is often called *stress* (e.g. the classical MDS stress (called C-STRESS in this paper) [4] and SSTRESS [5]). Since these criteria are based on pairwise disparities among all the objects, the computational cost of MDS increases quadratically according to the number of objects. Such cost is too high for large-scale problems. In our previous works [1] [6] [7], we proposed an efficient MDS algorithm named "global mapping analysis" (GMA for short) by applying stochastic gradient algorithm to the optimizations of various MDS stresses. In this paper, two variations of GMA (GMA on SSTRESS and GMA on C-STRESS proposed in [1]) are applied to the analysis of a massive document collections including about 100000 articles. The result shows that GMA can extract many interesting structures from such massive collections.

This paper is organized as follows. In Section 2, GMA on SSTRESS (Section 2.1) and that on C-STRESS (Section 2.2) are described in brief. In Section 3, GMA is applied to the analysis of a massive document collection (about 100000 articles posted to a Usenet newsgroup (NetNews)) and the result is compared with the SOM [2]. Lastly, Section 4 concludes this paper.

2   Global Mapping Analysis

2.1   GMA on SSTRESS

First, assume that there are N objects in the intrinsic M-dimensional space. Let x_{ip} be the p-th coordinate of the i-th object. Next, let an r-dimensional vector $\boldsymbol{Y}_i = (y_{ik})$ be the coordinates of the i-th object in a low-dimensional space $(r < M)$. Then, SSTRESS [5] is defined as

$$\text{SSTRESS} = \sum_{i,j} \left(d_{ij} - d_{ij}^* \right)^2, \tag{1}$$

where $d_{ij}^* = \sum_{p=1}^{M} \left(x_{ip} - x_{jp} \right)^2$ (the intrinsic disparity between the i-th and j-th objects) and $d_{ij} = \sum_{k=1}^{r} \left(y_{ik} - y_{jk} \right)^2$ (the square Euclidean distance in the r-dimensional space). In order to use stochastic gradient algorithm for minimizing Eq. (1), a probability distribution $P(\boldsymbol{X})$ is defined as a sum of N-dimensional δ functions:

$$P(\boldsymbol{X}) = \frac{1}{M} \sum_{p} \delta \left(x_{1p} - x_1, \cdots, x_{Np} - x_N \right), \tag{2}$$

where $\boldsymbol{X} \equiv (x_1, \cdots, x_N)$. Consequently, the following difference equation is derived (see [1] for the details of the derivations):

$$
\begin{aligned}
y_{ik} &(T+1) \\
= \quad & y_{ik}(T) \\
-c(T) &\sum_{l} \left(N y_{ik} y_{il}^2 - y_{il}^2 \alpha_k - 2 y_{ik} y_{il} \alpha_l + y_{ik} \beta_{ll} + 2 y_{il} \beta_{kl} - \gamma_{kl} \right) \\
+c(T) &\left(N x_i^2 y_{ik} - x_i^2 \alpha_k - 2 x_i y_{ik} \delta + y_{ik} \epsilon + 2 x_i \zeta_k - \eta_k \right),
\end{aligned} \tag{3}
$$

where (x_i) is given independently at each discrete time step T according to $P(\boldsymbol{X})$, $c(T)$ is the learning rate, and

$$\alpha_k = \sum_i y_{ik}, \qquad \beta_{kl} = \sum_i y_{ik} y_{il}, \qquad \gamma_{kl} = \sum_i y_{ik} y_{il}^2,$$

$$\delta = \sum_i x_i, \qquad \epsilon = \sum_i x_i^2, \qquad \zeta_k = \sum_i x_i y_{ik}, \qquad \eta_k = \sum_i x_i^2 y_{ik}. \tag{4}$$

If $\lim_{T \to \infty} c(T) = 0$ and $\sum_T c(T) = \infty$, SSTRESS is minimized for $(y_{ik}^{\text{conv}}) = \lim_{T \to \infty} (y_{ik}(T))$ (see [8]). Therefore, MDS on SSTRESS is solved incrementally by the difference equation Eq. (3). Obviously, the computational cost of Eq. (3) at each step is just the linear complexity.

2.2   GMA on C-STRESS

C-STRESS is defined as

$$\text{C-STRESS} = \sum_{i,j} \left(b_{ij} - b_{ij}^* \right)^2, \tag{5}$$

where

$$b_{ij} = \sum_{k} y_{ik} y_{jk}, b_{ij}^* = \sum_{p} \left(x_{ip} - \frac{\sum_i x_{ip}}{N} \right) \left(x_{jp} - \frac{\sum_i x_{ip}}{N} \right), \quad (6)$$

where x_{ip} and y_{ik} are defined in Section 2.1. (see [4] and [3]).

Similarly for SSTRESS, the following difference equation is derived:

$$y_{ik}(T+1) = y_{ik}(T) - c(T) \left(\sum_{l} y_{il} \beta_{kl} - x_i \zeta_k + \frac{x_i \alpha_k \delta}{N} + \frac{\zeta_k \delta}{N} - \frac{\alpha_k \delta^2}{N^2} \right), \quad (7)$$

where (x_i) is given by $P(\boldsymbol{X})$ and α_k, β_{kl}, and ζ_k are defined in Eq. (4). The incremental calculation of Eq. (7) solves the classical MDS. In addition, GMA on C-STRESS is essentially equivalent to Oja's symmetrical PCA network rule [8], and it is guaranteed to converge to the global optimum (see [1]).

2.3   Summary of GMA

1. Give $\boldsymbol{Y}$ randomly (for GMA on C-STRESS) or as the result of GMA on C-STRESS (for GMA on SSTRESS).
2. Repeat the following steps until convergence is reached:
   (a) Generate $\boldsymbol{X} = (x_i)$ by $P(\boldsymbol{X})$.
   (b) Calculate the parameters α-η by Eq. (4).
   (c) Update $\boldsymbol{Y}$ by Eq. (3) (on SSTRESS) or Eq. (7) (on C-STRESS).

3   Analysis of a Massive Document Collection

Here, a massive document collection is analyzed by GMA on C-STRESS, GMA on SSTRESS, and the SOM [2], and their results are compared.

3.1   Experimental Conditions

Raw data. 84817 articles posted to a Usenet newsgroup "`comp.os`" in Jan. 2001 - Feb. 2001 were used.

Extraction. Each article was decomposed into a set of words. Only the alphabetical words were extracted, and all the words were converted to their stems by WordNet [9]. In addition, the articles which consists of less than 5 words or more than 1000 words were removed. Consequently, 77280 distinct word stems and 84006 articles were extracted (namely, $N = 84006$). Almost all preprocessing methods were roughly based on [2].

Word selection. Two subsets of the 77280 words were used for the following experiments. The first subset (named "the rare words") is a set of words which occur in 10-20 articles. This set contains 6568 words. The second one (named "the ordinary words") is comprised of 1529 words which occur in 500-5000 articles. Each article is represented by a word vector whose i-th element is 1 if the i-th word is included in the article, and 0 otherwise. Consequently, 6568-dimensional word vector space was used for the rare words, and 1529-dimensional one for the ordinary words.

Random mapping. Each d_i was transformed into a 500-dimensional vector n_i by the random mapping method in the same way as in [2]. It is because quite high-dimensional (more than 500-) vectors are intractable for the SOM. The reduced document vector $n_i = (n_{ip})$ is given as $n_i = d_i R$, where R is a 6568×500 or 1529×500 random matrix. For each row of R, randomly-selected five components were set 1, and the rest were set 0. It is easy to show that the square Euclidean distance between n_i and n_j approximates that between d_i and d_j.

GMA on C-STRESS. (x_{ip}) was set (n_{ip}). r (the number of dimensions in the low-dimensional space) was set 5. $(y_{ik}(0))$ was set randomly. Then, Eq. (7) was calculated repeatedly with decreasing $c(T)$ until $(y_{ik}(T))$ converged. Lastly, five axes in the low-dimensional space were extracted by principal component analysis.

GMA on SSTRESS. In order to avoid local minima, the result of GMA on C-STRESS was given as the initial configuration $(y_{ik}(0))$. Then, Eq. (3) was calculated repeatedly in the same way.

SOM. Every $(x_{i1}, \cdots, x_{i500})$ was used as the training vectors. A rectangular map of 100×100 units was used. The reference vectors of the units were trained by the simple SOM algorithm of SOM_PAK [10]. The training lasted for 10000 steps during the first (ordering) phase. During the second (fine-tuning) phase, it lasted for 100000 steps.

3.2   Comparison of Results

Calculation time. All the calculations were carried out by a Pentium III 1GHz processor. The calculation time taken by the three methods are shown in Table 1. These results confirm in any case that GMA is faster than the original SOM algorithm. There have been many techniques to make the SOM faster [2], but GMA has an additional advantage. In these experiments, r in GMA ($= 5$) is much larger than r in the SOM ($= 2$). The computational cost of GMA increases just quadratically according to r, while that of the SOM increases exponentially.

Word vector space of rare words. The visualized results for the rare words are shown in the left side of Fig. 1. Though the result of GMA on C-STRESS (Fig. 1-(a)) spreads over the first axis and the borders of the clusters are not visible, it became clearer by GMA on SSTRESS (Fig. 1-(c)) and three distinct cluster were visible. These three clusters were found roughly by

Table 1. Calculation Time: This table shows the calculation time (minute) for the three methods. Because GMA on SSTRESS uses the results of GMA on C-STRESS as the initial configurations, the additional time (noted in the following parentheses) must be taken into account. This result verifies the efficiency of GMA.

	C-STRESS	SSTRESS	SOM
rare words	60	250 (+60)	470
ordinary words	50	110 (+50)	460

For Rare Words.   For Ordinary Words.
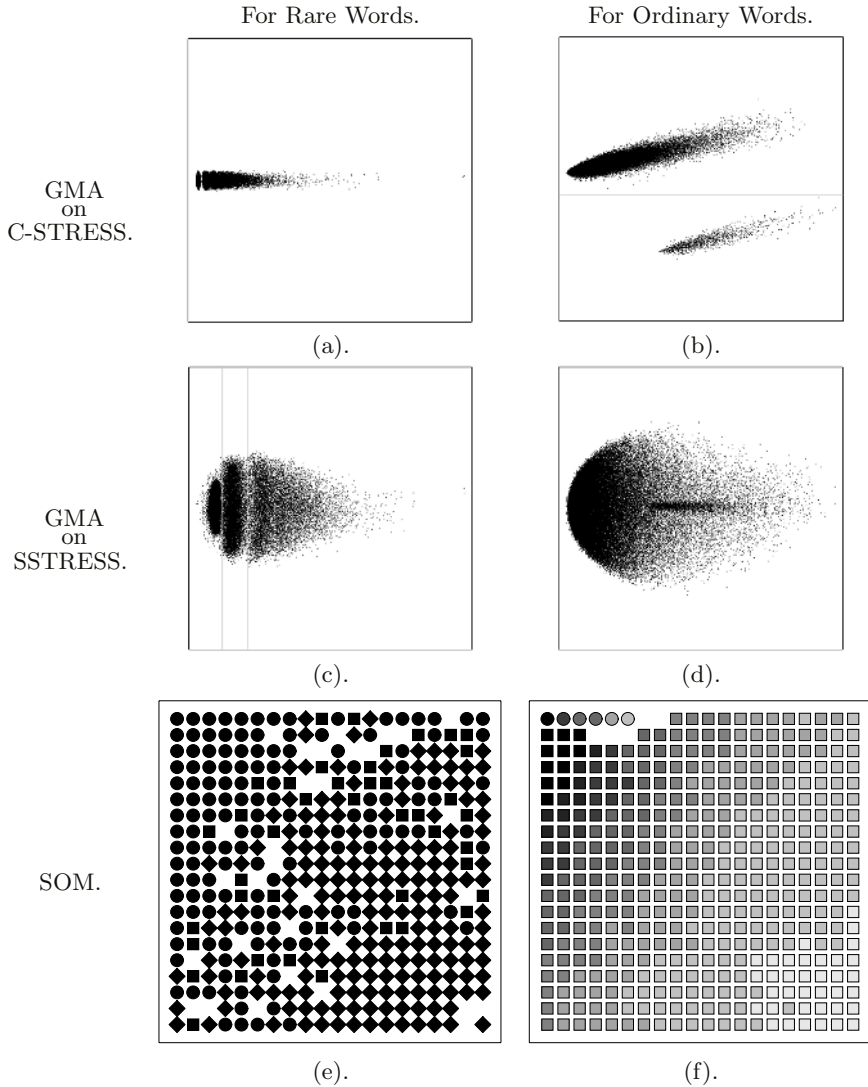
Fig. 1. Results of Analysis of Massive Document Collection: In (a)-(d) (the results of GMA), the first and second axes are along the x- and y- coordinates. An article is represented by a dot. In (c), The gray lines indicate the borders of the three clusters found by GMA on SSTRESS. In (e) and (f) (the results of SOM), the 100×100 SOM units were divided equally into 20×20 blocks. Each block (of 5×5 units) includes the articles nearest to any unit in the block. The cluster of each block was determined by a majority vote from its included articles. If there exists no article, the block is left blank. In (e), the left, middle, and right clusters in (c) correspond to the diamonds (◆), squares (■), and circles (●), respectively. In (f), the squares (■) and circles (●) correspond to the upper and lower clusters in (b). The darkness of each block represents the average number of words over the included articles.

the SOM also. The right lower, middle, and left upper regions in Fig. 1-(e) correspond to the left, middle, and right clusters in Fig. 1-(c), respectively. But, there is no visible border and the blocks of the middle cluster (■) were scattered. So, the results show that GMA can find the clusters more distinctly than the SOM. It is probably because of the problem with local minima inherent in the SOM.

Word vector space of ordinary words. The results for the ordinary words are shown in the right side of Fig. 1. GMA on C-STRESS (Fig. 1-(b)) found two (upper and lower) clusters quite clearly. In GMA on SSTRESS (Fig. 1-(d)), the upper and lower clusters (in Fig. 1-(b)) correspond to the left large semicircle and the horizontal tail in the middle, respectively. In addition, Fig. 2-(a) shows that the x-coordinates of Fig. 1-(b) and (d) (the first axes of the results) correspond to the number of words in an article. It is worth noting that the first axis is common to the two clusters. In other words, the number of words (the first axis) and the distinction between the two clusters (the second axis) were extracted independently. The SOM (Fig. 1-(f)) could distinguish the two clusters and could arrange the articles according to the number of words in *each* cluster. But, it failed to find a *common* axis. Even though the axis of the number of words probably existed in the intrinsic space, the SOM distorted it. Besides, Fig. 2-(b) shows that different weak clusters were found along the third axis of GMA on C-STRESS, while there are only two axes in the SOM.

| For Fig. 1-(b). | For Fig. 1-(d). | |
(a). The number of words and first axis.   (b). Other axes.
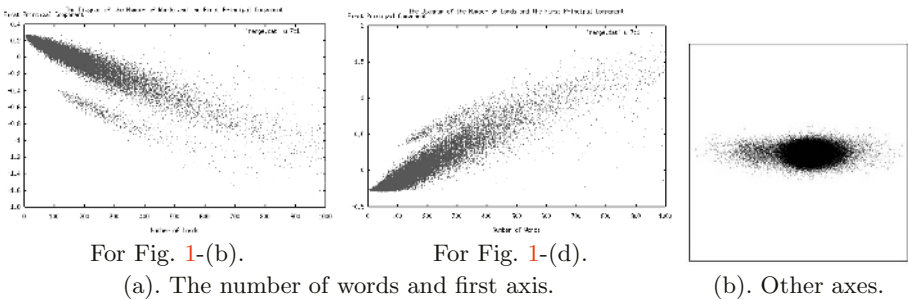
Fig. 2. Supplementary Results for Ordinary Words: (a). The number of words is in the x-coordinate. The first axis (the x-coordinate in Fig. 1-(b) (GMA on C-STRESS) or Fig. 1-(d) (GMA on SSTRESS)) is in the y-coordinate. Each dot corresponds to an article. (b). The third axis and the fourth one of GMA on C-STRESS are along the x- and y- coordinates.

Comprehensive discussion. The advantages of GMA are summarized as follows:

- GMA is more efficient than the simple SOM algorithm.
- GMA is applicable to larger r.
- GMA can detect clusters clearly without converging to a local minimum.
- GMA can find the axes in the intrinsic space without distortion.

These advantages confirm the validity of GMA.

4 Conclusion

In this paper, we applied GMA to the analysis of a massive document collection and found out some interesting structures efficiently. Because some "meanings" of the clusters of the Usenet articles found by GMA are not clear at present, we are making it clear by the further inspection of the articles in a way similar to the analysis in the SOM [2]. In addition, we are planning to apply GMA to more massive collections of documents on the Internet [11].

References

1. Matsuda, Y., Yamaguchi, K.: Global mapping analysis: stochastic gradient algorithm in SSTRESS and classical MDS stress. In: ICONIP2001 Proceedings, Shanghai, China (2001) 102–107
2. Kohonen, T., Kaski, S., Lagus, K., Salojarvi, J., Honkela, J., Paatero, V., Saarela, A.: Self-organization of a massive document collection. IEEE Transactions on Neural Networks **11** (2000) 574–585
3. Cox, T.F., Cox, M.A.A.: Multidimensional scaling. Chapman & Hall, London (1994)
4. Mardia, K.V.: Some properties of classical multidimensional scaling. Communications in Statistics: A, Theory and Method **A7** (1978) 1233–1241
5. Takane, Y., Young, F.W., Deleeuw, J.: Nonmetric individual-differences multidimensional-scaling - alternating least-squares method with optimal scaling features. Psychometrika **42** (1977) 7–67
6. Matsuda, Y., Yamaguchi, K.: Global mapping analysis: stochastic approximation for multidimensional scaling. International Journal of Neural Systems **11** (2001) 419–426
7. Matsuda, Y., Yamaguchi, K.: An efficient MDS-based topographic mapping algorithm. Neurocomputing **64** (2005) 285–299
8. Oja, E.: Principal components, minor components, and linear neural networks. Neural Networks **5** (1992) 927–935
9. Fellbaum, C., ed.: WordNet: an electronic lexical database. MIT Press, Cambridge, Massachusetts and London, England (1998)
10. Kohonen, T., Hynninen, J., Kangas, J., Laaksonen, J.: SOM_PAK: the self-organizing map program package. Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science (1996)
11. Baldi, P., Frasconi, P., Smyth, P.: Modeling the Internet and the Web: Probabilistic Methods and Algorithms. John Wiley & Sons, New York (2003)

(SE-33)Intelligent ID-Based Threshold System by an Encryption and Decryption from Bilinear Pairing

Young Whan Lee and Byung Mun Choi

Department of Computer and Information Security
Daejeon University, Daejeon, 300-716, Korea
{ywlee,bmchoi}@dju.ac.kr

Abstract. We suggest a new intelligent ID-based threshold secure system by an encryption and decryption protocol using (k, n)-threshold scheme which only k agents or more are engaged to reconstruct an encrypted data for a mobile intelligent system. Security issue protecting the network communication can be resolved by applying the Security Socket Layer. But the issue protecting the host against agent and the agent against malicious host can't be resolve with this SSL. We aim that a protocol has to allow hosts to send and receive safely an encrypted data from the agents, even if some of the agents are compromised under the multi-agent roaming system. Since our secure system have the verification, the hosts and agents can be protected from each other.

1 Introduction

The idea behind the (k, n)-threshold cryptosystem approach is to distribute secret information such as the secret key and computation among n parties in order to remove single point failure. The goal is to allow a subset of more than k players to jointly reconstruct a secret and perform the computation while preserving security even in the presence of an active adversary which can corrupt up to k parties. The secret information is distributed among n parties with the help of a trusted dealer or without it by running an interactive protocol among all parties.

1.1 Related Work and Discussion

Recently, several cryptographic protocols have been suggested many authors [4-15]. In particular, A. Boldyreva [2] proposed a threshold signature scheme that allows users to generated signature shares and to reconstruct the signature using bilinear pairing. B. Libert and J. J. Quisquater [10] suggested a pairing based (k, n) - threshold decryption scheme. This scheme is a threshold adaption of the Bonet-Franklin IBE scheme [3] where a fixed PKG plays the role of a trusted dealer. J. Baek and Y. Zheng [1] also presented a ID-based (k, n) - threshold decryption scheme. This provides the feature that an user who obtained a private

key from the PKG(Private Key Generate) can share the key decryption agents at will. After key generation the PKG can be closed. Also this protocol achieves chosen ciphertext security under BDH assumption in random oracle model.

1.2   Contribution

In this paper we present an ID-based threshold authentication protocol that allows users to send and receive encrypted data from the agents, even if some of the agents are compromised. We use a (k, n)-threshold scheme which only k agents or more are engaged to reconstruct a secret information. Our scheme is based on a bilinear pairing that could be built from Weil paring or Tate paring such that BDH problrm for this pairing is hard.

2   Preliminaries

In this section we shall summerize the properties we require of bilinear pairing and Gap Diffie-Hellman group much of details can be found [3]. Let G_1 be a prime order subgroup of an elliptic curve E over the field F_p for same large prime p where the order G_1 is q and G_2 a subgroup of $F_{p^2}^*$. Therefore, we view G_1 as an additive group and G_2 as a multiplicative group. A map $\hat{e} : G_1 \times G_1 \to G_2$ is said to be a bilinear pairing if $\hat{e}$ satisfies the following properties

1. Bilinearity : $\hat{e}(aP, bQ) = \hat{e}(P, Q)^{ab}$ for all $P, Q \in G_1$ and all $a, b \in Z_q^*$.
2. Non-degeneracy : there exists a $P \in G_1$ such that $\hat{e}(P, P) \neq 1$. Observe that since G_1 and G_2 are groups of prime order this implies that if P is a generator of G_1 then $\hat{e}(aP, bP)$ is a generator of G_2.
3. Computability : there exists an efficient algorithm to compute $\hat{e}(P, Q)$ for any $P, Q \in G_1$.

To construct the bilinear pairing, we can use the Weil pairing and Tate pairing [3]. Now we specify some version of Diffie-Hellman problems.

1. Given a generator P of a group G_1 and a 3-tuple (aP, bP, cP), the Decisional Diffie-Hellman (DDH) problem is to decide whether $c = ab$.
2. Given a generator P of a group G_1 and a 2-tuple (aP, bP), the Computational Diffie-Hellman (CDH) problem is to compute abP.
3. If G_1 is a group such that DDH problem can be solved in polynomial time but no probabilitic algorithm can solve CDH problem with non-negligible advantage with polynomial time, then we call G_1 a Gap Diffie-Hellman (GDH) group.
4. Given a generator P of a group G_1 and a 3-tuple (aP, bP, cP), the Bilinear Diffie-Hellman (BDH) problem for a bilinear pairing $\hat{e}$ is to compute $\hat{e}(P, P)^{abc}$.

The domains of bilinear pairings provide examples of GDH groups. The MOV reduction [11] provides a method to solve DDH in G_1 whereas there is no known

efficient algorithm for CDH in G_1. Note that an algorithm for CDH in G_1 and G_2 is sufficient for solving BDH for a map $\hat{e} : G_1 \times G_1 \to G_2$, but the converse is currently an open problem[3]. In this paper we assume that BDH problem is hard, which means there is no polynomial time algorithm to solve BDH problem with non-negligible probability. This assumption is called the BDH assumption.

3 Intelligent ID-Based Threshold Encryption and Decryption (IIDTED) Scheme

3.1 IIDTED Protocol Description

We let l be security parameter given the setup algorithm and let W be some BDH parameter generator.

System Setup: Given a security parameter l the algorithm works as follows:

1. Run W on input l to generate a prime $p > 2^l$ (2^l: bite-size of the prime number), and a bilinear pairing $\hat{e} : G_1 \times G_1 \to F_{p^2}^*$.
2. Choose two arbitrary generators P and $P_{pub} \in G_1$, where $P_{pub} = sP$ for some $s \in Z_q$ and the computing s given P and P_{pub} is infeasible.
3. Choose cryptographic hash functions $H : F_{p^2} \to \{0,1\}^n$ for some n, $G_1 : \{0,1\}^* \to G_1$ and $F : \{0,1\}^* \times G_1 \to Z_{q^*}$
4. The system parameters are

$$s - params =< p, \hat{e}, h, G, F, P, P_{pub} > .$$

5. Send $s - params$ and user's private key $sG(ID)$ where ID is user's identity.

Encryption Protocol: Denote n agents involving $s - params$ as $S_1, S_2, \cdots, S_n$ and the client playing the role of a dealer as Alice. Let ID_{Bob} be the Bob's identity and d_{Alice} the Alice's private key. We assume that Alice know Bob's public key ID_{Bob}. Alice then performs the following protocol at the client terminal:

Step1. The client Alice, as a dealer, distributes user's secret information to n agents:

(a) Select randomly $r \in Z_q^*$
(b) Choose a random polynomial $f(x)$ over Z_q of degree $k - 1$ such that

$$f(0) = a_0 = r \quad \text{and} \quad f(x) = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1}.$$

(c) Compute $P_i = f(i)P$
(d) Compute $g_{Bob} = \hat{e}(Q_{Bob}, P_{pub}) \in F_{p^2}$ where $Q_{Bob} = G(ID_{Bob})$
(e) Encrypt a message M by $C = M \oplus H(g_{Bob}^r)$.
(f) Compute $T_i = H(P_i)d_{Alice} + f(i)P_i$
(g) Compute $V = F(H(g_{Bob}^r), P)$
(h) Send enrollment parameters $e - params =< ID_{Bob}, P_i, T_i, C, V >$

Step2. After receiving all the information from Alice, S_i dose as follows:

(a) Verifies that
$$\hat{e}(T_i, P) = \hat{e}(H(P_i)Q_{Alice}, P_{pub})\hat{e}(P_i, P_i) \tag{1}$$

(b) If the above equation in (a) is verified to be false, response a complaint against Alice. Otherwise accept and store $e-params < ID_{Bob}, P_i, T_i, C, V >$ in a storage maintained an each S_i.

Step3. Alice discards all information, and completes the enrollment protocol. For the sake of conveniens, we assume that Alice has receive no complaint in Step2.

Decryption protocol: Denote k agents by $S = \{S_i | 1 \leq i \leq k\}$. To give authenticated retrieval and rescue the message, the client Bob performs the actions as follows:

Step1. Bob sends a request message to k agents:

(a) Select an uniformly distributed random number $x \in Z_q^*$, compute $X = xd_{Bob}$
(b) Send X and ID_{Bob} to each agent $S_i \in S$

Step2. On receiving the request, each agent S_i responds as follows:

(a) Retrieve $e - params =< ID_{Bob}, P_i, T_i, C, V >$ from the storage maintained securely on each S_i
(b) Compute $K_i = \hat{e}(X, P_i)$ and then reply $< K_i, C, V >$ to Bob.

Step3. Finally, Bob reconstructs secret information as follws:

(a) Compute $L_i = \prod_{j \in S, j \neq i} \frac{j}{j-i}$ for each ith agent
(b) Compute $K_i' = (K_i)^{L_i x^{-1}}$ and $K = \prod_{i \in S} K_i'$.
(c) If $V \neq F(H(K), P)$, abort. Otherwise compute $C \oplus H(K)$

4   Security Analysis

In this section we discuss with the security aspects of our proposed scheme IIDTED and give the complete security proof.

Theorem 1. *The protocol IIDTED from section 3 satisfies as follows;*
(1) Secrete values are uniformly distributed. Thus even an adversary with all the set of identity has no advantage in identifying any one of the participating senders and receivers over random guessing.
(2) All subsets of k shares provided by honest players define the same unique secrete value.

Proof. (1) The polynomial f with degree $k-1$ can be considered as a function chosen randomly from the collection of all polynomials over Z_q^* since a user Alice choose randomly a secrete values r and $a_1, a_2, \cdots, a_{k-1}$, during the execution of IIDTED. Thus random elements in F_{p^2}, g_{Bob}^r, $\bigcup_{i=1}^n \{P_i\}$ and $\bigcup_{i=1}^n \{T_i\}$ are uniformly distributed. Also a user Bob choose randomly a secrete value x and so a X is a random element in F_{p^2}. Thus secrete values in our scheme are uniformly distributed. This means that action of the adversary are independent of the secretes.

(2) As a result we can state that IIDTED can be resistant to corruption of even $k-1$ of $n \geq 2k-1$ servers. Even there exists an adversary who can corrupt at most $k-1$ servers among $n \geq 2k-1$, any subset of k servers constructs the secrete value K uniformly distributed in F_{p^2}. Since for all $a, b, \in Z_p^*$

$$\hat{e}(aP + bP, P) = \hat{e}(P,P)^{a+b} = \hat{e}(ap, P)\hat{e}(bP, P),$$

$$\hat{e}(T_i, P) = \hat{e}(H(P_i)d_{Alice} + f(i)P_i, P) = \hat{e}(H(P_i)Q_{Alice}, P_{pub})\hat{e}(P_i, P_i).$$

This means that all the honest players indeed hold the verification Eq.(1). Also they hold shares $\{P_i\}$ which contribute to reconstruct unique secret by combining with client's request message as in decryption step. For any S of k shares and extra value X from client's request message, the unique secret is computed as follows :

$$K = \prod_{i \in S} \hat{e}(X, P_i)^{L_i x^{-1}} = \prod_{i \in S} \hat{e}(x \cdot d_{Bob}, f(i)P)^{L_i x^{-1}}$$

$$= \prod_{i \in S} \hat{e}(f(i)Q_{Bob}, P_{pub})^{L_i} = \prod_{i \in S} \hat{e}(f(i) \prod_{j \in S, j \neq i} \frac{j}{j-i} Q_{Bob}, P_{pub})$$

$$= \hat{e}(\sum_{i \in S} f(i) \prod_{j \in S, j \neq i} \frac{j}{j-i} Q_{Bob}, P_{pub}) = \hat{e}(rQ_{Bob}, P_{pub}) = g_{Bob}^r,$$

where L_i is an appropriate Lagrange interpolation coefficient for the set S. Since the above holds for any set of k correct shares then K is uniquely defined, where the same extra value X which as said is derived from the Bob's private key.

Theorem 2. *Under BDH assumption, the protocol IIDTED is secure against the adversary A who is able to corrupt and to control k or more servers if he desires.*

Proof. We need to show that A can not reconstruct the secrete value K without knowing the user Bob's private key. In order to break the protocol, A tries to compute K' such that

$$K' = \prod_{i \in S} \hat{e}(X', P_i)^{L_i x'^{-1}}$$

with knowledge of system parameter params, any set S of t secret shares P_i for $i = 1, 2, \ldots, t (t \geq k)$ and client request messages X', where L_i is an appropriate Lagrange interpolation coefficient for the set S. Consider the three cases to compute K'.

First, A find an integer x such that $X' = x d_{Bob} = x s Q_{Bob}$. That is, A solve CDH problem. Since BDH problem is hard, CDH problem is also hard. Thus no adversary can compute X'.

Secondarily, A compute K' for given a 3-tuple (P, bP, xaP) such that $K' = \hat{e}(x d_{Bob}, P_{pub}) = \hat{e}(xaP, bP) = \hat{e}(P, P)^{xab} = g^{abx}$, where $d_{Bob} = aP$, $P_{pub} = bP$ and g is a generator of F_{p^2}. But it is impossible since BDH is hard.

Third, A guess Bob's private key d'_{Bob} and choose an element x' in Z_p^*. Now A compute

$$K' = \prod_{i \in S} \hat{e}(x' d'_{Bob}, P_i)^{L_i x'^{-1}} = \prod_{i \in S} \hat{e}(s' Q_{Bob}, f(i)P)^{L_i} = \hat{e}(r Q_{Bob}, s'P).$$

Since every secrets are uniformly distribute, A can not verify whether his guess is correct or not. Thus it is impossible to distinguish K' from K.

5   Comparison

With mainly compared to [1] and [10] our scheme have different features. J. Back and Y. Zheng [1] aimed at private key distribution. Thus they provided the feature that a user who obtained a private key from the PKG can share the key among ndecryption servers. B. Libert and J. J. Quisquater [10] proposed a pairing based (k, n) - threshold decryption scheme. This scheme provides the feature that the PKG plays the role of the trusted dealer and each distributed servers give the ciphertext to the recombiner who may be a designated player.

Our scheme is similar to that of [1] and [10], but with some different features: First, our scheme employs the concept of threshold scheme in a different way such as the dealer distribute secret shares to n agents with only receiver's identity and system parameters. And then the receiver reconstructs secret shares with his private key hardening protocol. That is, the receiver interact with servers to harden his private key into a strong secret without retrieving either his private key or the hardened result. Secondarily, servers and receivers in our scheme have parameters determining whether the message has been changed.

6   Conclusion

In this paper we present a new intelligent ID-based threshold encryption and decryption scheme using a bilinear pairing. When only k agents are involved in our protocol, the receiver can decrypt dealer's data with knowledge of his own hardening private key. Besides, even attacker succeeded in compromising k or more agents but without knowing the user's private key, she still cannot obtain any information about the user's credential. Since our secure system have the verification, the hosts and agents can be protected from each other. We have also presented simple solution to the security problems discovered in the scheme.

References

1. Baek, J., Zheng, Y.: Identity-Based Threshold Decryption. Cryptology ePrint Archive, Report 2003/164 available at http://eprint.iacr.org /2003/164, PKC 2004
2. Boldyreva, A.: Efficient Threshold Signature, Multisignature and Blind Signature Schemes Based on the Gap-Diffie-Hellman-Group Signature Scheme. PKC 2003, LNCS 2139, Springer-Verlag, (2003) 31–46
3. Boneh, D., Franklin, M.: Identity Based Encryption form the Weil Pairing. SIAM J. of Computing, Vol. 32, No. 3, Extended Abstract in Crypto 2001.(2003) 586–615
4. Boneh, D. , Lynn, B. , Shacham, H.: ; Short Signatures form the Weil Pairing. In Proceedings of Asiacrypt 2001
5. Boneh, D., Mironov, I. , Shoup, V.: A secure Signature Scheme form Bilinear Maps. CT-RSA-2003, 98–110
6. Chen, X., Zhang, F., Kim, K.: A New ID-Based Group Signature Scheme form Bilinear Pairings. Cryptology ePrint Archive, Report 2003/116, available at http://eprint.iacr.org /2003/116
7. Cocks C. : An Identity Based Encryption Scheme based on Quadratic Residues. In Cryptography and Coding, LNCS 2260, Springer-Verlag, (2001) 360–363
8. Galbraith, S., Harrison K., and Soldera, D.: Implementing the Tate Pairing. Algorithm Number Theory Symposium - ANTS V, LNCS 2369, Springer-Verlag(2002) 324–337
9. Hess, F.: Efficient Identity Based Signature Schemes Based on Pairings. SAC 2002, LNCS 2595, pp. 310-324, Springer-Verlag, (2002)
11. Libert, B. , Quisquater,J. J.: Efficient Revocation and Threshold Pairing Based Cryptosystems. PODC 2003, 163–171
11. Libert, B. , Quisquater, J. J.: New Identity Based Signcryption Schemes from Pairings. Cryptology ePrint Archive, Report 2003/023, available at http://eprint.iacr.org /2003/023
12. Paterson, K. G.: ID-Based Signatures from Pairings on Elliptic Curves Electron. Lett Vol. 38, No. 18, pp. 1025-1026, 2002. Also available at http://eprint.iacr.org /2002/004
13. Sakai, R. and Kasahara, M.: Cryptosystems Based on Pairing over Elliptic Curve SCIS 2003, 8c-1, Jan. 2003. Japan. Also available at http://eprint.iacr.org/2003/054
14. Sakai, R., Ohgihi, K. and Kasahara, M.: Cryptosystems Based on Pairing. SCIS 2000-c20, Okinawa, Japan, January 2000
15. Zhang, F., Safavi-Naini, R. and Susilo, W.: Efficient Verifiably Encrypted Signature and Partially Blind Signature from Bilinear Pairings. In Proceedings of Indocrypt 2003, LNCS, Springer-Verlag, (2003)

Development of an Intelligent Information Security Evaluation Indices System for an Enterprise Organization

Il Seok Ko[1], Geuk Lee[2], and Yun Ji Na[3]

[1] Dept. of e-Commerce, Chungbuk Provincial University of Science and Technology
40 Geumgu-Ri, Okchon-Eup, Okchon-Gun,Chungbuk, Korea
isko@ctech.ac.kr
[2] Dept. of Computer Engineering, Hannam University
133 Ojeong-Dong, Daedeok-Gu, Daejon, Korea
leegeuk@hannam.ac.kr
[3] Dept. of Internet Software, Honam University
Gwangsan-Gu, 59-1 Seobong-Dong, Gwangju, Korea
yjna@honam.ac.kr

Abstract. For effectively facing information security threats under rapidly changing information-oriented environment, we need intelligent measuring the information security level of the whole enterprise organization. Most of the evaluation systems have performed evaluation with an emphasis on information security products so far. However, evaluating information security level for an enterprise needs analysis of the whole enterprise organization, and a synthetic and systematic evaluation system based on it. This study has tried to grasp the information security level of the whole enterprise organization, and develop an evaluation system of information security level for suggesting a more developing direction of information security with an intelligent agent.

1 Introduction

The importance of information security has been increasing while the whole society has become rapidly information-oriented. Information security has got to be comprehensively considered all over the organization, not limited to an information system or its technology. Also, the higher the dependence of an enterprise on information processing, the higher the loss caused by the imperfect security countermeasure of an information system. In order to effectively achieve the object of information security of an organization, there should be criteria or evaluation models showing a direction to exactly evaluate the information security level of the whole organization, and improve it. Also, for this, there should be evaluation indices capable of evaluating and improving information security levels by sections.

The typical information evaluation criteria are ITSEC, TCSEC, BS7799, etc.. ITSEC, TCSEC, CC, etc. are evaluation criteria for evaluation of information security systems; however, in fact, they were developed for technological evaluation of the security function of information security products, and so have their limits. Also, BS7799 is authentication of a management system, not authentication of a product; therefore, it is hard to evaluate the function of a system and the whole enterprise.

The security of an enterprise should be based on analysis of the whole enterprise organization. The whole security of an enterprise organization can't be achieved only by combination of individual security products, and the information security of an organization can be effectively maintained only when the whole security & management system works well by proper union of security products/system and managerial security countermeasures. Also, the whole security level of an organization should be evaluated through proper embodiment and operation of security policy derived from operation environment and security requirements peculiar to the organization. In order to overcome the limits of the existing researches, the study has subdivided information security elements into 5 levels – planning, environment, support, technology, and management, and developed indices based on them, and finally, made the information security level of the whole enterprise organization possible to be measured with an intelligent agent. Applying the results of the study makes possible synthetic and systematic evaluation of the information security level of an enterprise.

2   Related Works

The information security evaluation in the existing research contents could be divided by two approaches. The first is an evaluation system for the security function and performance of a system or a product like TCSEC, ITSEC, CC, etc.. The existing evaluation criteria have some limits as they can't evaluate various products required in private fields because of using evaluation criteria by products. The second is an evaluation system for a managerial side like BS7799.

■   **TCSEC (Trusted Computer System Evaluation Criteria)**
In 1983, the US had NCSC(National Computer Security Center) produce the draft of TCSEC called "Orange Book", a evaluation criterion for a safe computer system, and 2 years later (in 1985), it was adopted as a US Department of Defense standard(DoD STD 5200.28). TCSEC divides a operation system into 6 classes (C1, C2, B1, B2, B3, A1) for effectively evaluating the security of a computer system and spreading to each organ a computer system whose safety and reliance are proven. TCSEC emphasizes, of security elements, especially secrecy. Looking into TCSEC features, it has many various data as the first evaluation criterion in the world, and some manuals applicable to several cases. And it accumulates the related experience through trial and error. But it is hard to apply to a private enterprise valuing availability and integrity as it uses only the fixed criteria without dividing the side of function and the side of guarantee and was revised for secrecy alone of security elements.

■   **ITSEC (Information Technology Security Criteria)**
In May, 1991, some European countries like France, Germany, Britain, and the Netherlands announced the draft of ITSEC, their common security evaluation criterion of IT system by harmoniously unifying and adjusting each country's security standard. These countries drew up ITSEC for avoiding trade barriers and using it as a guide on basic standard proposals and exams. And in case of multinational development and heterogeneous system, that's also for minimizing double efforts

made during evaluation. ITSEC is divided basically by security function criteria and guarantee requirement criteria. ITSEC, as the first international unification criterion in the world, has function and guarantee divided, maintains fluidity by generally defining criteria, and mostly includes the contents of TCSEC. However, general description of criteria is hard to understand, and no subdivision by security level can make it include subjective views in time of evaluation.

■   **BS7799**

Supervised by the British Board of Trade, together with major enterprises like BT, HSBC, Marks and Spencer, Shell International, Unilever, etc. under the title, 'A Code of Practice for Information security management,' BS7799 was developed as a universal document that managers, who should embody and maintain the information security of their organization, can refer to and devised as the basis of an organizational security standard. BS7799 was first made in 1995, and revised in 1999; is used in Britain and other countries such as Australia, Brazil, the Netherlands, New Zealand, Norway, Finland, India, etc.; in October, 1999, was suggested as ISO standard into ISO/IEC 17799-1.
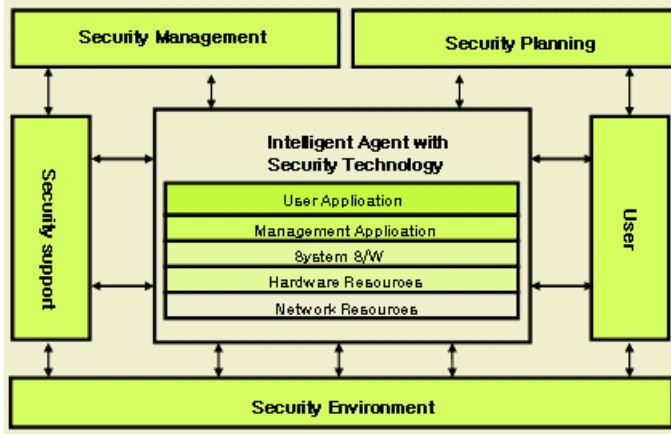
Fig. 1. Intelligent Information Security System on the organization viewpoint

3   An Information Security Index System

As shown in Chapter2, the existing evaluation systems have difficulty grasping the information security level of the whole enterprise synthetically and effectively. The study has developed a system that can evaluate the information security level of the whole enterprise from the whole viewpoint of an enterprise organization.

Fig.1 is an index system of information security shown from the whole viewpoint of an organization with an intelligent agent. This analyzes the whole vulnerable points of information security from five viewpoints like planning, environment, supporting, technology, and management level; then, evaluates the level of information security according to the analyzed results. Each index system is based on the essential success elements (CSFs) of information security planning level, and indices used in BS7799.

3.1   Information Security Planning Level

The index items of information security level are composed of security policy and security plan. Security policy evaluates the general items related to information security policy. The index items of information security policy evaluate whether to establish information security policy, whether to examine and evaluate them, and the documented part of information security policy. Also, it evaluates establishing action guides or not, establishing accident-settling procedure or not, stating exceptions or not, managing information and assets or not, drawing up the written oath of manpower related to information security or not, observing information security policy or not, including items for evaluating by CSFs the essential success elements of information security environment level. The indices of security plan index items also evaluate information security plan and information security investment, including items for evaluating by CSFs the essential success elements of information security environment level.

Table 1. Index System of Information Planning Level

Information security planning level	
Information security policy	Security plan
- Establishing information security policy or not / - Verifying information security policy / - Evaluating information security policy / - Reflecting the importance of information security / - Stating information security - Verifying, and evaluating information security policy documents - Checking technological observance / - Observing security policy - Establishing action guide / - Establishing accident-settling procedure / - Stating exceptions / - Managing information and assets / - Drawing up the written oath of related manpower / - Observing information security policy - Security policy documents	- Applied-technology development investment amount - Basic-technology development investment amount - Information security investment expense - Information security plan investment amount - Drawing up a security plan

Table 2. A Information Security Environment Level Index System

Information security environment level		
Equipment security	Personnel security	
	Organizational security	Human security
- Admission-ticket management - Automatic locking device / - Facilities security / - Network management - Equipment security / - Tangible assets - Intangible assets / - Material availability security - Equipment security	- Operating an information security committee - The existence of business alternation policy - Organization management - Organizational security - Security organization - Selecting those in charge of equipment, and assigning responsibility	- Duty definition and resource allocation security - Coping with security accidents and errors - Personnel management - Human security - Licensor management

3.2   Information Security Environment Level

The index items of information security environment level are composed of equipment security and personnel security as in table 2. Equipment security evaluates the items of information security facilities and basic equipment environment as in equipment control and facilities security. And it evaluates the present situation of basic environment such as admission-ticket management, automatic locking device,

facilities security, network management, internet connection specification, information security system, etc.. Personnel security is composed of organizational security and human security: organizational security evaluates observing or not detailed information security procedure like selection of those in charge of equipment within an organization and allotment of responsibility; personnel security evaluates individual-centered information security items such as a procedure settling security accidents, reporting vulnerability, managing licensors, etc..

Table 3. Information Security Support Level Index System

Information security support level		
Support organization	Support activity	
Support activity	Emergency countermeasure	Education training
- Development and support process security - Plant/support equipment - Development and support process security - Appointing those in charge of information security - Stating the roles of those in charge of information security - Unification/adjustment of information security activity - Security-related license level	- Establishing a emergency-settling plan or not - Agreements against emergency - Obstacle restoration - Settling preservation accidents and errors - Backup and restoration - Settling preservation accidents and function obstacles	- Education operation - Education organization - Education content - The selection method of educatees - The yearly mean number of education days - User education training - Education/training - User education/training

3.3 Information Security Support Level

The index items of information security support level are composed of support organization and support activity as in Table 3. Organization management evaluates development and support process security, appointment of those in charge of information security and their roles, security-related license level, etc; organization operation evaluates its whole matters like information activity unification/adjustment, etc.; outsourcing evaluates its own rate and sphere. Security support activity is composed of emergency countermeasures, information security education/training; emergency countermeasures evaluate the existence of emergency countermeasures and the emergency-settling ability of an enterprise. Last, information security education/training evaluates the selection method of educatees and information security education/training scope.

Table 4. Information Security Technology Index System

Information security technology level	
Access control operation	System function
- Access authority control / - Access control function - N/W access control / - Outsider access security - Operation system access control / - User access management - Application access control / - Network access control - Physical access security / - User approach management - User responsibility / - Network access control / - Operation system access control / - Applied-system access control / - System access and use supervision	- Account and password management - Applied-system security - Software security / - Authentication technique - Obstacle restoration / - Virus prevention - Encryption function / - Key management - The backup management of important files

3.4   Information Security Technology Level

The index items of information security technology level are divided into access control operation and system function as in table 4. The index items of access control operation are composed of indices evaluating the whole information security technology level. The index items of system function level are composed of indices evaluating the functional side (necessary to access control) and general items (on the whole account management and password).

3.5   Information Security Management Level

The indices system of information security management level, as in Table 5, is recomposed of management process requirements, documentation requirements, and BS7799 on the index items of managerial information security management control.

Table 5. Information Security Management Level Index System

Index items	Indices	Index items	Indices
Information security policies	Approval and announcement of policy The system of policy Maintenance and management of policy an organizational system Responsibility and roles	Human security	Establishing a education and training program Enforcement and evaluation Responsibility assignation and stipulation Managing a qualification test and those in charge of major duties Secret-keeping
Transaction security	A written exchange agreement Electronic-transaction security management Electronic mail The security management of open server User official announcement	Operation management	Operation procedure and responsibility System operation Network operation Media and document management Virulent software control
Access control	Cipher policy Cipher use Key management Access control policy User access management Access control scope		Establishing a business continuity management system Establishing and embodying a business continuity plan Planning, testing, maintaining and managing business continuity Mobile computing and remote working Information-asset investigation and responsibility assignation Information-asset classification and treatment
		Outsider security	Contract and service level agreement security management Outsider security practice management
Security-accident management	Countermeasure plan and system Countermeasure and restoration Post management	Physical security	Physical security measures Data- center security Equipment security Office security Analysis and design security management Embodiment and performance security management Change management
verification monitoring, and inspection	Observing and verifying legal requirements Observing and verifying information security policy Monitoring Security Inspection		

3.6   Comparison with the Existing Methods

Table 6 has compared the proposed method with the existing evaluation systems. TCSEC emphasizes secrecy, of information security elements, and is hard to apply to an enterprise organization. ITSEC evaluates all information security products, with a single criterion based, and performs evaluation of products as a security guarantee part. Also, these two have evaluation objects limited to products and a system, and have common features of emphasizing each country's traits and functional elements. BS7799 is authentication of a management system; therefore, it is hard to evaluate the function of a system and the whole enterprise. Each of the existing three evaluation systems is possible to use complementarily. The system proposed by the study has

subdivided each evaluation element into 5 spheres through complementation and extension of these systems; accordingly, made possible the information security level evaluation of the whole enterprise organization.

Table 6. Comparison of the evaluation systems.

Items	TCSEC	ITSEC	BS7799	Proposed method
Features	- Emphasizing secrecy, of information security elements. - Hard to apply to an enterprise organization - Limiting the object of evaluation to products and a system - Emphasizing each country's traits and functional elements	- Evaluating all information security products with a single criterion based - Performing evaluation of a product as a security guarantee part - Limiting the object of evaluation to products and a system - Emphasizing each country's traits and functional elements	- Authentication of a management system - No authentication of a product - Not enough to evaluate system function - Hard to evaluate the whole information security of an enterprise	- Possible to evaluate the information security level of the whole enterprise organization - Possible to evaluate a management system - Possible to evaluate products and a system - Subdividing each evaluation element into 5 levels

4   Conclusion

The study has divided information security elements of the whole enterprise into 5 levels to overcome the limits of the existing studies, developed indices system that is based on them with an intelligent agent, and so, made the information security level of the whole enterprise organization possible to be measured. Applying the results of the study has made the information security level of an enterprise possible to be evaluated synthetically and systematically. Also, applying the evaluation methods of information security level can have evaluation progress objectively. Establishing evaluation items considering features by industry and complementing the methods could maximize the effect.

References

1. PWC (1999), secure, defend, and transform: the complete e-business legal strategy, PWC
2. PWC (1999), security basics: a whitepaper, PWC
3. Roseann Day, John Daly, and Christian A. Christiansen (1999), eSecurity the essential eBusiness enabler, IDC
4. B.B Jenkins(1998), security risk analysis and management, countermeasures, Inc
5. Common Criteria Project (1998), common criteria for information technology security evaluation, common criteria
6. B. Guptill, C. Price (1996), a security framework for enterprise using the internet, Gartner Group
7. NIST (1995), an introduction to computer security : the NIST handbook, NIST(national institute of standards and technology)
8. Robin Moses (1995), corporate risk analysis and management strategies, European convention on security and detection, conference publication No. 408
9. Lynette Barnard et al. (1998), The evaluation and certification of information security against BS7799, Information Management & Computer Security 6/2, pp.72-77.
10. Rossouw von solms(1998), Information Security Management (3) : the code of practice for Information Security Management (BS7799), Information Management & Computer Security 6/5, pp.224-225.

A Study on the Centralized Database of the Multi-agents Based Integrated Security Management System for Managing Heterogeneous Firewalls

Dong-Young Lee

Dept. of Information and Communication, Myong-ji College
356-1 Hongeun3-Dong, Seodaemun-Gu, Seoul 120-776, Korea
dylee@mail.mjc.ac.kr

Abstract. In this paper, we present the merits of centralized approach for managing heterogeneous firewalls using multi-agents and implement the prototype of the central policy database that is a component of the PB-ISMSF (Policy-based Integrated Security Management System for Firewalls). Multi-agents initiate with SNMP security MIB, monitors the status of firewalls, and executes control requests from the PB-ISMSF engine. And also, we classified the policy conflicts, the condition of policy conflicts and briefly described the resolution methods for specified policy conflicts.

1 Introduction

To protect networks from these attacks, many vendors have developed various security systems such as firewalls, Intrusion Detection Systems (IDSs), and Virtual Private Networks (VPNs). However, managing these systems individually requires too much work and high cost. Thus, integrated security management and the establishment of a consistent security policy for various security products have become more important. The PB-ISMSF (Policy-based Integrated Security Management System for Firewalls) supports monitors and controls heterogeneous firewalls through central policies. In general, firewalls [1-3] are placed between an organization and the outside world, but a large organization may need internal firewalls as well to isolate security domains. A security domain is a set of machines under common administrative control, with a common security policy and security level. Integrated security management includes the functions of transparent management of security products, easy ways of defining security policies and the simple expansion of managed ranges. The PB-ISMSF consists of three separable parts - client, integration engine, and intelligent agent [4-5]. The PB-ISMSF client displays a conceptual management view of security services to security managers and sends monitoring or control requests to the engine using TCP/IP. The PB-ISMSF engine receives and processes requests from clients, and stores status information of security products from agents. It also generates SNMP [6-7] messages and dispatches them to the appropriate agents. And, the PB-ISMSF agents collect and directly control data form security products upon the requests from the engine. They also provide the engine with SNMP security MIB. Fig. 1 shows the detailed architecture of the PB-ISMSF.
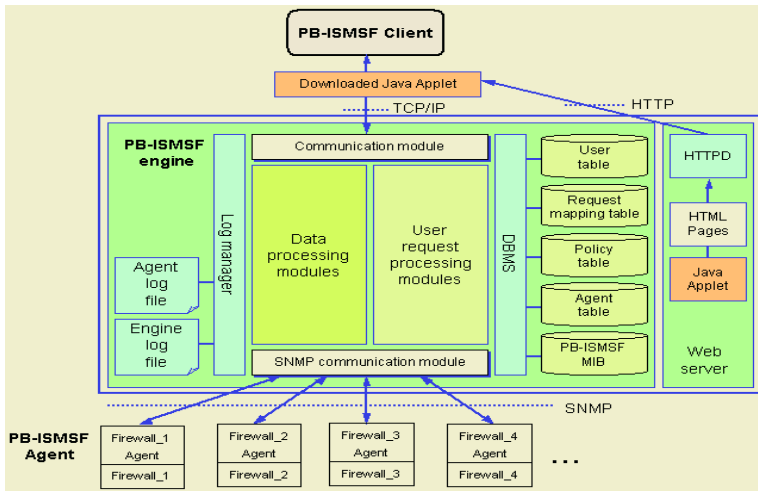
Fig. 1. Conceptual architecture of PB-ISMSF

The remaining paper is organized as follows. Section 2 describe the data structure of central policy database of PB-ISMSF. The classification of policy conflicts and the resolution methods for specified policy conflicts are described in Section 3. Finally, Section 4 discusses conclusion and future works.

2   Central Policy Database of PB-ISMSF

2.1   User Table

The user table is used for authenticating users and deciding appropriate permission to managing security systems. Users are classified into three levels and the user's scope of permission is differentiated according to these levels. User(security managers) are classified as network manager(NM), general security manager(GSM), and top-level security manager (TSM) based on the level of security management authority. Table 1 shows data structure of the user table of the PB-ISMSF.

Table 1. Data structure of the user table

Field	Data type	Description
user_id	INT	User identifier
Name	VARCHAR(20)	Login ID of user
Password	CHAR(10)	Login password
Level	ENUM	User level(NM, GSM, TSM)
network	CHAR(15)	Network address this user responsible for
description	TINYTEXT	Description for user

2.2   Agent Table

The agent table holds the information about firewall agents which directly control and manage the firewall. This table consists of the agent identification number (agent_id), the agent name, the type of managed firewall, the external address of the firewall, the

internal address of the firewall, and the SNMP community for the agent. Table 2 shows data structure of the agent table of PB-ISMSF.

Table 2. Data structure of the agent table

Field	Data type	Description
agent_id	INT	Agent identifier
name	VARCHAR(20)	Agent name
type	ENUM	Type of managed firewall
ext_addr	CHAR(15)	External network address of firewall
int_addr	CHAR(15)	Internal network address of firewall
community	VARCHAR(20)	SNMP community of agent

2.3   Policy Table

The policy table is used for managing whole policies of all the firewalls installed over an organization's network. Namely, it holds all firewall policies for central management. And, the policy table contains information related to the firewall policy like policy type, source address, source port number, destination address, destination port number, protocol type, and so forth. Table 3 shows data structure of the agent table of PB-ISMSF.

Table 3. Data structure of the policy table

Field	Data type	Description	OID
Index	INT	Policy index number	policyIndex
Rule	ENUM	Access control rule(permit, deny)	PolicyAccessCtrlRule
State	ENUM	Current state of policy	PolicyState
Src_addr	CHAR(15)	Source IP address	PolicySrcAddr
Src_port	CHAR(15)	Source port number	PolicySrcPort
Dst_addr	CHAR(15)	Destination IP address	PolicyDstAddr
Dst_port	CHAR(15)	Destination IP port number	PolicyDstPort
Protocol	ENUM	Protocol type	PolicyProtocol
Service	CHAR(20)	Service name	PolicyService
S_time	DATETIME	Start time for applying policy	PolicyStartTime
E_time	DATETIME	End time for applying policy	PolicyEndTime
Week_day	ENUM	Day of the week for applying policy	PolicyWeekDay
Notice	ENUM	Notification type(E-mail, Alarm, Log)	PolicyNotificationType
C_time	DATETIME	Time when policy is created	PolicyCreatedTime
M_time	DATETIME	Time when policy is modified	PolicyModifiedTime
Agent_id	INT	ID of agent related to this policy	PolicyAgentId
User_id	INT	ID of user who created this policy	PolicyUserId
comment	TINTEXT	Description of this policy	policyDescription

2.4   Procedure for the Policy-Request Processing

The PB-ISMSF engine, if it receives a policy addition request from a web client, processes the request in the procedure described below. The PB-ISMSF engine man-

ages the session information related to the corresponding web client from the user's login to logout (or closing the web browser).

The PB-ISMSF engine

① Extracts the host/network address from the policy addition request message.

② Checks whether the current user has proper permission to manage the policy for corresponding host/network

③ Selects the agent responsible for the access control policy for the acquired host/network address.

④ Checks whether the same policy exists or not in the selected firewall agent.

⑤ Extracts other values from the policy addition request message (Source/destination port number, protocol, service name, and so on.)

⑥ Builds the SNMP SetRequest message by pairing each value with appropriate OID.

⑦ Saves the policy to the database if it receives a successful response message from the agent. Fig. 2 describes the procedure for request processing.
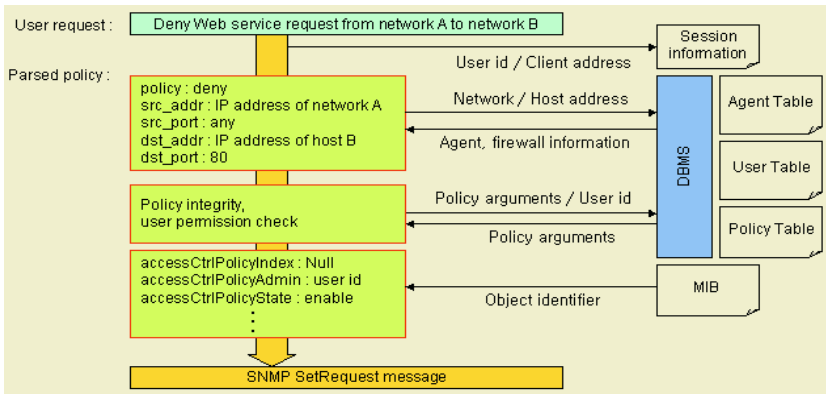
Fig. 2 Procedure for request processing

3   Policy Conflicts: Classification, Detection and Resolution

Whenever several policies are applied to the same target there is potential conflict between them. A preliminary work on conflict classification is reported in[8-10] where different types of conflict are distinguished according to the overlaps among the subject, action and target scopes. The policy of PB-ISMSF P(x) is defined by the existing policy(old) and the newly requested policy(new). A policy P(x) consists of policy manager M(x), Target T(x) and the action of policy A(x). That is, P(new) and (old) is defined as P(new) = {M(new), T(new), A(new) }, P(old) = {M(old), T(old), A(old)}

Condition 1. Equivalence; P(old) = P(new)
An equivalent policy conflict occurs when two policies have the same values of policy manager M(x), Target T(x) and the action of policy A(x). This can be resolved by

the user levels – network manager(NM), general security manager(GSM) and top-level security manager(TSM).

Condition 2. Contradiction; P(old) ↔ P(new)

A contradictable policy conflict occurs when both positive and negative policies of the same kind (i.e, permit policies or deny policies) exist. This conflict is very serious because there are no means of deciding whether the action is to be permitted or denied. A contradictable policy conflict can be resolved by user level and the priority of policy. The classification of contradiction conflict is shown in Table 5. Fig.3 shows the example of contradictable policy conflict. And, Fig. 4 shows the state diagram of detecting the contradictable policy conflict and resolving process.

(a) M(old) = M(new), T(old) = T(new) and contradictable relationship between A(new) = Negative_Action and A(old) = Positive_ Action
(b) M(old) ≠ M(new), T(old) = T(new) and contradictable relationship between A(new) = Negative_Action and A(old) = Positive_ Action
(c) M(old) ≠ M(new), T(old) = T(new) and contradictable relationship between A(new) = Negative_Action and A(old) = Positive_ Action at the same time.
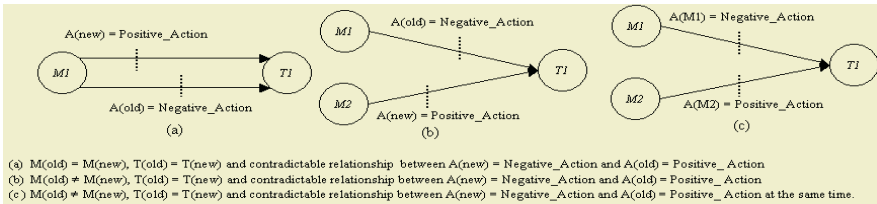
Fig. 3. Examples of the contradictable policy conflict
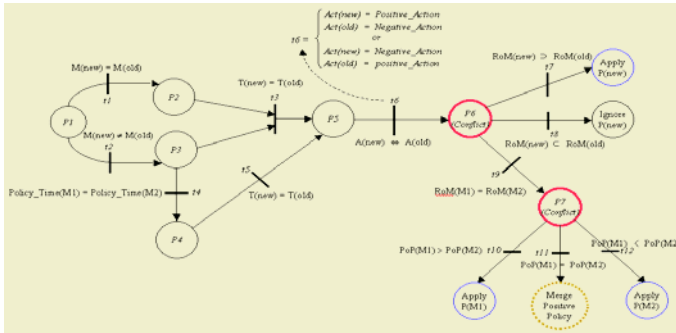
Fig. 4. State diagram for detecting & resolving for the contradictable policy conflict.

Table 4. Classification of contradiction conflicts

M(x); Manager	T(x); Target	A(x); Action
M(old) = M(new)	T(old) = T(new)	A(new) = Positive_Action A(old) = Negative_Action
		A(old) = Negative_Action A(old) = Positive_Action
M(old) ≠ M(new)		A(new) = Positive_Action A(old) = Negative_Action
		A(new) = Negative_Action A(old) = Positive_Action

Condition 3. Inclusion; {P(old)⊃P(new) ∨ P(old)⊂P(new)} ∧ {p(old) ↔ P(new)}
Examples of inclusive policy conflict are shown in Figure 7. This policy conflict occurs when the inclusive relationship with contradictable relationship between existing policy P(old) and newly requested policy P(new) exist. This problem can also be resolved by user level and the priority of policy. The classification of inclusive policy conflict with contradictable relationship is shown in Table 6 and Fig.5 shows the example of inclusive relationship policy conflict. And, Fig. 6 shows the state diagram of detecting the inclusive policy conflict and resolving process.

Table 5. Classification of inclusive policy conflicts

M(x); Manager	T(x); Target	A(x); Action
M(old) = M(new)	T(new) ⊂ T(old)	A(new) = Positive_Action A(old) = Negative_Action
		A(new) = Negative_Action A(old) = Positive_Action
	T(new) ⊃ T(old)	A(new) = Positive_Action A(old) = Negative_Action
		A(new) = Negative_Action A(old) = Positive_Action
M(old) ≠ M(new)	T(new) ⊂ T(old)	A(new) = Positive_Action A(old) = Negative_Action
		A(new) = Negative_Action A(old) = Positive_Action
	T(new) ⊃ T(old)	A(new) = Positive_Action A(old) = Netative_Action
		A(new) = Negative_Action A(old) = Positive_Action

(a) M(old) = M(new), T(old) ⊂ T(new) and contradictable relationship between A(new) = Negative_Action and A(old) = Positive_ Action
(b) M(old) ≠ M(new), T(old) ⊂ T(new) and contradictable relationship between A(new) = Negative_Action and A(old) = Positive_ Action
(c) M(old) ≠ M(new), T(old) ⊂ T(new) and contradictable relationship between A(new) = Negative_Action and A(old) = Positive_ Action at the same time.

Fig. 5. Examples of inclusive policy conflicts

4   Conclusion and Future Works

In this paper, we described the architecture and detailed features of the centralized database of the multi-agents based integrated security management system for managing heterogeneous firewalls in distributed network environments and illustrated the tree-tired system composed of client, engine and intelligent agent. The PB-ISMSF engine manages all policies of multiple heterogeneous firewalls using central policy database, and operates on policy table inside the database in response to the user's request for policy addition, deletion and modification. And, we selected SNMP for protocol between the PB-ISMSF engine and the firewall agent to give flexibility for adding new firewalls and deleting needless firewalls. It also enables a firewall agent
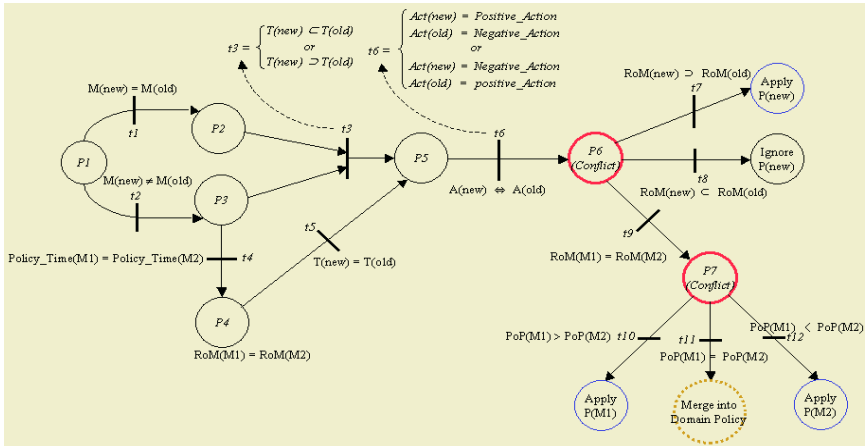
Fig. 6. State diagram for detecting and resolving for the inclusive policy conflict

to cooperate with other SNMP-based network management application. We also classified the policy conflicts, the condition of policy conflicts and briefly described the resolution methods for specified policy conflicts. In the future, we must secure communication channels between the client and the PB-ISMSF engine as well as between the PB-ISMSF engine and firewall agent using various channel security algorithms. Additionally, we are currently expanding our system to manage other security systems, for example, the vulnerability scanner, host-based access control utilities, intrusion detection systems, etc. We must apply a stronger user authentication mechanism to our management system.

References

1. William R. Cheswick, Steven M. Bellovin, Firewalls and Internet Security : repelling the willy hacker, Addison Wesley, 1994.
2. D. Brent Chapman, Elizabeth D. Zwicky, Building Internet Firewalls, O Reilly & Associations, Inc., Janyary 1996.
3. Chris Hare, Karanjit Siyan, Internet Firewalls and Network Security - 2nd ed., New Readers, 1996.
4. D. Y. Lee, D. S. Kim, K. H. Pang, T. M. Chung, "Web-based integrated security management system using SNMP", KNOM Review Vol. 2, No. 1, pp.1167-1171 April 1999.
5. D. Y. Lee, D. S.. Kim, K. H. Pang, H. S. Kim, T. M. Chung, "A Design of Scalable SNMP Agent for Managing Heterogeneous Security Systems", NOMS2000, 10-15 April 2000.
6. William Stallings, SNMP, SNMP v2, SNMP v3, and RMON 1 and 2 - 3rd ed.. Addison Wesley, 1999.
7. David Perkins, Even McGinnis, Understanding SNMP MIBs, Prentice Hall PTR, 1997
8. J. Moffett, Morris S. Sloman, "Policy Conflict Analysis in Distributed System Management", Journal of Organizational Computing, Vol.4, No.1, pp.1-22, 1994.
9. E. Lupu, M.Sloman, "Conflict Analysis for Management Policies "International Symposium on Integrated Network Management IM'97, pp.430-443, 1997.
10. Emil C. Lupu, Morris Sloman, "Conflicts in Policy-Based Distributed Systems Management", Journal of IEEE Transaction on Software Engineering, Vol. 25. No.6, pp.852-869, 1999.

Adaptive Modulation and Coding Scheme for Wireless Mobile Communication System

Jae-Sang Cha[1] and Juphil Cho[2]

1 Seo-Kyung University, Korea
`chajs@skuniv.ac.kr`
[2] Kunsan National University
San 68, Miryong-dong, Gunsan, Jeonbuk, 573-701, Korea
`stefano@kunsan.ac.kr`

Abstract. Adaptive modulation and coding (AMC), which has a number of variation levels in accordance with the fading channel variation, is a promising technique for communication systems. In this paper, the AMC method for OFDM system with cluster in pursuit of bandwidth efficiency and performance improvement is presented. And the AMC schemes applied into each cluster or some clusters are determined by the minimum or the average SNR value among all subcarriers within the corresponding cluster. It is important to find the optimal information related to cluster because AMC performance can be varied according to the number and position of cluster.

1 Introduction

Orthogonal Frequency Division Multiplexing (OFDM) technique has received considerable attention especially in the field of wired and wireless communication systems for its efficient usage of frequency bandwidth and robustness to frequency selective fading. In the conventional OFDM system, the lowest order modulation scheme has to be employed for all subcarriers so as to keep the overall objective BER performance [1]. OFDM system in conjunction with adaptive modulation (AM) technique is well known as one of the methods to solve the above problems. In the OFDM system with AM technique, the optimum modulation scheme is assigned for each subcarrier according to its instantaneous fading characteristics. In that scheme, a number of required AM information bits would be directly proportional to the great number of subcarrier including one OFDM symbol and the available modulation schemes considered in our system. Therefore, the increase of AM information bits will lead to degradation of overall system performance and unnecessary increase of both frequency bandwidth and transmission power. To solve this problem, in this paper, we use the cluster, which combines a number of subcarriers in the frequency domain, for OFDM system.

J. Campello proposed the algorithm having the optimal performance by feedback of channel information of every subcarriers and selection of appropriate modulation method to each subcarrier [2]. However, it is impossible to feedback the channel information of every subcarriers by the limited control channel of the uplink. Therefore, in order to overcome this problem, it is efficient to feedback a number of channel information with a little information. In this paper, using cluster concept, we can reduce the amount of AMC bits at a cost of slight decrease of channel capacity. Simu-

lation results show how we can use the limited bits of UFCH for efficient AMC method and obtain the optimal position and a number of clusters for the best AMC performance. For the analysis of the presented method, we apply the same AMC scheme to the subcarriers included in a cluster or subcarriers in some cluster groups.

2 Adaptive Modulation Method for OFDM System Using Cluster

Table 1 shows the system parameter and channel model used in this paper. The increase of AM information bits will lead to degradation of overall system performance and unnecessary increase of both frequency bandwidth and transmission power. To solve this problem, we use the cluster which represents the group of several subcarriers in the frequency domain. The subcarriers included in one cluster or subcarriers in some clusters can use the same modulation level and coding method. Using cluster concept, we can reduce the amount of AMC bits at a cost of slight decrease of channel capacity. Within the system considered in this paper, we use the UFCH to send the AMC information for D/L transmission.

Table 1. System parameter and channel model

Parameter	Value
Bandwidth	20MHz (D/L)
Carrier frequency	2GHz
Number of FFT	2048
Modulation	QPSK, 16QAM, 64QAM
Coding	1/2 rate convolution code (k=7)
Fading channel model	ITU-R channel (Ped. A & B)

6 kinds of configurable cluster allocation, 1, 2, 3, 4, 6 and 12 clusters, are considered in this paper and the performances of 4, 6 and 12 clusters are especially compared by the simulation. In case the number of cluster (NoC) is 1, 2 and 3, overall used numbers of bits (NoB) are 4, 8 and 12, respectively. Because of the limited 16 bits of UFCH, it is important to arrange the clusters effectively for efficient AMC information assignment. For example, if NoC is 3 and AMC level is 16, 512 information subcarriers per one cluster are used and overall bits are 12 bits ($3 \times 4 = 12$). In this case, the residual 4 (16-12) bits can be considered an inefficient bits regarding the channel structure.

Above example could be also applied to NoC = 1 and 2 cases, and 1536 and 728 subcarriers per a cluster in each case are assigned, respectively.

Maximum 32 AMC levels are considered in this paper. If the 32 and 16 AMC levels are applied, then the 5 and 4 bits for AMC information are needed for a cluster or cluster group, respectively. Therefore, in order to keep the defined 16 bits of UFCH, we should group the overall clusters into maximum 4 cluster group (CG) in case of 16 AMC levels because the allocation bits per cluster is 4. And, we actually require 2 bits for each AMC process because we use the 4 AMC levels.

So we assume the residual bits except 2 bits among UFCH bits are null bits. For example, when 4 bits for AMC information are needed for cluster or CG and 4 CG are applied, in each CG, each 2 bits are recognized null bits. In Table 4, option 1 represents that 12 clusters are used and these clusters are grouped by 4 CG. Like the clustered OFDM system [3], we restrict the CG as a group of contiguous subcarriers. That is, only adjacent subcarriers are included in a cluster or CG as illustrated in Fig. 1.

Table 2. Example of NoC, PoC and NoB (Bits for AMC info. : 5)

Option	NoC	PoC	NoB
1	12	1/(1,2,3,4), 1/(5,6,7,8), 1/(9,10,11,12)	15
2		1/(1,2,3,4,5,6), 1/(7,8,9,10,11,12), 1/(1,2,3,4), 1/(7,8,9,10)	10
3		1/(1,2,3,4,5,6,7,8,9,10,11,12)	5
4	6	1/(1,2), 1/(3,4), 1/(5,6)	15

Table 3. 5 Cases of PoC for T4O5

No. of Case	PoC
1	1/(1), 1/(6), 1/(2,3), 1/(4,5)
2	1/(3), 1/(4), 1/(1,2), 1/(5,6)
3	1/(1), 1/(2), 1/(3,4), 1/(5,6)
4	1/(1), 1/(4), 1/(2,3), 1/(5,6)
5	1/(5), 1/(6), 1/(1,2), 1/(3,4)

In Table 2, the allocated bits for a CG, which is the combination of some clusters, are 5. For example, in option 1, 5 bits for the cluster group 1 (CG1), which includes the clusters 1, 2, 3 and 4, are assigned and the total 15 ($3 \times 5 = 15$)bits for option 1 are allocated. Also, cluster group 2 (CG2) and CG3 are made up of the clusters $5 - 8$ and $9 - 12$, respectively. In Tables 2 - 4, PoC represents the position of cluster and 1/(A, B, C) means that the representative 4 or 5 bits for AMC are transmitted by obtaining minimum or average SNR value of the cluster A, B and C, which are the elements of a CG.

In these tables, a number of allocated bits per a CG increase in proportion as the number of clusters in (A, .. ,Z) becomes larger. And the performances of AMC schemes tend to be deteriorated according as the allocated bits per a CG or a cluster increase. Though the frequency hopping is considered in the frame structure, the frequency hopping is not applied actually to show the performance in this paper. Because the overall performance of system can be varied by how the clusters are allocated, we should select carefully the optimal number of clusters considering the trade-off between performance and simplicity.

Table 4. Example of NoC, PoC and NoB (Bits for AMC info. : 4)

Option	NoC	PoC	NoB
1	12	1/(1,2,3),1/(4,5,6),1/(7,8,9),1/(10,11,12)	16
2		1/(1,2,3,4),1/(5,6,7,8),1/(9,10,11,12)	12
3		1/(1,2,3,4,5,6), 1/(7,8,9,10,11,12)	8
4		1/(1,2,3,4,5,6,7,8,9,10,11,12)	4
5	6	1/(1,2), 1/(3,4), (5), (6)	16
6		1/(1,2), 1/(3,4), 1/(5,6)	12
7		1/(1,2,3), 1/(4,5,6)	8
8		1/(1,2,3,4,5,6)	4
9	4	(1), (2), (3), (4)	16
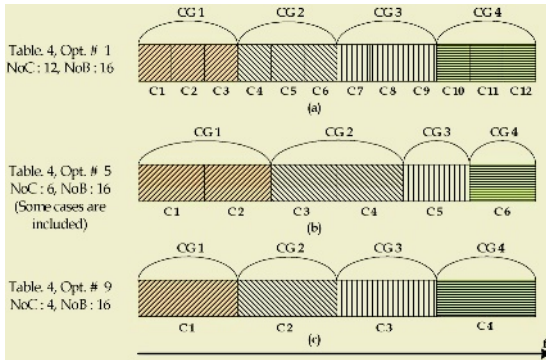
Fig. 1. Concept of cluster allocation in Table 4

The AMC scheme applied by each cluster or some clusters is decided by the mini-mum signal to noise ratio (SNR) or the average SNR value among all the subcarriers within the corresponding cluster. When we use the minimum SNR value, we compare the SNR value of each subcarrier in a cluster or each CG and the final minimum SNR value are applied into the rest of a cluster or a CG. For example, in Fig. 1 (a), the minimum SNR value of CG1 is obtained by the comparison of minimum SNR values of cluster 1 (C1), C2 and C3. And in this case, subcarriers per each cluster and a CG are 128 and 384, respectively. Similar to Fig. 1 (a), the 256 subcarriers for CG3 and CG4 are assigned in Fig. 1 (b). If the frequency hopping is not considered, Figs. 1 (a) and (c) have the same AMC performance because the 384 subcarriers are allocated for each CG. So, in this paper, we omit the simulation performance for the Fig. 1 (a). In Fig. 1 (b), the 512 subcarriers are inserted between C3 and C4. In Table 3, T4O5

means the option 5 of Table 4. These options are the representative parts of the cluster positions can be considered in T4O5.

3   Simulation Results

Here we present the simulation results of OFDM system using cluster under ITU-R Pedestrian channel environments and evaluate the performance of the presented method described in a previous section. Performances of the presented AMC scheme for OFDM system are analyzed by outage probability and bandwidth efficiency results. The outage is defined as follows. Outage = $N_{\text{Packet error \#}}/N_{\text{Packet trans \#}} + P$ unsatisfying packets of constant bit rate 3.84Mbps , where $N_{\text{packet\_error \#}}$ is the total number of received packets incorrectly and $N_{\text{packet\_trans\#}}$ is the total number of transmitted packets. Outage is averaged over 10000 OFDM blocks and bandwidth efficiency is defined as

$$BW\ efficiency = Allocated\ bits\ in\ one\ frame / (N \times symbols\ in\ frame)$$

In general, simulation results based on main options of Tables 2 ~ 4 show that we can get a better outage probability and bandwidth efficiency performance by reporting the average SNR value more than the minimum SNR value to UFCH information in pedestrian channel environment. Also, in the same table, we can see that the performance can be differentiated by the other positions of clusters. For example, in option 5 of Table 4, there are 5 main cases classified by different cluster allocation. Comparing with the performances of these 5 cases and fixed QPSK modulation, we can see that fixed modulation method outperforms the 5 cases of Table 3, which are based on Fig. 1 (b), in terms of outage probability. However, AMC shows the better characteristic than the fixed modulation when it comes to bandwidth efficiency performance. And, based on the simulation result of Table 3, we can see that 5 cases have different performance curves due to the distance among clusters even though they are in the same option. That is, the distance among clusters can be varied according to cases because 512 virtual carriers are inserted between C3 and C4. In Figs. 2 ~ 3, T2 and T4 represent the Table 2 and 4, respectively, and the O5 and O9 mean the option 5 and 9, respectively. For example, T4O5C1 implies the case 1 of option 5 in Table 4.

Fig. 2 demonstrates the outages of the AMC with different NoCs and the fixed modulation under ITU-R pedestrian A & B channels. Outage probability is 1% and E_s/N_o is 220 dB and SNR value for decision of AMC level is minimum SNR value. The transmitting power of Base-Station (BS) is determined by $P_{\text{1/2-QPSK > received SNR of cell}}$ boundary users < 0.01. In this simulation, the tx power of BS is 220 dB. Compared with the results of Ped. A channel, the outages of Ped. B channel show about $0.5 \times 10^{-2} - 2 \times 10^{-2}$ degradation at 5km, except for T4O5C5 and T2O4. And these constant gaps of performance are maintained under overall cell environment. The outages of every AMC cases with average SNR value are much smaller than those of AMC with minimum SNR values, consequently, compared with the latter, the outages for former schemes show about $1 \times 10^{-2} - 5 \times 10^{-2}$ improvement when cell radius is more than 5km. For the Ped. A and B channel under the same environments of Fig. 2, an over 2.5 improvement of bandwidth efficiency can be obtained by using AMCs instead of fixed QPSK modulation, as shown by Fig. 3. Compared with other AMC cases, the bandwidth efficiencies for T4O5C1, T4O9 and T4O5C3 are improved by

about $0.2 \sim 0.5$. As shown in Fig. 3, those 3 cases have much better performances than other schemes including QPSK in proportion to the cell radius. Based on the trade-off between outage probability and bandwidth efficiency results, in Ped. A channel, T4O9 is better than other AMC cases and fixed QPSK modulation. Also, in Ped. B channel, T4O5C5 and T4O5C3 outperform other cases in terms of outage and bandwidth efficiency, respectively. Simulation results in Ped. B channel are caused by the distance effect among clusters.
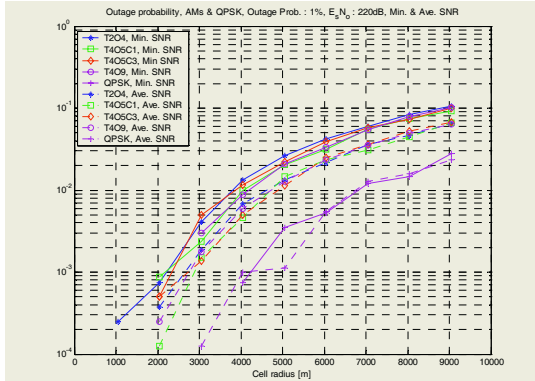
Fig. 2. Outage of AMC and fixed QPSK modulation with minimum and average SNR. (Outage probability : 1%)
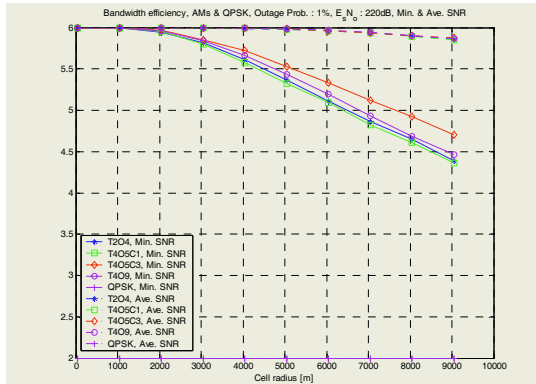
Fig. 3. Bandwidth efficiency of AMC and fixed QPSK modulation with minimum and average SNR. (Outage probability : 1%)

Comparing the performances of different UFCH allocation bits using average SNR vlue, although the outage probability for T4O5C1 shows the least value about 1.25×10^{-4} at 2km, we can see that 2 cases of T4O5 and T4O9 have similar characteristics from 3 to 10km. Because the number of clusters in two AMC schemes are 4 and 6, there are no big differences between clusters and subcarriers per cluster. Hence, both the T4O5 and T4O9 with average SNR value are more robust to the channel delay profiles than those using minimum one. Fig. 3 shows the bandwidth efficiency

under the same environment as Fig. 2. Compared with minimum SNR values, the bandwidth efficiency with average SNR shows about maximum 1.5 improvement such as outage performance. Furthermore, the bandwidth efficiencies for AMC using average SNR values are exactly same and so robust to channel delay profiles.

4 Conclusions

In this paper, it is demonstrated by computer simulation that AMC scheme with average SNR value for selecting AMC levels shows better performances than those of method using minimum SNR value. The performance of AMC method can be affected by the position and number of clusters owing to the subcarriers within the corresponding cluster or CG as well as the distance among clusters. Also, T4O9 among all AMC schemes with cluster can be considered as the optimal scheme in terms of outage and BW efficiencies for our system.

References

1. T. Ue, S. Sampei and N. Morinaga, "Symbol rate and modulation level-controlled AM/TDMA/TDD system for high-bit-rate wireless data transmission," *IEEE Trans. Veh. Technol.,* vol. 47, no.4, pp. 1134-1147, Nov. 1998.
2. J. Campello, "Practical bit loading for DMT," in *Proc. ICC'99,* June 1999, pp. 801-805.
3. Y. Li and N.R. Sollenberger, "Clustered OFDM with channel estimation for high rate wireless data," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2071-2076, Dec. 2001.

Performance Analysis of a Antenna Array System Using a New Beamforming Algorithm in the CDMA2000 1X Wireless Communication

Sungsoo Ahn[1], Minsoo Kim[2], Jungsuk Lee[3], and Dong-Young Lee[1]

[1]Dept. of Information Technology and Communication, Myongji College, Korea
{ssan64,wismss}@empal.com
[2] Dept. of Information and Communication, Donghae University, Korea
mskim1019@hanmir.com
[3] Dept. of Mechatronics, Inha Technical College, Korea
ungbolee@inhatc.ac.kr

Abstract. In order to achieve the maximum gain along the desired signal, this paper propose beamforming algorithm that utilize Generalized on-off algorithm. The new algorithm that provides a maximum SINR(Signal to Interference plus Noise Ratio) based on the conventional on-off algorithm for obtaining optimal weight vector. Also, in this paper, we present a novel demodulation method enhancing the performance of smart antenna by using the pilot channel of CDMA2000 1X channel to obtain the exact weight vector. From various computer simulation, we could conform that the proposed algorithm has superior performance in terms of SER(Symbol Error Rate) than conventional algorithm with multipath fading.

1 Introduction

This paper presents a novel beamforming algorithm of smart antenna system[1][2] in mobile communication systems. Smart antenna system obtains the optimum receiving gain in direction of desired user by adjusting the phase of each element of antenna array and combining the phase coherently compared to conventional one antenna base station system. A novel beamforming algorithm used in this paper is based on generalized on-off algorithm which utilizes the conventional on-off method[3]. While conventional on-off algorithm is to maximize the SNR(Signal to Noise Ration) of array output, a proposed novel algorithm is to maximized the SINR(Signal to Interference plus Noise) of array output to update the optimum weight vector. We made a performance analysis under multipath fading communication environment[4] in up-link of CDMA2000 1X demodulation method by using pilot channel in up-link channel system. Especially, in this paper, we propose a of CDMA2000 1X environment to demodulate the received signal accurately. There is 3 different data channels to support variable high speed data communication and pilot channel which has no data information in CDMA2000 1X system. Processing Gain(PG) in pilot channel can be set to any value arbitrarily since there is no information in pilot channel.

In this paper, we present a novel beamforming method enhancing the performance of antenna array by calculating the integration interval of pilot channel to obtain the exact weight vector.

2   The Generalized On-Off Algorithm

The proposed Generalized on-off algorithm is realized by maximizing the following equation called as generalized eigen problem[4].

$$\mathbf{R}_{yy}\,\underline{w} = \lambda_{\max}\,\mathbf{R}_{xx}\,\underline{w} \tag{1}$$

Where $\underline{w}$ is the weight vector to be computed, $\mathbf{R}_{yy}$ and $\mathbf{R}_{xx}$ are the autocovariance matrices of the despread signal $\underline{y}$ and un-despread signal $\underline{x}$, $\lambda_{\max}$ is the maximum eigenvalue corresponding to the maximum eigenvector.

In this section, we present an adaptive procedure of implementing the Generalized on-off algorithm for generating the weight vector which maximizes the ratio of the weighted despreaded $\underline{w}$, signal power to the weighted un-despreaded signal power as follows

$$P = \frac{w^H\,\mathbf{R}_{yy}\,w}{w^H\,\mathbf{R}_{xx}\,w} \tag{2}$$

Proposed algorithm depends on the sign of the gradient of the power ratio, the algorithm decides whether the present phase value should be increased or decreased.

The ^{m}th element of gradient $\underline{v}$ of P is represent as

$$\begin{aligned}
\nabla_m &= \frac{\partial P}{\partial \phi_m} = \frac{\partial P}{\partial w_m}\frac{\partial w_m}{\partial \phi_m} + \frac{\partial P}{\partial w_m^*}\frac{\partial w_m^*}{\partial \phi_m} \\
&= 2\,\mathrm{Im}[\frac{w_m^* \cdot y_m \cdot Y^* \cdot X \cdot X^* - w_m^* \cdot Y \cdot Y^* \cdot x_m \cdot X^*}{(X \cdot X^*)^2}]
\end{aligned} \tag{3}$$

Where w_m is the weight vector of ^{m}th element at the present snapshot, ϕ_m is phase of m th antenna element, X and Y are   defined as being $\underline{w}^H \underline{x}$ and $\underline{w}^H \underline{y}$. $\mathrm{Im}[\,^\bullet\,]$ means the image part of the complex $^\bullet$.

Then, the phase is updated by the positive or negative amount of the adaptive gain, which should be preset properly for the convergence. The phase vector of each antenna element is updated as

$$\underline{\phi}(n+1) = \underline{\phi}(n) + \mu \cdot sign(\underline{\nabla}) \tag{4}$$

Where n denotes the iteration number, μ is the adaptive gain by which the phase delay is increased or decreased depending on the sign of the gradient of the output power. Finally, the weight vector $\underline{w}$ is updated as

$$\underline{w}(n+1) = e^{j\underline{\phi}(n+1)} \tag{5}$$

3   Demodulation Methods in CDMA2000 1X

Let's consider a uniform linear array consisting of N antenna elements with half-wavelength spacing, where N denotes the number of antenna elements. Figure 1 is the block diagram of the method for updating the weight in the adaptive array antenna

system using the reverse pilot channel of IS2000 1X channel. In Figure 1, r_m is the received signal at the m th antenna element of the Base Station, P_I and P_Q are PN code, T of integrator denotes the integration period for walsh demodulation, and W_m^l represents Walsh code, whose length is l, of the m th low in Hadamard matrix. Note that Walsh codes are applied to each channel in order to separate channels because they are orthogonal functions. Received signal r_m is distinguished from real part and complex part and divided into PCH(Pilot CHannel), FCH(Fundamental CHannel), SCH(Supplemental Channel)1, and SCH(Supplemental CHannel)2 through the procedures of PN correlation and Walsh demodulation.
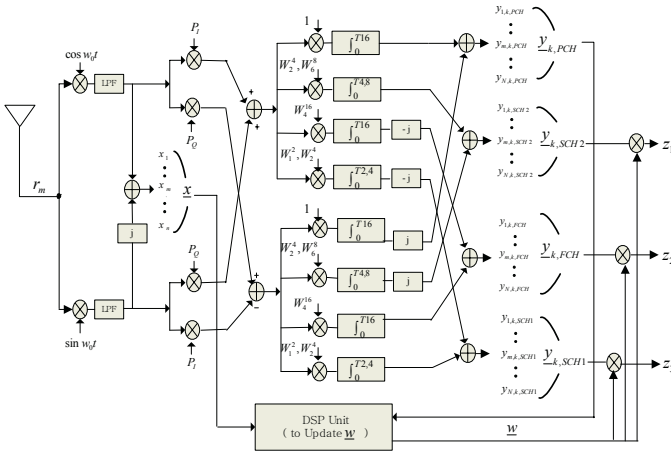
Fig. 1. Proposed demodulation structure of CDMA2000 1X

From figure 1, the un-despread signal of the m th receiving antenna, which extracted through down converting from r_m, is written as [5][6]

$$x_m(t)=\sum_{j=1}^{M}\sum_{k=1}^{K_j}(\sum_{q=1}^{L_k}s_j(t-\tau_{j,k,q})e^{j2\pi(f_d\cos\theta_{j,k,q}t-f_c\tau_{j,k,q})})e^{-j(m-m_0)\pi\sin\theta_{j,k,q}}+n_m(t))$$ (6)

where j is the user index, k is the finger index for the number of multipath from the j th user, L_k is the number of scattered component, f_d is Doppler frequency, f_c is carrier frequency, $\tau_{j,k,q}$ is propagation delay, m is the antenna index, m_0 denotes the reference antenna element, $\theta_{j,k,q}$ is incident angle of scattered signal, and $s_j(t)$ is the transmitted signal of j th user. All of the scattered components in the k th cluster of the j th user are assumed having the same propagation delay $(\tau_{j,k,q}\cong\tau_{j,k})$.

The despread signal associated with the m th antenna channel can obtained as

$$y_m(t) = \int_0^{T_s} (x_m(t) \times PN_1(t) \times W(t))dt$$

$$\text{for } m = 1, 2, \ldots, N$$

(7)

where T_s is the symbol period and PN_1 is the PN (Pseudo Noise) sequence of the first user which means the desired signal.

Despread signal vector, $\underline{y}$ is generated by despreader using the un-despread signal vector $\underline{x}$ received from the antenna array. The despread signal of the pilot channel of each antenna, $\underline{y}_{pilot}$ is transmitted to the input signal of weight update block with the receiving signal vector $\underline{x}$. The weight update block operates to find the optimal weight vector $\underline{w}$. As pilot channel does not include data, the spreading gain in the pilot channel can be controlled arbitrarily.

The updated weight vector is transmitted into the modem receiver block and multiplied to the despread signal of data channels, called weighting procedure, $z = \underline{w}^H \underline{y}_{data}$. Consequently, the final array output signal, z is generated.

4   Performance Analysis

Note that CDMA2000 1X system provides a multi-channel to each user with the spreading rate being different depending on the data rate of the data to be transmitted. We consider 3 channel, i.e., fundamental, supplemental 1, and supplemental 2, at each interfering user while the desired user transmits the signal through its fundamental channel. For this simulation, we consider signal environments encountering the following parameters. First, the integration period in the pilot channel, which determined the processing gain in computing the weight vector, has been set to 384-chip duration. Second, the angle spread is set to $\pm 10^{\circ}$ from the center DOA (Direction Of Arrival) for the desired signal. Third, the number of fingers for RAKE reception has been set to 2. Finally, the number of interferers is set to 20 and the target subscriber changes its DOA by 0.01°/snapshot where the snapshot period is set to 1ms. Figure 2 and 3 illustrate the SER performance of the proposed method with a comparison to the conventional method in CDMA2000 1X signal environments with angle spread effect[7]. Data channel performance is shown in Figure 2 and Figure 3 when only fundamental channels are used as data channels. In Figure 2 in which the angle spread for desired signal is assumed to be zero, the SER performance of the two methods are quite comparable in the CDMA2000 1X environments. However, as the signal environment is corrupted with the angle spread as shown in Figure 3, which caused the environmental statistics regarding the desired signal to vary more randomly at each snapshot, the SER performance of the conventional method is degraded much more sensitively than the proposed method. Therefore, as shown in Figure 2 and Figure 3, the performance difference between proposed algorithm and conventional algorithm is enlarged
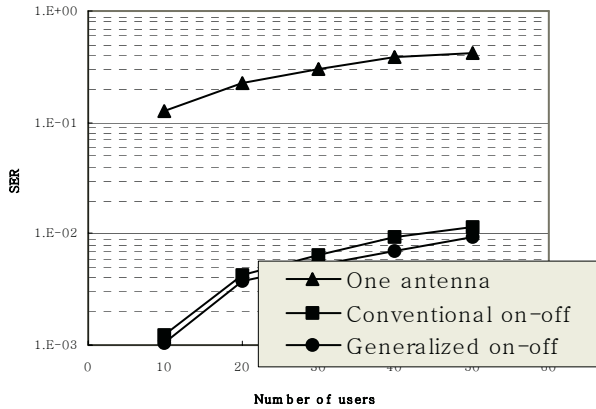
Fig. 2. SER performance with no Angle Spread
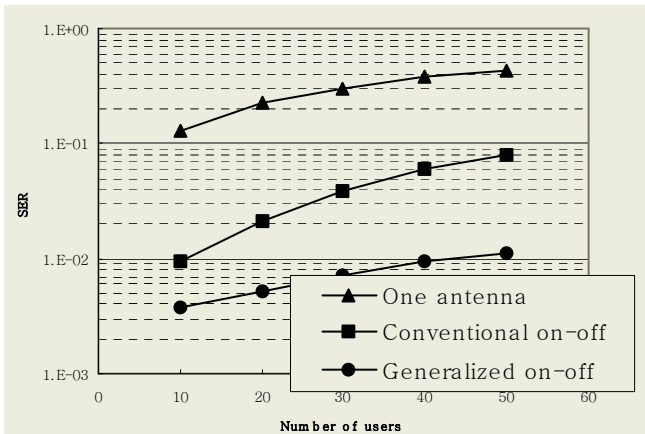
Fig. 3. SER performance with $\pm 20^0$ Angle Spread

as angle spread increases, $\pm 20^0$, SER is enhanced about 6 times when adopting two-finger rake receiver.

Figure 4 shows the performance of each case corresponding to the change of power ratio of supplemental 1 when there exist angle spread. In figure 4, assuming that the first user of the supplemental1 channels was fixed as desired signal and both other users of the fundamental channels and other channels were regarded as interferers. In case of CDMA2000 1X, high-speed data are transmitted through the supplemental channels. Therefore, they are transmitted with greater power than fundamental channels. Power ratio of simulations in this paper is assumed as 1: 4: 16 in the order of fundamental, supplemental 1, and supplemental 2 channel, respectively. As shown in Figure 4, SER performance of the proposed Generalized CGM algorithm outperforms the conventional CGM in this signal environment as well.
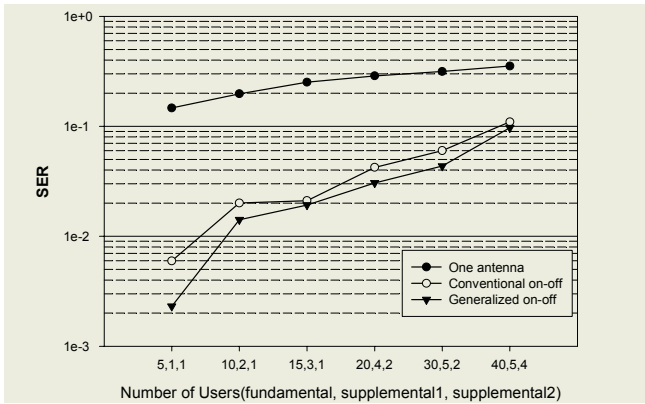
Fig. 4. SER performance when supplemental1 channel is desired signal

5   Conclusion

We have presented a new adaptive beamforming technique based on the Generalized on-off algorithm computing the weight vector of an antenna array system operating in a CDMA2000 1x environment. Also, it is confirmed that proposed demodulation method is enhanced the performance of antenna array by using the pilot channel of CDMA2000 1x channel to obtain the exact weight vector. From various simulations, it is confirmed that the proposed Generalized on-off algorithm has superior perform-ance in case of severe angle spreads because proposed method has excellent charac-teristic in spatial filtering.

References

1. J. H. Winters, "Smart Antennas for Wireless Systems", IEEE Person. Commun. Mag., pp.23-27, February 1998.
2. M. Barrett and R. Arnott, "Adaptive Antennas for Mobile Communications", Electronics & Communicational Engineering Journal, pp.203-214, Aug. 1994.
3. S. Ahn, S. Choi, and T. K. Sarkar, "An Adaptive Beamforming Algorithm with a Linear complexity for a Multipath Fading CDMA Channel", IEICE Trans. on Communication, ,Vol. E84-B, No. 8,pp2317-2320, Aug. 2001.
4. S. Choi and D. Yun, "Design of Adaptive Antenna Array for Tracking the Source of Maxi-mum Power and Its Application to CDMA Mobile Communications," IEEE Trans on An-tenna and Propagations, Vol. AP-45, No. 9, pp. 1393-1404, Sep. 1997.
5. J. S. Lee and L. E. Miller, CDMA Systems Engineering Handbook, Artech House Publish-ers, 1998.
6. A. F. Naguib, Adapative Antenna for CDMA Wireless Networks, Ph.D. Dissertation, De-partment of Electrical Engineerging, Standford University, Aug. 1996.
7. G. V. Tsoulos, "Smart Antennas for Mobile Communication Systems", Electronics and Communication Enginnering Journal, Vol. 11, No. 2. pp. 84-94, Apr. 1999.
8. Texas Instruments, Reference guide TMS320C62x/67x CPU and Instruction Set, Texas Instruments Inc. 1998.

Intelligent Tool for Enterprise Vulnerability Assessment on a Distributed Network Environment Using Nessus and OVAL*

Youngsup Kim[1], Seung Yub Baek[1], and Geuk Lee[2]

[1] Dept. of Computer Eng., Hannam University, DaeJeon, South Korea
{c4i2,psy9511}@is.hannam.ac.kr
[2] Dept. of Computer Eng., Hannam University, DaeJeon, South Korea
leegeuk@hannam.ac.kr

Abstract. This paper designs and implements a unified and intelligent tool for vulnerability assessment which can efficiently assess vulnerabilities of computer systems in which functions of the vulnerability assessment were distributed. Developing a vulnerability assessment tool needs two steps; constructing a vulnerability analysis database and implementing a vulnerability assessment tool. The vulnerability analysis database is constructed based on CVE to report vulnerabilities in standard assessment result forms by updating newest information, and is organized to be suitable to the distributed network. The vulnerability assessment tool is implemented using Nessus and OVAL. The vulnerability assessment tool suggested in this paper can provide fast and more accurate vulnerability assessment and proper guidelines to corresponding vulnerabilities.

1 Introduction

A purpose of the vulnerability assessment tool on a distributed network environment is to detect vulnerabilities of systems and to suggest appropriate security guidelines so that each system and information can be secure.

The vulnerability assessment tools are broadly categorized into a host-based assessment tool and a network-based assessment tool, and existing commercial tools and public tools were developed in one of both tools but not an unified one. And there is a problem to keep individual assessment results of each tool because most tools use their own vulnerability databases. Therefore it is necessary to construct an unified vulnerability assessment tool which can provide a standardized vulnerability database and interface with other tools [1].

Enterprise Vulnerability Assessment Tool (EVAT) mentioned in this paper has an unified vulnerability database for the network vulnerability assessment and the host vulnerability assessment. The unified vulnerability database is based on CVE(Common Vulnerabilities and Exposures), which is a vulnerability standard naming tool, and includes proper guidelines of corresponding vulnerabilities. The network vulnerability assessment tool used Nessus and the host vulnerability assess-

---

ment tool was implemented using OVAL.EVAT can assess vulnerabilities of both of networks and hosts at a same time in an unified way.

2   Related Works

2.1   Vulnerability Database

The vulnerability database is a database storing classified vulnerability information of computer systems and one of the most important parts of the vulnerability assessment tool. The vulnerability database is used to query about collected vulnerabilities and to report the assessment results[2]. The research of the vulnerability database is world-widely being performed, and some of leading research centers are Security Research Team COAST at University of Purdue and ICAT Meta Vulnerability Database of NIST.

2.2   Vulnerability Assessment Tool

The vulnerability assessment tool is a tool which investigates security weaknesses broadly distributed in networks and hosts, analyzes those weaknesses, and provides vulnerability information. The vulnerability assessment tool is normally called a scanner, and classified into a network scanner and a host scanner[1].

The network scanner assesses vulnerability of network services of a system. Scanning objects are unnecessarily opened ports, DoS, RPC, HTTP, SMTP, FTP, FINGER, and etc. Typical network scanners are Nessus, SAINT, ISS network scanner.

The host scanner assesses security weaknesses within hosts. The host scanner has a strong point for detecting and diagnosing weaknesses at a host level in detail and accurately. But it is mostly platform dependent, should be installed at all target hosts, and needs administrator's permission to be executed[3]. Typical host scanners are COPS, Tiger, STAT, CyberCop but most of them have limitations of checking vulnerability in a distributed environment.

2.3   CVE(Common Vulnerabilities and Exposures)

CVE is a naming-schema to give standard names to vulnerabilities of a computer system. Now, a lot of security companies take part in standardizing names of vulnerabilities [4]. Vulnerability assessment tools not using the CVE based vulnerability database have an interoperable problem with other vulnerability assessment tools since those tools assess vulnerability by their own security policies.

2.4   OVAL(Open Vulnerability Assessment Language)

OVAL, suggested by MITRE, is a standard language to check the presence of vulnerability based on specific information of a system, and its standardization is currently being under going [2, 5]. OVAL uses XML script and SQL script to describe OVAL definition which includes vulnerability definitions and conditions of vulnerabilities on

the target system. Different OVAL definitions are necessary for different operating platforms because each platform may have different conditions of the presences of vulnerabilities. It does not directly apply simulation attacks to detect vulnerabilities but uses vulnerability information existing at the software of the system.

3   Design and Implementation of EVAT

3.1   Structure of Vulnerability Assessment System

EVAT consists of a vulnerability database, vulnerability assessment tools (a network scanner and a host scanner), an agent, and a web manager (Fig. 1).
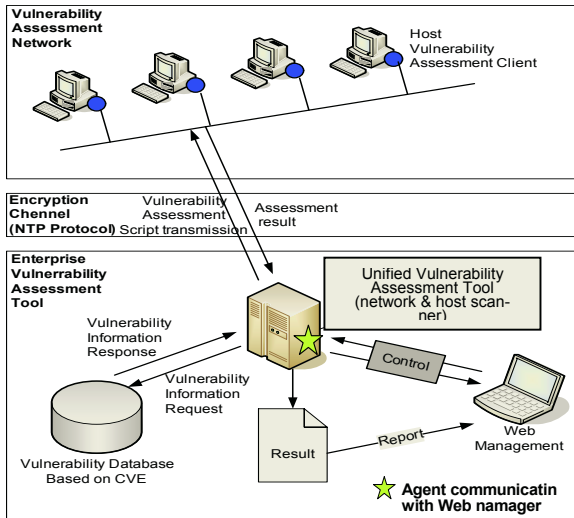
Fig. 1. System structure of EVAT

The web manager uses the agent to make both scanners assess vulnerabilities of the system. The scanners detect vulnerabilities, query the vulnerabilities to the vulnerability database, and report the assessment results to the web manager. Because of using NTP protocol and encrypting data packets in SSL, vulnerability information transferred within EVAT is surely secure against packet sniffing [6].

3.2   Vulnerability Database

The vulnerability database consists of schemas described in Table 1, and can be opened to the external through a web interface. It uses CVE ID as an identifier of the vulnerability, and provides 100% CVE compatibility with other vulnerability databases because of using CVE. It can be optimized by parsing I-CAT DB, and includes web clients which contain a module performing real-time update of the database.

3.3   Enterprise Vulnerability Assessment Tool

The unified vulnerability assessment tool consists of the network scanner and the host scanner, interfaces with the vulnerability database, diagnoses vulnerabilities, and reports the results.

Table 1. Entities of the vulnerability database

Entity	Contents
Vulnerability	General contents of the vulnerability. CVE ID, Type, Risk Level, Damage Type, Domain Type, requirement for attack, and etc
Related Reference	Corresponding reference information to the vulnerability
Vulnerable Software	Information of vulnerable Software. Same as a reference entity, one vulnerability may include more than one vulnerable software.
Vulnerability Name	Name of the vulnerability. Classified into English and Korean
Vulnerability Summary	Summary of the vulnerability
Vulnerability Description	Description of the vulnerability
Vulnerability Solution	Solution of the vulnerability
Vulnerability Author	User who updated information of the vulnerability
User	User information

3.3.1   Network Vulnerability Assessment Tool(Network Scanner)

The network scanner uses Nessus [7]. Nessus can periodically and rapidly update the vulnerability script by the user group, includes lots of plug-in scripts, and can diagnoses large scale networks. The network scanner communicates with the web manager using NTP protocol and SSL encryption.

3.3.2   Host Vulnerability Assessment Tool(Host Scanner)

The host scanner is implemented based on OVAL on Redhat9 Linux platform. The host scanner collects vulnerability information of hosts using the host vulnerability assessment client installed on each host, analyzes those information using the OVAL interpreter, queries to the vulnerability database, and then reports the assessment results [6]. The host scanner can detect all the vulnerabilities registered on CVE and describe vulnerabilities clearly and logically because the vulnerability assessment script is written in XML and SQL [2, 5].
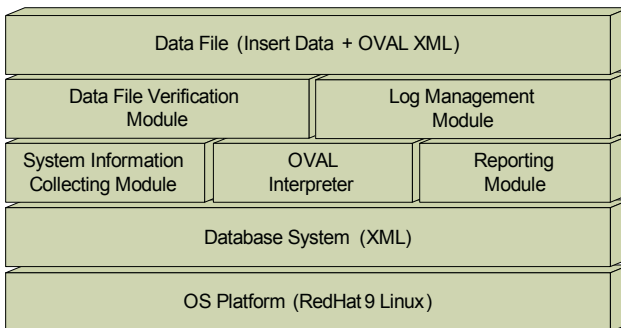
Fig. 2. Overall Structure of Host Vulnerability Assessment Tool

The host scanner includes system information collecting module, OVAL defini-
tion, and OVAL definition interpreter (Fig. 2). The system information collecting
module collects host information such as system, file, network demon, process, RPM
package, shadow password, operating platform, and etc. The OVAL definition is
specified using vulnerability definitions which MITRE standardized and has names of
"OVAL + ID-number"[4]. The OVAL definition interpreter interprets the OVAL
definition, decides what kinds of information to be collected, makes the system in-
formation collecting module collect the information, compares collected information
with conditions of vulnerability described in the OVAL definition, and then forwards
results to the reporting module.

3.4   Web-Manager System

The web manager communicates with scanners using the agent. The web manager
requests scanner to check vulnerabilities of the system, and the scanners report the
assessment results to the web manager. At any computers in which the web manager
is installed, the vulnerability assessment is available in a real-time just after logging
in the web manager. The web manager can adjust the environment setup and plug-in
setup for assessing the vulnerability.
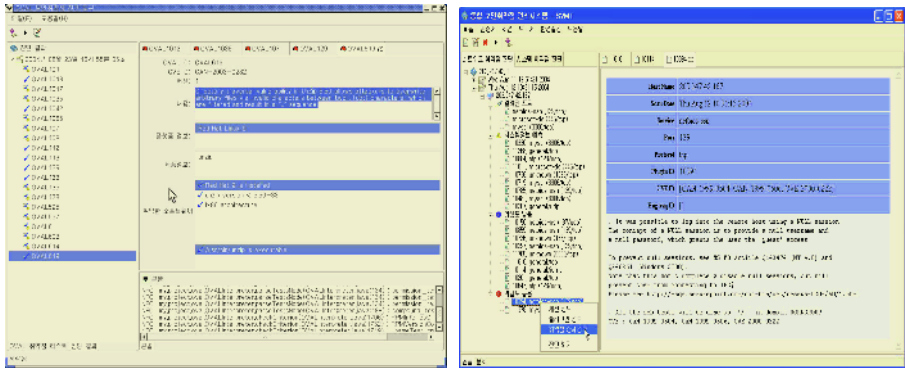
Fig. 3. Result of Web-Manager's vulnerability assessment

Examples of vulnerability assessment results displayed on web manager's screen
are in Fig. 3. The left is a snapshot of the host vulnerability assessment on RedHat9
Linux platform. It tells, by OVAL619 schema of the OVAL definition, a vulnerability
of using /usr/bin/unzip file which version is 2.50-30 or below.

The right is a snapshot of the network vulnerability assessment by performing a
port scanning. It shows a possibility of remote log-in using a NULL session and a
guideline to prevent the NULL session.

4   Conclusion

The purpose of the vulnerability assessment tool, EVAT, implemented in this paper is to remotely detect the vulnerabilities of networks and hosts and to provide proper guidelines through WEB system on the distributed network environments.

Because EVAT uses CVE to construct standard vulnerability database, it can interface with other tools using CVE based vulnerability database and easily updated.

EVAT can reduce time-cost and increase maintenance efficiency of a system since it is a centralized system which can assess networks and hosts at a same time and provide vulnerabilities and guidelines together. EVAT is implemented on Redhat9 Linux platform andthe vulnerability assessment scripts are periodically updated; the network vulnerability script is by Nessus user group, and the host vulnerability script is by MITRE.

Data packets transferred within EVAT are secure because of using NTP protocol and encrypting in SSL.

References

1. Jung Hee Kim, "The Trends of International Standardization on Vulnerability Assessment of Computer", http://www.secureosforum.org/download/data_read.htm?id=56&start=0, The Trends in International Information Security, Korea Information Security Agency, June 2003
2. Ragi Guirguis, Network and Host-based Vulnerability Assessments, http://www.securitytechnet.com/resource/security/consulting/1200.pdf, February 2004.
3. UNIX Security Checklist v2.0, http://www.cert.org/tech_tips/usc20_full.html
4. Introduction to CVE, The Key to Information Sharing, http://cve.mitre.org/docs/docs2000/key_ to_info_shar.pdf
5. M. Wojcik, T. Bergeron, T. Wittbold and R.Roberge, Introduction to OVAL, http://oval.mitre.org/documents/, MITRE Corporation, November 2003.
6. Young Mi Gwon, Hui Jae Lee, Geuk Lee, " A Vulnerability Assessment Tool Based on OVAL in Linux System", IFIP International Conference, LNCS3222, pp653-660, October 2004
7. Introduction to Nessus, http://www.securityfocus.com/infocus/1741

Design and Implementation of SMS Security System for Wireless Environment

Yan-Ha[1], Hea-Sook Park[1], Soon-Mi Lee[1], Young-Whan Park[1], and Young-Shin Han[2]

[1]Dept. of Computer & Information Technology Kyungin Womens's College
548-4 Gyesan-dong, Gyeyang-gu, Incheon 407-740 Korea
`{edpsphs,white,leesm,yhpark}@kic.ac.kr`
[2] Center for Computer Graphics and Virtual Reality Ewha Womens University
11-1 Daehyun-dog, Sudaemun-gu, Seoul 120-750 Korea
`yshan95@ewhain.net`

Abstract. This paper aims at developing communication module and application program for client management module and developing database management module and managing wireless communication facilities for server systems. To construct these aims, we have adapted DES algorithm and researched on encrypting and decrypting module development applicable to SMS Security System and optimize module size and processing speed. In addition, the XML technology is applied to encrypt only the contents of the message in a way to reduce the volume of data to process and shorten the time to encrypting.

1 Introduction

1.1 Background of the Study

SMS(Short Message Service) is a two-way mobile data communication service for mobile (cell-phones, PDA) may exchange short messages without additional equipment [1-2]. SMS is widely used domestically due to the accuracy and promptness and other advantages of SMS, and GSM(Global System Association) reported that 19 billion messages were transmitted per month worldwide last year [3]. In addition, the on-going researches are focusing on the system research and development for M-Commerce by using the advantage of SMS under the mobile Internet environment. However, there are still insufficient efforts in resolving the shortage of SMS, the issues of security (exposing or altering the message).

These shortages may become the factors in interfering with the safe e-commerce environment. Therefore, if the security of SMS is strengthened to remove the risks on information exposure and alteration, it would be applied effective on the areas of services on order of goods for M-Commerce and E-Commerce, proceed payment, confirmation of payment, forwarding of electronic receipt and others, services on balance inquiry, transfer, introduction of new financial product on mobile internet banking service, and services on various auction service, reservation service area.

In this thesis, in order to present the message exposure by a third party and ensuring alteration of messages from the SM transmission, the security system for SMS with the basis of applying the encryption technology is to be developed. For this pur-

pose, the DES algorithm is to be applied with the appropriate algorithm to use the small resource of mobile device effectively.

Furthermore, this thesis applied the XML when encrypting/decrypting of the messages. The advantages of applying XML are, first, which the volume of data to encrypt/decrypt decreases. Therefore, it reduces the time required and use of the system resource less. Second, in the event that the information of the part not encrypted is required, the necessary information may be used without the decryption process. Examples are the sender ID, receiver ID, sending date, and title of message and others. Third, the encrypted part is not to be parsing prior to decrypt that the efficiency of transmission can be heightened [4-6].

2   Related Studies

SMS related researches are mainly the studies on system structuring. That, on [7], the system to deliver the stock information with voice is developed and it designed the Voice XML server module for stock information, SMS server module to perform the authentication procedure, the PSTN network, CTI module for interface and others. At this time, the SMS certification based on random disposable password is attempted as for the method of certification.

From [8], the existing SMS is designed to operate on the main memory data structure that the DBMS concept is applied to point out and supplement the shortcomings of recovery function simultaneous control function deficiency. On [9], the PGP is applied on the web-based e-mail service system to improve the security, and use the applet and sublet communication method using the java language to resolve the security issue between the client and server. At this time, in order to encrypt the message, it applies the IDEA algorithm of secret key method, and at this time, the RSA algorithm of open key method is applies to encrypt the secret key. [10] proposed the safe instant messenger system that an individual user encrypts the individual information that is sent to the messenger server, and selectively encrypts the information delivered between the users to resolve the potential security issues. And, the system design is made on the basis of diffused processing environment by the JAVA RMI that saves the communication overhead comparing to the client-server structure, and the DES encryption algorithm is used for the encryption of transmitted details, and the exchange of the DES encryption key uses the Diffie-Hellman key exchange method and the RSA encryption algorithm.

3   The Design and Implementation of SMS System

3.1 Security Requirements of System

In this thesis, the security requirement needed to the system under the mobile environment is presented as follows. First, in order to maintain the confidentiality, the secret key uses the random number generator to generate 1024 bits of random numbers, and in order to keep the secret key, it uses the Diffie-Hellman key distribution algorithm. Second, in the event the receiver gets the message when designing the

database for the non-repudiation function, it shall be able to confirm. Third, for the mobile authentication and user authentication, the database is designed accordingly.

3.2 Structure of System

The entire structure of the SMS system is like the Fig. 1.
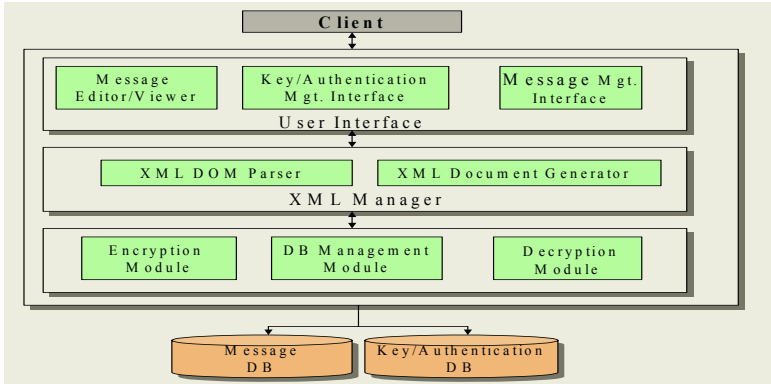
Fig. 1. System Architecture of SMS

The client performs the functions of message management and transmission, user and group management. If the client wishes to transmit the message, it logs in on the server, prepares and encrypts the message and the encrypted message is transmitted to the server. The message prepared at this time becomes the encryption with the secret key of the client. If another client receives the messages from the server, receive the encrypted message after logging in. The message encryption in its own secret key is decrypted.

The message management function of server is largely divided into two stages. Server decrypts the encryption message transmitted from the client, then encrypts again with the secret key of the receiver to store in the database. Server makes the contents of transmitting messages unable to read. If the receiving client wants the message reception, the encrypted message is forwarded after the client certification. At this time, the secret key of the XML Manager encrypts the message. If the receiving client wishes to receive the message from the server, log in first and then use the individual secret key to decrypt the encrypted message.

4   Implementation

The server of the developed SMS security system is implemented on the computer with Intel Pentium IV 2GHz 256 DDR RAM, and the development tool is used of Visual Studio 6.0. And, in order to develop the client program, the Embedded Visual Studio 3.0 and Pocket PC SDK are used on the above PC. And, in order to test the program, the server program developed from the above PC is implemented and the Win CE implements the client program on the Compaq IPAQ used as the OS.

4.1 Client Implementation

4.1.1 Client Message Reception Function

Fig. 2 is the interface of Client to receive the messages. Through these interface, the received message may be read or processed.
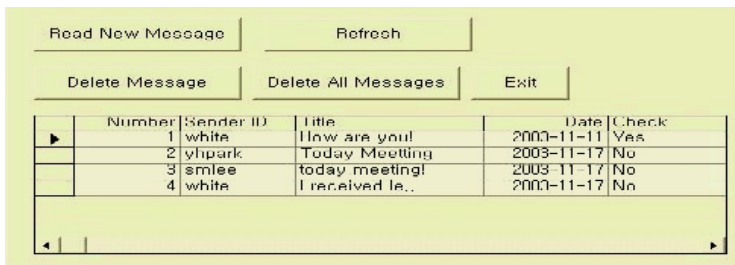
Fig. 2. Decrypted and Received Messages

Fig. 3. Transmission of Message

4.1.2 Message Transmission Function

Fig. 3 shows the user interface to forward new messages. Through this screen, the message is transmitted to the entire group designated or to one particular user. At this time, press the transmission button to forward the message that is inputted on the "content" window is encrypted on the DES algorithm by using the "user password" of the user. The example of a message inputted through the user interface to a XML document form is shown in Table 1.

Table 1. XML Document

```
<?xml version="1.0" encoding="euc-kr"?>
<msginfo>
     <Num>Se_0001N</Num>
     <Send_ID>edpsphs</Send_ID>
     <Recv_ID>white</Recv_ID>
     <Recv_Group>sms</Recv_Group>
    <Msg_Title>Full paper and registration!</Msg_Title>
    <Msg>Full paper and registration are due by January 31, 2004.</Msg>
     <Send_Date>2003-11-17</Send_Date>
</msginfo>
```

The XML document that encrypted the XML document of Table 1 is the same as Table 2.

Table 2. The Encrypted XML Document

```
<?xml version="1.0" encoding="euc-kr"?>
<DocEncryptedElement algorithm="DES/CBC/PKCS5Padding" contentType="text/xml"
encoding="base64" iv=
"TPjfzrhsNvA">AtIvI4n+GOoo6p/WV)dNJyZY4AD4Tt6Q/frbBJS7FaKjouxojyBHw==</D
ocEncrptedElement>
<msginfo>
<Num>Se_0001N</Num>
<Send_ID>edpsphs</Send_ID>
<Recv_ID>white</Recv_ID>
<Recv_Group>sms</Recv_Group>
<EncrptedElement algorithm="DES/CBC/PKCS5Padding"
 contentType="text/xml"encoding="base64" iv hrc7PbRXy
   MY=">NfPTROLpBJR4aE55bMnihYaOjakrMMcdGPOpFKysOw=</EncrypteElement>
<Send_Date>2003-11-17</Send_Date>
</msginfo>
```

Table 3. Comparison With Other Messenger System

Compared Item	Messenger System		SMS System	
	System with-out security function	System providing security function	System without security function	System providing security function
Implemented System	ICQ[11], MSN[12], UIN[13]	MaXIM[14]	Many	System proposed in this study
Applying Algorithm	None	RSA, DES	None	DES
XML Security	Not supported	Not supported	Not supported	Supported
Message Confidentiality	Not supported	Supported	Not supported	Supported
Sender Authentication	Not supported	Supported	Not supported	Supported
DB Storage Function	Not supported	Supported	Not supported	Supported
Mobility	Not supported	Not supported	Supported	Supported
Equipment	Desktop	Desktop	PDA, Mobile Phone	PDA, Mobile Phone

4.2 Server Implementation

When the server program is implemented, it generates the screen as shown on Fig. 4. The server manager may confirm the sender, receiver, group ID, message, sent date, received date and others related to the wait message through this screen. In addition, the server manager may delete the unnecessary messages from the database. At this time, the message displayed on the screen is encrypted that the server manager also is not able to confirm the message.
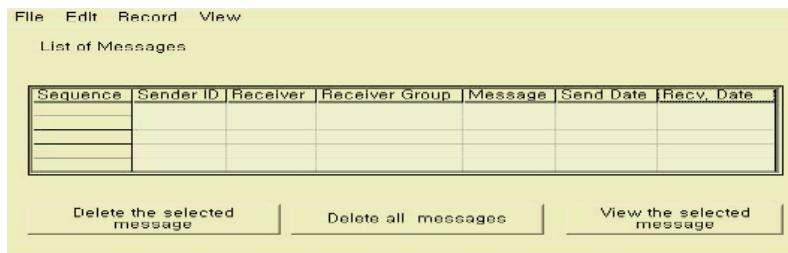
Fig. 4. Application of Server Side

4.3   Comparison with Other Messenger System

This study is a study proposed to overcome the shortcomings in the security aspect comparing to the possibility of utilization for the SMS system that the intent is to evaluate the functional aspects related to the security.

5   Conclusion and Further Studies

In this study, safe message use environment is implemented by applying the security technology on the SM that is broadly used under the mobile environment today. For this purpose, it applies the security algorithm, DES, of symmetric key method, and implements the SMS. SMS is consisted of DB server module, encryption/decryption module and server interface module. And it is designed and implemented by encrypting the message prepared by the sender to store in the database of the SMS server, and perform the reception and decryption of the message by the receiving party. SMS that is developed by this study is evaluated as satisfying all sectors, such as message confidentiality, sender authentication, mobility and others when compared to the other messenger systems. The system implemented by this thesis is evaluated as applicable to many fields serviced under the mobile Internet environment. In particular, the XML technology is applied when encrypting message, it encrypts/decrypts only the contents of the message to cut down the volume of data. It increases the convenience to confirm the contents not encrypted without decrypt entire volume.

References

1. H. S. Choi, "Reports of Mobile Services 2003", Softbank Korea, pp.20-78.
2. K. C. Kim, "The trend of Mobile Service", Journal of KIPS, Vol. 9. No 2, March 2002, pp.17-23.
3. "The survey of Mobile Service", http://www.mic.org, June 2003.
4. Miyazawa T., Kushida T., "An advanced Internet XML/EDI Model Based on Secure XML Documents", Parallel and Distributed System, IEEE, Iwate, Japan, June 2000, pp.295-300.
5. Tidwell D., "The XML Security Suite: Increasing the security of e-business", http://www4.ibm.com/software/.
6. Maruyama H., Imamura T., "Element-Wise XML Encryption", http://lists.w3.org/Archives/Public/xml-encryption/2000Apr/att-0005/01-xmlenc.html, April 2000.

7. S. I. Ou, J. H. Go, "Voice Portal based on SMS Authentication at CTI Module Implementation by Speech Recognition", Journal of KISS, Vol.28 No.1, April 2001, pp. 439- 441.

8. W. H. Lee, "Performance Evaluation of Short Message Service System based on Main-Memory DB and Disk-Resident DB", Journal of KISS, Vol.27 No.2-1, Oct. 2000, pp. 216-218.

9. T. H. Kim, H. K. Kang, K. J. Kang, "Implement of Multimedia Messaging Service", Journal of KISS, Vol. 27 No.2-3, Oct 2000, pp. 343- 345.

10. S. B. Kim, K. S. Kim, "Design of Algorithm for Preventing Decryption of Inaccurate Cipher text", Journal of KCIE, Vol.3, No.5, 2002.

11. ICQ, http://www.icq.com

12. MSN Messenger Service, http://messanger.msn.com

13. UIN Messenger, http://uin.com

14. MaXIM, http://www/dreamsecurity.com/products/products_

Intelligent Method
for Building Security Countermeasures
by Applying Dr. T.H. Kim's Block Model

Tai-hoon Kim[1] and Seung-youn Lee[2]

[1] San-7, Geoyeo-Dong, Songpa-Gu, Seoul, Korea
taihoon@empal.com
[2] SKK Univ., Dept. of Information & Communication Eng. Kyonggi, 440-746, Korea
syoun@ece.skku.ac.kr

Abstract. Because the networks and systems become more complex, the implementation of the security countermeasures becomes more critical consideration. The designers and developers of the security policy should recognize the importance of building security countermeasures by using both technical and non-technical methods, such as personnel and operational facts. Security countermeasures may be made for formulating an effective overall security solution to address threats at all layers of the information infrastructure. This paper uses the security engineering principles for determining appropriate security countermeasures. This paper proposes a intelligent method for building security countermeasures by modeling and dividing IT systems and security components into some blocks.

1 Introduction

When we design general or some special IT systems, we may provide a framework for the assessment of quality or security characteristics by considering some approaches and methods. And this framework can be used by organizations involved in planning, monitoring, controlling, and improving the acquisition, supply, development, operation, evolution and support of IT systems.

In the general cases, security countermeasures for IT systems are implemented in buying and installing some security products such as Firewall, IDS and Anti-virus systems. But the scope of IT systems is being extended and the security holes are increased. Most of the threat agents' primary goals may fall into three categories: unauthorized access, unauthorized modification or destruction of important information, and denial of authorized access. Though any cases are occurred, the compromise of IT system may be connected to loss of money or job. Therefore, Security countermeasures must be implemented to prevent threat agents from successfully achieving these goals [1-4].

This paper proposes a method for building security countermeasures by modeling and dividing IT systems and security components into some blocks. In facts, IT systems are very complex and consist of very many components. So we can't help dividing IT systems into some parts. And we categorize security components into some groups.

Security countermeasures should be considered with consideration of applicable threats and security solutions deployed to support appropriate security services and objectives. Our Block model may be used to make security countermeasures in any cases. Because the size of each block expresses parts insufficient in security.

2   Design Flow of IT Systems Security Countermeasures

We published a design method for IT systems security countermeasures a few months ago. In that model, we identified some components we should consider for building security countermeasures. In fact, the Procedure we proposed is not perfect one yet, and the researches for improving are going on.

2.1   Determination of Robustness Strategy

The robustness strategy is intended for application in the development of a security solution. An integral part of the process is determining the recommended strength and degree of assurance for proposed security services and mechanisms that become part of the solution set. The strength and assurance features provide the basis for the selection of the proposed mechanisms and a means of evaluating the products that implement those mechanisms.

Robustness strategy should be applied to all components of a solution, both products and systems, to determine the robustness of configured systems and their component parts. It applies to commercial off-the-shelf (COTS), government off-the-shelf (GOTS), and hybrid solutions. The process is to be used by security requirements developers, decision makers, information systems security engineers, customers, and others involved in the solution life cycle. Clearly, if a solution component is modified, or threat levels or the value of information changes, risk must be reassessed with respect to the new configuration.

Various risk factors, such as the degree of damage that would be suffered if the security policy were violated, threat environment, and so on, will be used to guide determination of an appropriate strength and an associated level of assurance for each mechanism. Specifically, the value of the information to be protected and the perceived threat environment are used to obtain guidance on the recommended strength of mechanism level (SML) and evaluation assurance level (EAL) [5].

2.2   Consideration of Strength of Mechanisms

SML (Strength of Mechanism Levels) are focusing on specific security services. There are a number of security mechanisms that may be appropriate for providing some security services. To provide adequate information security countermeasures, selection of the desired (or sufficient) mechanisms by considering particular situation is needed. An effective security solution will result only from the proper application of security engineering skills to specific operational and threat situations. The strategy does offer a methodology for structuring a more detailed analysis. The security services itemized in these tables have several supporting services that may result in recommendations for inclusion of additional security mechanisms and techniques.

2.3   Selection of Security Services

In general, primary security services are divided five areas: access control, confidentiality, integrity, availability, and non-repudiation. But in practice, none of these security services is isolated from or independent of the other services. Each service interacts with and depends on the others. For example, access control is of limited value unless preceded by some type of authorization process. One cannot protect information and information systems from unauthorized entities if one cannot determine whether that entity one is communicating with is authorized. In actual implementations, lines between the security services also are blurred by the use of mechanisms that support more than one service.

2.4   Application of Security Technologies

An overview of technical security countermeasures would not be complete without at least a high-level description of the widely used technologies underlying those countermeasures. Next items are some examples of security technologies.

- Application Layer Guard, Application Program Interface (API).
- Common Data Security Architecture (CDSA).
- Circuit Proxy, Packet Filter, Stateful Packet Filter.
- CryptoAPI, File Encryptors, Media Encryptors.
- Cryptographic Service Providers (CSP), Certificate Management Protocol (CMP).
- Internet Protocol Security (IPSec), Internet Key Exchange (IKE) Protocol.
- Hardware Tokens, PKI, SSL, S/MIME, SOCKS.
- Intrusion and Penetration Detection.
- Virus Detectors.

2.5   Determination of Assurance Level

The discussion of the need to view strength of mechanisms from an overall system security solution perspective is also relevant to level of assurance. While an underlying methodology is offered by a number of ways, a real solution (or security product) can only be deemed effective after a detailed review and analysis that consider the specific operational conditions and threat situations and the system context for the solution.

Assurance is the measure of confidence in the ability of the security features and architecture of an automated information system to appropriately mediate access and enforce the security policy. Evaluation is the traditional method that ensures the confidence. Therefore, there are many evaluation methods and criteria exist. In these days, the ISO/IEC 15408, Common Criteria, replaces many evaluation criteria such as ITSEC and TCSEC.

The Common Criteria provide assurance through active investigation. Such investigation is an evaluation of the actual product or system to determine its actual security properties. The Common Criteria philosophy assumes that greater assurance results come from greater evaluation efforts in terms of scope, depth, and rigor.

Next figure is the summarized concepts we proposed.
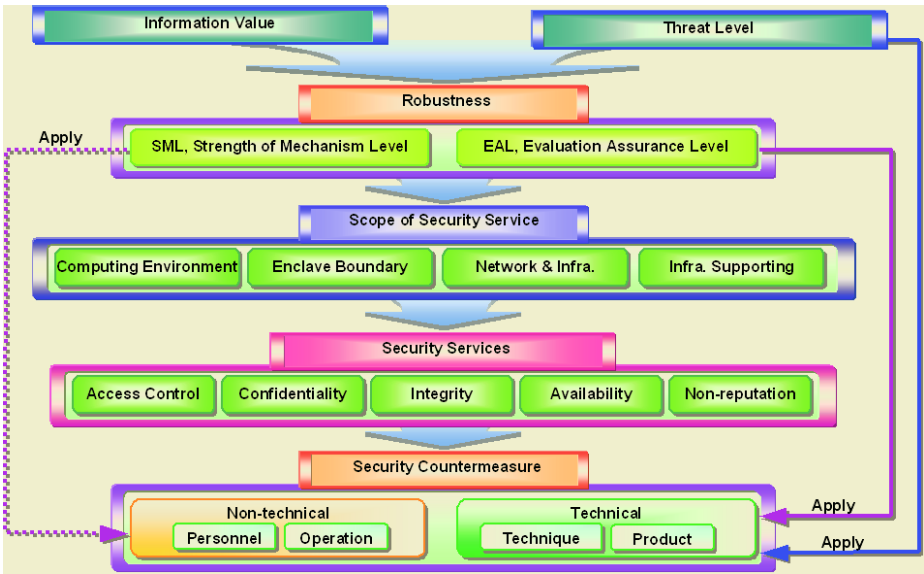
Fig. 1. Example of Design Procedure for Security Countermeasures

Fig.1 is an example of design procedures for security countermeasures. Now we are researching about this model for more perfect one.

3   Dividing IT Systems

Implementation of any security countermeasure may require economic support. If your security countermeasures are not sufficient to prevent the threats, the existence of the countermeasures is not a real countermeasure and just considered as like waste. If your security countermeasures are built over the real risks you have, maybe you are wasting your economic resources.

First step is the division of IT systems into some parts (See Fig.2). In this paper, we divide IT systems into 4 parts. But we think this partition is not perfect one and we are now researching about that.

Each part may have three common components such as Technique, Product, and Operation and Personnel.
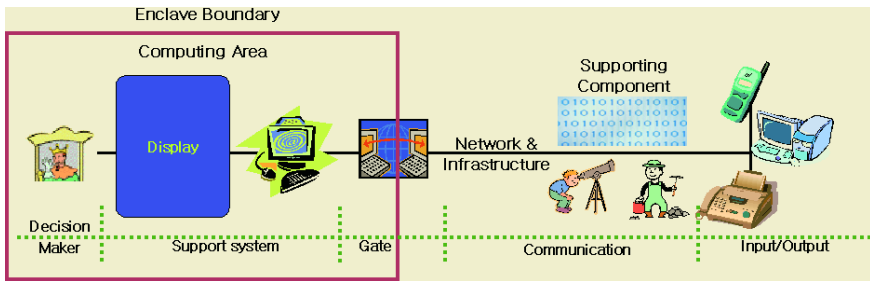
Fig. 2. Division of IT systems

Next step is construction of block matrix by using the parts of IT systems and common components we mentioned above (See the Fig. 3). This method for construction of block for complex system was proposed by Dr. Tai-hoon Kim, an author of this paper.
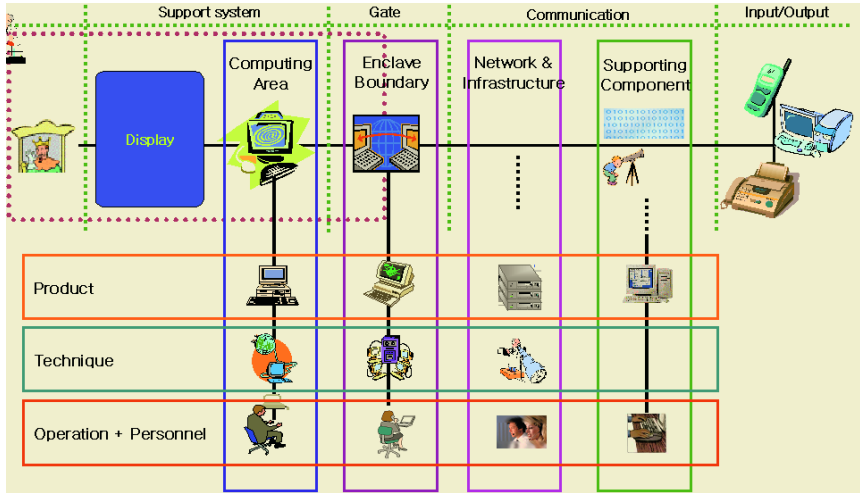
Fig. 3. Block matrix

Each cross point area of Fig.3 may be generalized and reduced into Block and matrix of Fig.4. Each Block may mean the area require security countermeasures or security method.

Next step is determination of security assurance level of IT systems. Security assurance level is related to the robustness decided in Fig.1. In the concept of our Block model, all cross point area should be protected by security countermeasures.
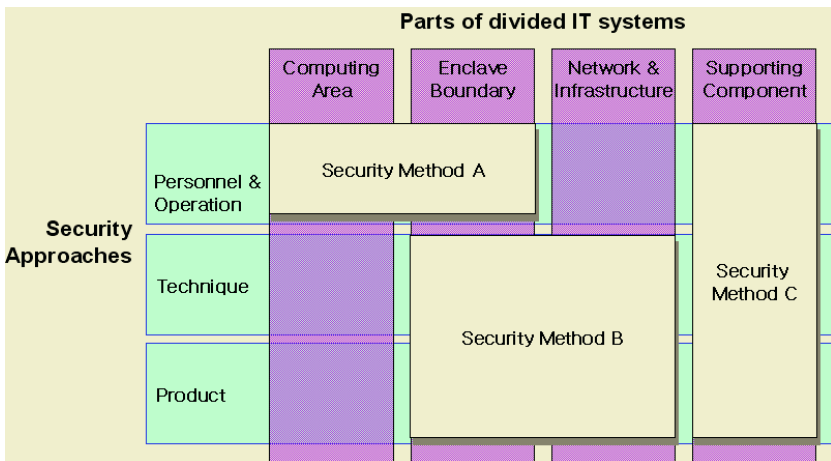
Fig. 4. Security Approaches and Methods for divided IT systems
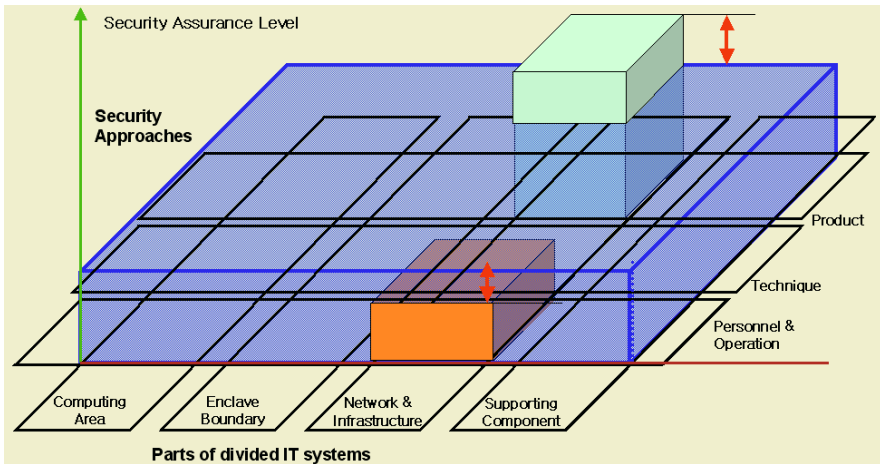
Fig. 5. Building security countermeasures by using Block Region

Robustness is connected to the level or strength of security countermeasures and this idea is expressed like as Fig.5. The last step may be building security counter-measures by using Block Region.

4   Conclusions

This paper proposes a method for building security countermeasures by modeling and dividing IT systems and security components into some blocks.

In facts, IT systems are very complex and consist of very many components. So we can't help dividing IT systems into some parts. And we categorize security components into some groups.

Our Block model may be used to make security countermeasures in any cases. Because the size of each block expresses parts insufficient in security. For more detail work, we are now researching about the Block Model for Building Security Countermeasures.

References

1. Tai-Hoon Kim, Byung-Gyu No, Dong-chun Lee: Threat Description for the PP by Using the Concept of the Assets Protected by TOE, ICCS 2003, LNCS 2660, Part 4, pp. 605-613
2. Tai-hoon Kim and Haeng-kon Kim: The Reduction Method of Threat Phrases by Classify-ing Assets, ICCSA 2004, LNCS 3043, Part 1, 2004.
3. Tai-hoon Kim and Haeng-kon Kim: A Relationship between Security Engineering and Security Evaluation, ICCSA 2004, LNCS 3046, Part 4, 2004.
4. Eun-ser Lee, Kyung-whan Lee, Tai-hoon Kim and Il-hong Jung: Introduction and Evalua-tion of Development System Security Process of ISO/IEC TR 15504, ICCSA 2004, LNCS 3043, Part 1, 2004
5. Tai-hoon Kim , Chang-wha Hong and Sook-hyun Jung: Countermeasure Design Flow for Reducing the Threats of Information Systems, ICCMSE 2004, 2004

6. Tai-hoon Kim, Tae-seung Lee, Kyu-min Cho, Koung-goo Lee: The Comparison Between The Level of Process Model and The Evaluation Assurance Level. The Journal of The Information Assurance, Vol.2, No.2, KIAS (2002).
7. Tai-hoon Kim, Yune-gie Sung, Kyu-min Cho, Sang-ho Kim, Byung-gyu No: A Study on The Efficiency Elevation Method of IT Security System Evaluation via Process Improvement, The Journal of The Information Assurance, Vol.3, No.1, KIAS (2003).
8. Tai-hoon Kim, Tae-seung Lee, Min-chul Kim, Sun-mi Kim: Relationship Between Assurance Class of CC and Product Development Process, The 6th Conference on Software Engineering Technology, SETC (2003).
9. Ho-Jun Shin, Haeng-Kon Kim, Tai-Hoon Kim, Sang-Ho Kim: A study on the Requirement Analysis for Lifecycle based on Common Criteria, Proceedings of The 30th KISS Spring Conference, KISS (2003)
10. Haeng-Kon Kim, Tai-Hoon Kim, Jae-sung Kim: Reliability Assurance in Development Process for TOE on the Common Criteria, 1st ACIS International Conference on SERA

Print and Generation Copy Image Watermarking Based on Spread Spectrum Technique

Ping Chen[1,2], Kagenori Nagao[2], Yao Zhao[3], and Jeng-Shyang Pan[4]

[1] Institute of Information Science, Beijing Jiaotong University, Beijing 100044 China
chpcn@126.com
[2] Fuji Xerox Co. Ltd
kagenori.nagao@fujixerox.co.jp
[3] Institute of Information Science, Beijing Jiaotong University, Beijing 100044 China
[4] National Kaohsiung University of Applied Sciences, Taiwan

Abstract. After a watermarked image is printed, copied several times and scanned, the watermark usually cannot be extracted rightly, as well as distorted by noise. In this paper, a spread spectrum technique was introduced to image watermarking which can be applied to print and generation copy. The watermarking message is modulated by the key-dependent pseudo-random sequences to produce a spread spectrum signal. The watermark is embedded into the cover image by modifying the gray value of each pixel. The cover image is not needed when the watermark is detected. The watermark is estimated exploiting the properties of HVS(Human Visual System) and extracted through the same key-dependent 2D sequences. Experimental results show that the scheme can resist print, generation copy and scan process.

1 Introduction

Watermarking techniques allow embedding information into digital media. The technologies are becoming more focus. An amount of algorithms have been proposed to hide data into multimedia [1]. But most of the commonly used watermark applications are referred to digital watermarking. While many important, long-lived documents are still available only on paper and most office activities still involve with paper documents. Current technology, however, has made some problems easier, such as leaking of information and counterfeit paper documents. These dishonest actions can be discouraged by enabling office equipment to check embedded watermarks of the documents.

Today the print, generation copy and scan processes are commonly used for image reproduction and distribution. It is popular to transform images between the digital format and the printed, generation copied format. Several articles can be found for printed-and-scanned image, e.g., [2]. But it didn't mention generation copy. This paper introduces spread spectrum technique[3] to image watermarking and this method is useful for the watermarking applications mentioned above, especially for printed and generation copied watermarked image. It can resist print, at least 2-generation copy and scan process[4]. In this method, watermarking message is modulated by the key-dependent pseudo-random sequences to produce a spread spectrum signal. The watermark is embedded into cover image in spatial domain exploiting the properties of HVS. The cover image is not needed when the watermark is detected.

After print, generation copy and scan process, the correlation between estimated watermark and original key-dependent 2D sequences is computed.

In Section 2, we describe the watermark embedding method. Watermarking detection process is given in Section 3. In Section 4, we give the experimental results and discuss the validity of the method and future work.

2   Embedding Process

Figure 1 shows the watermarking embedding system. First, the watermark message is encoded as a signal. Second, the signal is modulated by the key-dependent pseudorandom sequences to yield a spread spectrum signal. Finally, the modulated signal is added to the cover image exploiting the properties of HVS to obtain the watermarked image.
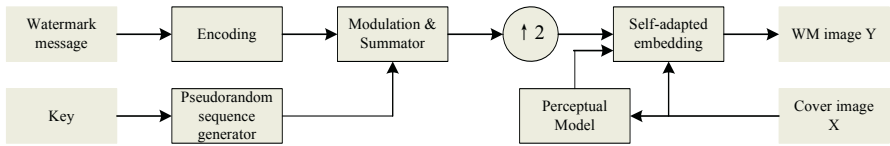
Fig. 1. Watermark embedding system

2.1   Watermark Generation

For getting a high correlation at watermark detection process, we generate watermark signal by spread spectrum technique. Suppose $b = \{b_i \mid b_i \in \{0,1\}, i = 1,2,\cdots L\}$ is a binary sequence that represents watermark information. We map each "0" bit of the sequence b to "-1", yielding a new binary polar sequence of $\{-1,1\}$, $b' = \{b_i' \mid b_i' \in \{-1,1\}, i = 1,2,\cdots L\}$. Afterward every b_i' is modulated by one sequence of modulation sequences set which were defined as $p = \{p_k(x,y) \mid k = 1,2,\cdots,L\}$, where $p_k(x,y)$ is a real number sequence according to Gaussian distribution $N(0,1)$ and it was generated by the key. In our experiments, the size of $p_k(x,y)$ is settled to 128×128 (a quarter of the cover image).

So the watermark signal will be

$$w(x,y) = \sum_{i=1}^{L} b_i' p_i(x,y) \tag{1}$$

After interpolate the watermark signal $w(x,y)$, we can get the final watermark W, which is the same size with the cover image.

2.2   Watermark Embedding

In order to make the watermark as robust as possible and avoid the visible distortion that are introduced by the watermarking process, we consider the scheme on the basis

of the utilization of luminance sensitivity function of the human visual system (HVS)[5]. The noise visibility function (NVF) related to the texture masking function is proposed by [6]. The most known form of the empirical NVF is widely used in image restoration applications [8]:

$$NVF \ (i, j) = \frac{1}{1 + \theta \sigma_x^2 (i, j)} \tag{2}$$

Where θ is a tuning parameter which must be chosen for every particular image. And to make θ image-dependent, it was proposed to use:

$$\theta = \frac{D}{\sigma_{x_{max}}^2} \tag{3}$$

Where $\sigma_{x_{max}}^2$ is the maximum local variance for given image and $D \in [50, 100]$ is an experimentally determined parameter.

Using the proposed content adaptive strategy, the final embedding equation is given [8]:

$$y(i, j) = x(i, j) + (1 - NVF \ (i, j)) \cdot w(i, j) \cdot \alpha \tag{4}$$

Where α denotes the watermark strength. $x(i, j)$ is cover image and $y(i, j)$ is watermarked image. The above rule embeds the watermark in highly textured areas and areas containing edges stronger than in the flat regions.

In the very flat regions, where NVF approaches 1, the strength of the embedded watermark approaches zero. As a consequence of this embedding rule, the watermark information is (nearly) lost in these areas. Therefore, to avoid this problem, we increase the watermark strength in these areas to a level below the visibility threshold by[8]:

$$y(i, j) = x(i, j) + (1 - NVF \ (i, j)) \cdot w(i, j) \cdot \alpha + NVF \ (i, j) \cdot w(i, j) \cdot \beta \tag{5}$$

Where β is about 3 for most of real world and computer generated images.

3   Detection Process

Watermark extraction system is shown in Figure 2.
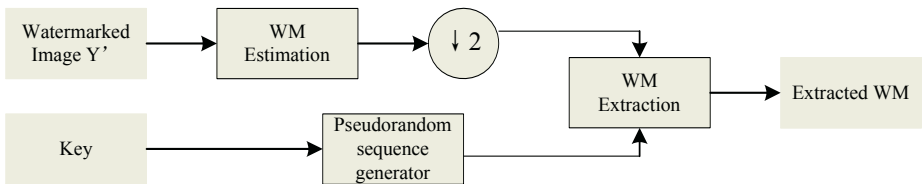
Fig. 2. Watermark extraction system

3.1   Watermark Estimation

According to the MAP criterion, the watermark can be estimated as [7]:

$$\hat{w} = \arg \max_{\widetilde{w} \in \Re^N} L(\widetilde{w} \mid y) \tag{6}$$

Suppose the cover image and the watermark message are independent identically distributed(i.i.d.), i.e. $x \sim N(\overline{x}, R_X)$, $w \sim N(0, R_W)$. Here, NVF is the output of the perceptual model so we can get

$$\hat{w} = \frac{R_W}{R_W + R_X}(y - \overline{y}) \tag{7}$$

Where we suppose $y \approx \overline{x}$. $\overline{y}$ denotes the local mean of the processed watermark image. $\hat{R}_X = \max(0, \hat{R}_Y - R_W)$ is the maximum likelihood(ML) estimation of the local image variance.

In our experiments, for the nice edge preserving properties we estimate local watermarked image mean y by median filtering in the 3×3 neighborhood around the corresponding pixel in the image. We replace the term $\frac{R_W}{R_W + R_X}$ with the value $NVF(y - \overline{y})$ and get a nice result. So watermark estimation as [8]:

$$\hat{w} = NVF(y - \overline{y}) \cdot (y - \overline{y}) \tag{8}$$

3.2 Watermark Extraction

Each watermark message bit detection is performed by calculating cross-correlation between watermarked image and the modulation sequences.

$$\begin{aligned} r_i &=< Y(x,y)*h, p_i(x,y) > \\ &=< W(x,y)*h, p_i(x,y) > + < X(x,y)*h, p_i(x,y) > \\ &=< w(x,y), p_i > + < X(x,y)*h, p_i(x,y) > \end{aligned} \tag{9}$$

Where h denoted the preprocess, including watermark estimation, interpolation and subsample. The first term of above formulation is watermark projection on modulation sequences, and the second term denoted original image projection on modulation sequences. In the above deduction, because p_i is a zero mean pseudo-random sequence, and it is independent of preprocessed image sequence $X(x, y)$, the term on the right tends to be zero. Due to the independency between different modulation sequences, we get:

$$\begin{aligned} r_i &=< w, p_i > \\ &=< \sum_{j=1}^{L} b'_j p_j, p_i > \\ &= b'_i \|p_i\|^2 \end{aligned} \tag{10}$$

And every watermark bit is extracted by the sign of r_i:

$$b_i = \begin{cases} 1 & if \ r_i \geq 0 \\ 0 & if \ r_i < 0 \end{cases} \tag{11}$$

4 Experimental Results

A watermark message of 64 bits is embedded into the cover image, whose size is 256×256 pixels. In our experiments, the parameters are $\alpha = 0.5, \beta = 0.2$. The

watermark signal $w(x, y)$ is shown in Figure 3(a) and the Figure 3(b) shows the watermark after modulation according to equation (5), that is, $(1-NVR(i,j))\cdot w(i,j)\cdot \alpha + NVR(i,j)\cdot w(i,j)\cdot \beta$
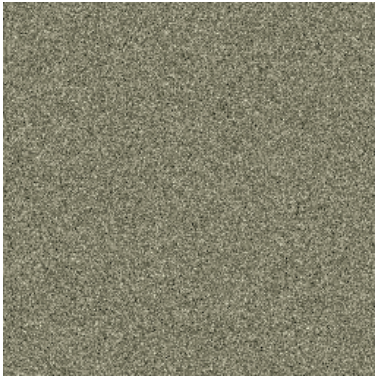
Fig. 3(a). Watermark signal
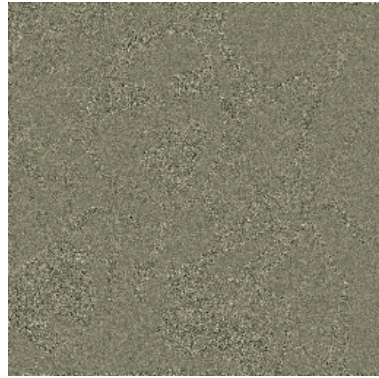
Fig. 3(b). watermark after modulation according to equation (5)

The original(cover) image and watermarked image are shown in Figure 4(a) and Figure 4(b).

Fig. 4(a). Original(cover) image

Fig. 4(b). Watermarked image

The printed watermarked image(0 generation copy) and generation copy watermarked images are shown in Figure 5(a) and Figure 5(b), 5(c), 5(d), 5(e), 5(f).

All of watermark bits can be detected by zero bit error in the 3rd generation copy. Table 1 shows the experiment result with the above original image.

In our experiments, the modulation sequences $p = \{ p_k(x, y) \mid k = 1, 2, \cdots, L \}$ are chosen from a normal distribution with mean zero, variance one and standard deviation one. And the watermark is embedded in the coordinate domain as a pseudo-random spatial spread modulated sequence. These explain why the watermark can be extracted from the printed and generation copied watermarked images.

Fig. 5(a). Printed watermarked image

Fig. 5(b). The 1st generation copy of the printed watermarked image

Fig. 5(c). The 2nd generation copy

Fig. 5(d). The 3rd generation copy

Fig. 5(e). The 4th generation copy

Fig. 5(f). The 5th generation copy

Table 1. Experiment results

Generation copy	Error bit in 64 bits	BER(%)
0 (only print)	0 bit	0
1st	0 bit	0
2nd	0 bit	0
3rd	0 bit	0
4th	1 bits	1.5625
5th	0 bits	0

Running the test we have successfully extracted the watermark message in the 3rd generation copy image. The test results also show that this method seems to be sensitive to the pixel dithering. We can extract all watermark bits on 5th generation copy but there is one bit error on 4th generation copy image. The possible reasons are the pixel dithering after the image is scanned and/or the error which is introduced when we search the border of the scanned watermarked image. It is, therefore, necessary to consider the enhancement of pixel synchronization and the effects of scanning. Based on the image pixel synchronization problem and the possible effects of scanning, we are going to develop the corresponding commercial schemes to solve it. In the future, we will apply the presented model also for commercial application.

References

1. M.D.Swanson, M.Kobayashi, and A.H.Tewfik: Multimedia Data-Embedding and Watermarking Techniques. *In Proc. of the IEEE*, vol.86, No. 6,Jun.1998, pp. 1064-1087.
2. C.-Y.Lin: Public Watermarking Surviving General Scaling and Cropping: An Application for Print-and-Scan Process. Multimedia and Security Workshop at ACM Multimedia 99, Orlando, FL, USA, Oct.1999
3. Ir.J.Meel (1999): Spread Spectrum (SS). [Online] Available http://fpzwireless.fpz.hr/dokumentacija/Spread_Spectrum.pdf
4. Ching-Yung Lin, Shih-Fu Chang: Distortion Modeling and Invariant Extraction for Digital Image Print-and-Scan Process. ISMIP99, Taipei, Taiwan, Dec. 1999
5. M.Kutter: Watermarking Resisting to Translation, Rotation and Scaling. *In Proc. of SPIE,* Boston, USA, Nov. 1998
6. S.Voloshynovskiy, A.Herrigel, N.Baumgaertner, and T.Pun. A Stochastic Approach to Content Adaptive Digital Image Watermarking. in Third International Workshop on Information Hiding, Dresden, Germany, 1999, Vol. LNCS 1768 of Lecture Notes in Computer Science, pp. 270-285
7. S.Efstratiadis, A.Katsaggelos. Adaptive Iterative Image Restoration with Reduced Computational Load. Optical Engineering, Vol.29, No.12, Dec. 1990, pp.1458-1468
8. Sang Maodong, Zhao Yao Digital image watermarking resisting to geometrical attacks. Journal of electronics & information technology, Vol.26, No.12, Dec. 2004, pp.1875-1881

Audio CoFIP (Contents Fingerprinting) Robust Against Collusion Attack

Kotaro Sonoda, Ryouichi Nishimura, and Yôiti Suzuki

Research Institutes of Electrical Communication
Graduate School of Information Sciences, Tohoku University
2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan
kotaro@ais.riec.tohoku.ac.jp
http://www.ais.riec.tohoku.ac.jp

Abstract. Digital fingerprinting method allows individualization of all content purchased by users. In a multi-cast network, however, distributed contents must be identical among all users. Therefore, the process of individualizing content should be given on the user side while the original signal should not be delivered to users for security reasons related to copyright protection. Content fingerprinting (CoFIP) for audio signals was proposed to meet both requirements. In CoFIP, a problem might occur when users collude. Synchronous addition of several individualized audio signals makes it difficult to identify each individualized contents composed in it. This article proposes a countermeasure against such collusion, providing imperceptible phase shift in the individualization. Listening tests and computer simulations show that the fingerprint can be detected correctly while the sound quality of the colluded signal is effectively degraded.

1   Audio CoFIP

Digital multimedia contents are now widely distributed through the Internet because of the progress of broadband networking technologies. As a result, management and protection of copyrights of such contents have become highly problematic. A possible solution to many such obstacles is fingerprinting, by which a unique fingerprint is stamped on each digital multimedia content. However, this is infeasible in the case of a non-interactive delivery system, e.g. a broadcast system, because broadcast data should be identical for all users.

"Content Fingerprinting (CoFIP)" has been proposed [1] to solve this problem. After CoFIP was originally proposed for still images, it has been implemented for audio signals [2]. In the original CoFIP algorithm, contents are distributed to users in the form of a package consisting of "parts" and a "body." "Parts" are a set of relatively small encrypted objects to be watermarked; the "body" is an unencrypted remaining object (Fig. 1(a)). Each "part" is stamped with several different watermarks. The key point of CoFIP is that all of these differently watermarked "parts" are packaged for distribution. Here, suppose that there are M "parts" and that each M "part" is stamped with one of N different watermarks. Because a distributed package contains $N \cdot M$ watermarked

(a) Original CoFIP for still images (b) Implementation of CoFIP for audio signals
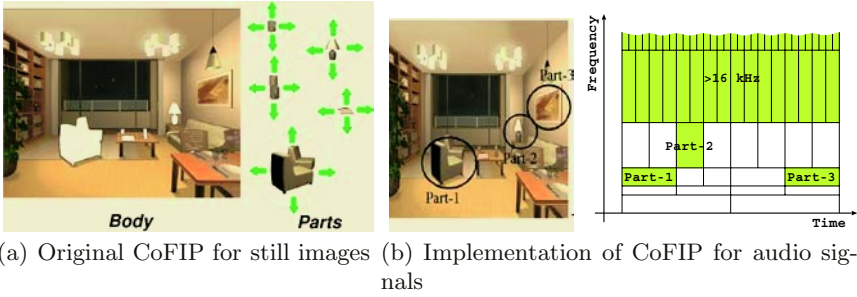
Fig. 1. Concept of individualization using CoFIP.

"parts," we reason that N different watermarks for M "parts" can generate N^M different fingerprints. Note that the "parts" are encrypted when the package is distributed. A user purchases a digital key to decrypt only one combination of the watermarked "parts" to generate individualized content by combining them with the "body."

In the audio CoFIP algorithm, the time-frequency representation of an audio signal is compared to a still image[2]. Figure 1(b) illustrates the comparison of the original CoFIP and that implemented for sound signals. We applied the wavelet transform for the time-frequency representation of audio signals. Prominent areas in the time-frequency representation that have large wavelet coefficients are chosen as "parts." In addition, the high-frequency region (> 16 kHz) is also chosen as a "part." These principles to choose the "parts" allow the fingerprints to survive even after some possible operations (attacks) such as band limitation and down-sampling have been undertaken.

The original audio CoFIP employed the time-spread echo watermarking method [3][4]. This method reduced the amplitude of echoes by spreading it over the time domain using a PN sequence. The robustness and secrecy were enhanced compared with the well known echo watermarking method. Figure 2(a) depicts an impulse response used to embed a watermark in the echo method. This impulse response is hereafter called a "kernel." Figure 2(b), on the other hand, depicts a kernel used in the time-spread echo method.

(a) Kernel of echo hiding   (b) Kernel of the spread-echo method
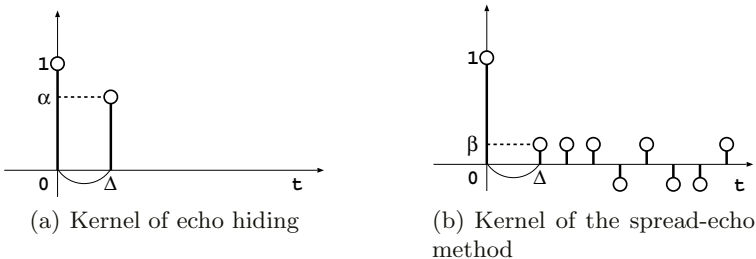
Fig. 2. Impulse response used to embed a watermark.

2   Problem and Solution

One drawback of the audio CoFIP is its vulnerability to collusion attacks. If various contents with different fingerprints are simply added after these are formally decrypted, detection of the watermarks may be quite difficult by the dilution (relative attenuation) of each watermark while the added contents are heard without any distortion. This article proposes a method to cope with this simple collusion attack by greatly degrading the sound quality after collusion.

2.1   Collusion Attack for Audio CoFIP

Time-spread echo method is realized by convolving a time-spread echo kernel depicted in Fig. 2(b) with an original signal. The impulse at $t = 0$ of this kernel produces the original signal. Other impulses, which are indeed a PN sequence, produce imperceptible echoes that correspond to fingerprints. Variations are produced by differing the PN sequences. Consequently, a watermarked signal $s'(t)$ consists of the original signal $s(t)$ and the spread echo signal $w(t)$. Using these notations, a simple collusion attack can be represented as

$$u(t) = \frac{1}{n} \sum_{i=1}^{n} s'(t) = \frac{1}{n} \sum_{i=1}^{n} (s_i(t) + w_i(t)) = s(t) + \frac{1}{n} \sum_{i=1}^{n} w_i(t), \qquad (1)$$

where $u(t)$ indicates a colluded signal. Consequently, each watermark $w_i(t)$ is reduced in its amplitude by $1/n$ and becomes an obstacle for detecting other watermarks. Because the original signal remains unchanged after this collusion attack, the attacker can obtain a signal from which no watermark can be actually detected.

2.2   Proposed Method

To overcome this problem, we propose a means to cause severe deterioration in the sound quality for a signal obtained by a collusion attack. This can be realized by modifying the phase of the signal relative to the original. This operation can be represented as

$$s_i'(t) = \mathrm{APF}_i[s(t)] + w_i(t), \qquad (2)$$

where APF[ ] indicates an all-pass filter with phase characteristics that differ for each i. Because the power spectrum is unchanged by Eq. (2), and the human auditory system is insensitive to phase, listeners are inferred to be incapable of discerning a signal obtained by this modification from the original signal. During a collusion attack, adding these signals contaminates the signal with distortion because its power spectrum is changed from the original by phase manipulation.
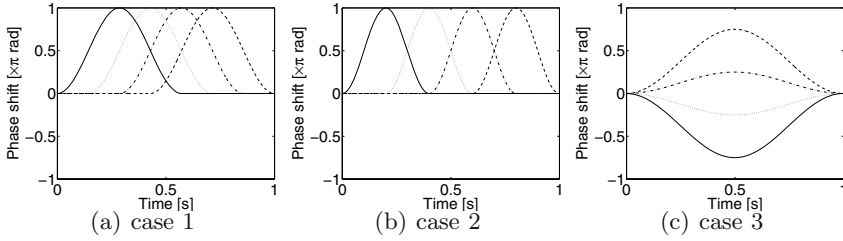
Fig. 3. Phase modification as a function of time for $N = 4$.

Although phase modification characteristics used in this method are arbitrary in a sense, three specific cases shown below are examined experimentally.

$$f(t,i) = \begin{cases} \pi \frac{1-\cos(2\pi \frac{(2N-1)(t-1)-(i-1)L}{NL})}{2} & \frac{(i-1)L+(2N-1)}{2N-1} \leq t \leq \frac{(i-1+N)L}{2N-1}, \\ 0 & 1 \leq t < \frac{(i-1)L+(2N-1)}{2N-1}, \quad \frac{(i-1+N)L}{2N-1} < t \leq L, \end{cases} \tag{3}$$

$$f(t,i) = \begin{cases} \pi \frac{1-\cos(2\pi \frac{(N+1)(t-1)-(i-1)L}{2L})}{2} & \frac{(i-1)L+(N+1)}{N+1} \leq t \leq \frac{(i+1)L+(N+1)}{N+1} \\ 0 & 1 \leq t < \frac{(i-1)L+(N+1)}{N+1}, \quad \frac{(i+1)L+(N+1)}{N+1} < t \leq L, \end{cases} \tag{4}$$

$$f(t,i) = \pi \frac{(2i-N-1)(1-\cos(2\pi t/L))}{2N} \qquad 1 \leq t \leq L, \tag{5}$$

where L and N respectively indicate the length of the original signal $s(t)$ and the number of variations. These functions for $i = 1,\ldots,4$ and $N = 4$ are depicted as a function of time in Fig. 3, where the abscissa is normalized by L.

In case 1 (Eq. (3)), periods where the phase of the original signal are modified differently among variations. This causes temporally localized phase distortion when the collusion attack is attempted. Case 2 (Eq. (4)) is similar to case 1 in the sense that periods where the phase of original signal are modified differently among variations. The difference between cases 1 and 2 is the amount of phase change when the collusion attack is delivered; the change becomes constant for the collusion attack in case 2, but it depends on the combination of the attack in case 1. Accordingly, in certain combinations in the collusion attack, sound distortion is expected to be less in case 2 than in case 1.

On the other hand, in case 3, the period where the phase of original signal is modified is common, but the amount of the phase shift differs among variations. In this case, as is apparent from Fig. 3, the temporal change in phase becomes mild, possibly resulting in less audible distortion than in the other two cases.

Table 1. Parameters set in listening tests.

# of watermarked objects M	1
# of variations N	2, 3, 5, 8
Time-spread echo method	
Delay Δ	56 / 44100 s
Gain β	0.005
Length of PN sequence	1023
Chip length	2

Table 2. Evaluation scale described in ITU-R BS.1116.

Scale	Description
5.0	Imperceptible
4.0	Perceptible, but not annoying
3.0	Slightly annoying
2.0	Annoying
1.0	Very Annoying

3   Evaluation by Listening Experiments

3.1   Experimental Conditions

Listening tests are carried out to evaluate the secrecy of the fingerprints in the proposed method. Computer simulations are also performed to evaluate the watermark detection performance. Moreover, sound quality deterioration as a result of the collusion attack is also evaluated by listening tests.

Sound sources used in the experiments are classical music and jazz pieces. These signals are excerpted from the RWC music database supplied by National Institute of Advanced Industrial Science and Technology, Japan (AIST) [5]. They are two-channel stereo samples with a sampling frequency of 44.1 kHz and a quantization resolution of 16 bits. The duration of samples used in the experiments is 20 s. Other parameters set in the experiments are summarized in Table 1.

3.2   Auditory Secrecy

Listening tests were carried out based on "Double-blind, triple-stimulus, with hidden reference" recommended in ITU-R BS.1116 [7]. In this test, subjects listened to three stimuli: 'A', 'B' and 'R.' 'R' is the original signal in all cases. 'A' and 'B' are either the original signal, the same with 'R' or a test signal, a watermarked signal in a random order. Those three signals are continuously output synchronously so that a listener can select and listen to an arbitrary stimulus among the three at any timing. The listener was asked to separately evaluate the sound quality degradation of both 'A' and 'B' relative to 'R' with respect to Table 2 in the precision of first decimal. Because either 'A' or 'B' is the same with 'R', a listener should give a rating of 5 to the stimulus judged to be the original. Listeners participating in this experiment were three well trained students with normal hearing ability in their twenties. Sound stimuli were presented through a headphone in a sound-proof room. Every set of 'A', 'B' and 'R' appeared five times in one session.

Results are depicted in Fig. 4. The notation of the abscissa indicates i over N in Eqs. (3)-(5). The ordinate indicates "diffgrade," which is calculated by subtracting the scale for the test signal from the scale for the hidden reference. Consequently, the lower the diffgrade is, the worse the sound quality. From Fig. 4,
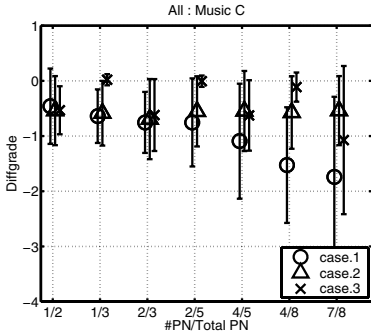
Fig. 4. Results of listening tests for secrecy of fingerprints.
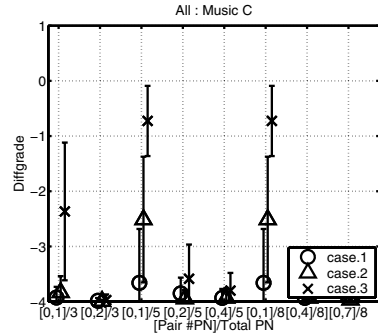
Fig. 5. Results of listening tests for some pairs of collusions.

it is clear that sound quality degradation caused by fingerprinting with the phase modification is not large.

3.3   Degradation of Sound Quality Against the Collusion Attack

Listening tests were carried out to examine sound quality degradation against simulated collusion attacks. Experimental conditions were identical to those in the previous experiments except for stimuli, i.e. the stimuli were signals obtained by simulated collusion attacks in the present experiment. Figure 5 shows the results. For simulating the collusion attack, eight kinds of combinations were tested, as shown in this figure. Only a collusion of two signals was examined. When the combination is represented as $(\xi, \eta)/\zeta$ in this figure, it means that from ζ different watermarked signals, a watermarked signal labeled number ξ and a watermarked signal labeled number η are summed and averaged. Comparison of Figs. 4 and 5 clarifies that the signals' sound quality after the collusion attack is remarkably deteriorated.

3.4   Watermark Detection

Computer simulations were carried out to evaluate the influence of phase modification on watermark detection. In the computer simulations, both the hit (positive) rate and the false alarm (false positive) rate were calculated. A method using log scaling along time-axis proposed by Ko *et al.* [6] was employed to detect watermarks. This method makes it possible to detect watermarks under attacks such as scaling or time-scaling, which are fatal attacks for the original time-spread echo method, in return for the worse detection performance in the case of no attack. Figures 6 depict results for a classical music piece without any attacks. The notation of the abscissa indicates i over N in Eqs. (3)-(5). From this figure, obtained hit rate scores are around 50% while the false alarm rate scores are less than 10%. Because one watermark is embedded several times into a music signal to allow watermark detection, these scores are practically realistic.

(a) Hit rate
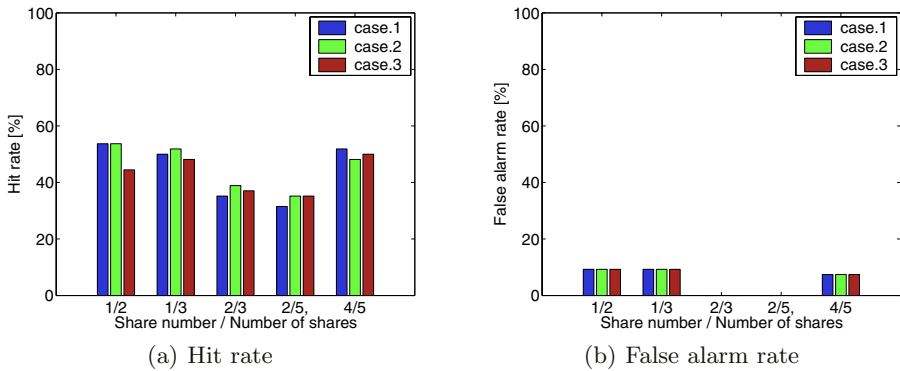
(b) False alarm rate

Fig. 6. Hit rate and false alarm rate for a classical music piece.

4   Conclusions

This article presented a method to deal with the collusion attack in audio CoFIP. The proposed method deteriorates sound quality during a collusion attack.Computer simulations revealed that watermarks remain detectable after phase modification in the proposed method. Moreover, listening tests demonstrated that a signal obtained by the assumed collusion attack shows remarkably deteriorated sound quality when compared with the original signal.

References

1. Takahashi, Y., Aoki, T., Yasuda, H.: A study on copyright protection based on CoFIP. Forum on Information Technology 2002 **N-19** (2002) (in Japanese)
2. Sonoda, K., Nishimura, R., Suzuki, Y.: Implementation of CoFIP for music signals. International Congress on Acoustics (2004) 1011–1014
3. Ko, B.-S., Nishimura, R., Suzuki, Y.: Time-spread echo method for digital audio watermarking using PN sequences. ICASSP **II** (2002) 2001–2004
4. Ko, B.-S., Nishimura, R., Suzuki, Y.: Time-spread echo method for digital audio watermarking. IEEE Trans. on Multimedia **7** (2005) 212–221
5. [web] http://staff.aist.go.jp/m.goto/RWC-WDB/
6. Ko, B.-S., Nishimura, R., Suzuki, Y.: Log-scaling watermark detection in digital audio watermarking. ICASSP **III** (2004) 81–84
7. Methods for the subjective assessment of small impairment in audio systems including multichannel sound systems. ITU-R BS.1116-1

A Low Cost and Efficient
Sign Language Video Transmission System

Mohsen Ashourian[1], Reza Enteshari[1], and Ioannis Lambadaris[2]

[1] Islamic Azad University of Iran, Majlesi Branch,
P.O.Box 86315-111, Isfahan, Iran
`mohsena@iaumajlesi.ac.ir`
[2] Department of System and Computer Engineering,
1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6, Canada
`ioannis@sce.carleton.ca`

Abstract. In this paper we propose a video in video hiding technique for embedding and transmission of sign language video signal through another video communication channel. Our proposed system does not need modification of the main video coding and transmission system, and it can be implemented by addition of a pre- and post-processing block. We encode the sign language video and embed the encoded bit-stream in the main host video. We develop a low complexity and robust encoding scheme for sign language video by applying three-dimensional wavelet decomposition on the sign language video, and encoding the important subbands using multiple description scalar quantizers. We also apply three-dimensional wavelet decomposition on the host video and embed the encoded bit-stream of the sign language video in the host signal subbands. We select blocks of the host video with high texture content and embed the information in the middle frequency subbands to have highest robustness and with low visible distortion. At the receiver the system reconstructs the sign language video by recovering the embedded data in subbands. We examine the system performance by embedding video sequences of sign language messages into host video sequences. At the receiver we evaluate the percentage of correct sign language messages recovery. We further evaluate the system performance when the host video undergone compression and addition of noise during transmission.

1   Introduction

Sign language video transmission offers great possibilities for better communication among deaf people.However, transmission of sign language video at low bit-rate and with low cost is still a challenging area. Conventional video conferencing technology generally addresses the limitation of channel capacity by drastically reducing the frame rate, while preserving image quality. This produces a jerky image that disturbs the trajectories of the hands and arms which are essential in recognizing sign language.

One potential demand is addition of a simple sign language video transmission system to an already established multimedia transmission system. In our

application, we assume that there is an established high quality digital video transmission system, and we would like to use it for transmission of a sign language video without interrupting the main transmission system. As depicted in Figure 1 , we propose a sign language transmission system by encoding the sign language video and embedding the encoded bit-stream in another digital video signal.
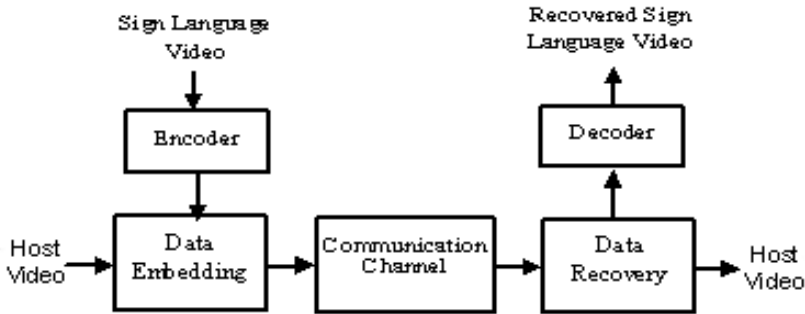
Fig. 1. Overview of the Proposed System.

Embedding a multimedia signal into another multimedia signal has been popularly used in digital watermarking and data hiding applications [2]. In data hiding applications, we need to recover the embedded information with high quality. However; in these applications, unlike watermarking systems, the host signal usually does not face active and severe attacks by unauthorized people for destroying the embedded data. The host multimedia data may only face signal processing operations, such as compression and addition of noise during transmission.

The main problem of hiding video in video is the large amount of data that requires a special data embedding method with high capacity as well as transparency and robustness. There have been few reports on large capacity data embedding [3], [4], [5]. Chae and Manjunath used the discrete wavelet transform (DWT) and lattice code for embedding a signature image/video into another image/video [3]. They further improved their system by using joint source-channel quantizers. However; the channel-optimized quantizer is not suitable in data hiding applications, where intentional or non-intentional manipulations are variable and not known in advance. In another approach Swanson et. al. [5] designed a method for embedding video in video based on linear projection in the spatial domain. The method is flexible and simple to implement, but like other spatial domain embedding techniques, it is not robust to compression.

In [7] we propose an image data hiding scheme that can hide a gray-scale image in another gray scale image with bigger size. The key advantage of the developed system is using multiple description coding which is a joint source-channel coding method [8]. In this paper, we extend the image hiding scheme for embedding a video with lower resolution (sign language video) in another video with higher resolution (host video). In the following sections, at first we explain

sign language encoding scheme in Section 2. In Section 3 we explain the method of hiding and extraction of video. In Section 4 we provide the experimental results, and finally in Section 5 we give conclusion and suggestion for further research.

2 Sign Language Encoding

Many video conferencing systems currently installed are not optimized for sign language transmission. Accurate sign language communication places specific demands on a visual media application. The sign language transmission needs special spatial/temporal resolution. In sign language small details may carry important information, like the eye gaze direction, used for example to indicate references to persons or things recently mentioned. Also rapid movements can carry important meaning, like finger-spelling, where a word or name is spelled out by the hand forming specific positions for each letter [9]. Figure 2 shows the results suggested in ITU standard [9] for sign language and lip reading real time conversation using low bit rate video communication.
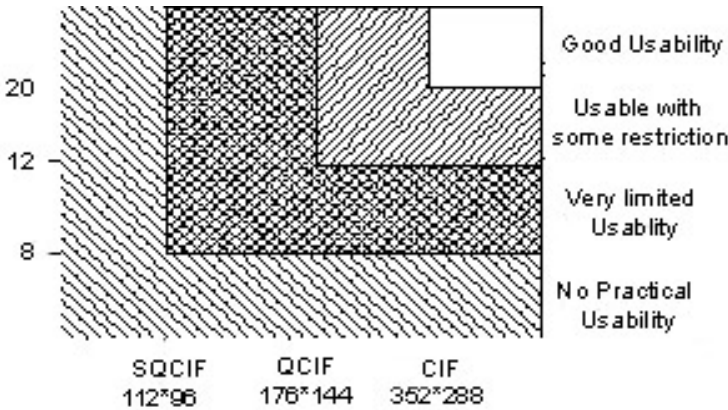
Fig. 2. Resolution requirements for sign language and lip-reading in person-to-person conversation.

We develop the sign language transmission system using QCIF format 176×144 with 15 frames per second. The host video has CIF format 352×288 with 30 frames per second. International video compression standards like MPEG-4, H.263 and the newly emerged H.264 standard describe mechanisms for coding, multiplexing and presenting video images at low bit rates. In general these coders use DCT coding for inter-frame compression and motion estimation/compensation for intra-frame compression. For further efficient sign language coding, it possible to identify visually important regions of video for deaf communications, and use object based coding methods.

In our proposed application we embed the bit-stream of encoded video in another host video that might face various signal processing operations like compression and addition of noise. Therefore, we need to use a robust video coding scheme. The hybrid coders have high encoding efficiency, but they are susceptible to noise due to using variable length code, and using motion estimation/compensation that increases the possibility of error propagation in group of frames. With knowledge of communication channel, it is possible to use various channel coding schemes for protection of the encoded bit-stream in a hybrid coder. However, in data hiding application usually there is not enough information about the type of processing that host signal faces, and the effect of this operation on the embedded information can not be easily estimated. Therefore, we preferred to use a robust source coding scheme, despite its lower coding efficiency.

We use a three-dimensional (3-D) subband coder with fixed rate quantizers [6]. Splitting the spatio-frequency information among subbands, and the fixed rate quantizers increase the encoder resistance to loss of information that can be happened due to various processing on carrier (or host) video.
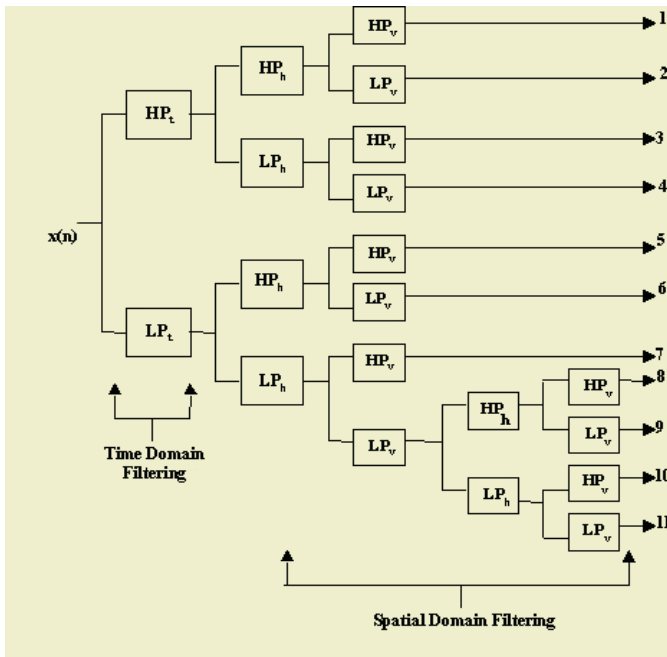
Fig. 3. Three-Dimensional Filter Bank Structure.

Figure 3 shows the structure of 3-D subband decomposition used for both video signals. A three-dimensional subband coder uses a unique approach for encoding intra-frame and inter-frame redundancy in a video sequence. The video signal passed through a 3-D filter bank, and then different subbands are encoded
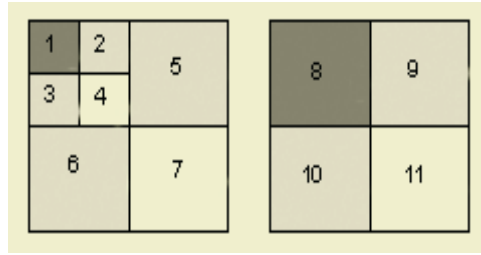
Fig. 4. Three-Dimensional Filter Bank Frequency Map.

based on their visual importance [6]. The terms HP and LP refer to high-pass filtering and low-pass filtering, where the subscripts t, h, and v refer to temporal, horizontal, and vertical filtering respectively. The selected subband framework consists of 11 spatio-temporal frequency bands. The temporal frequency decomposition is restricted to only two subbands due to potential delay problems in a practical implementation and reducing dependency in coding consecutive frames. The image frames are filtered temporally using the two-tap Harr basis functions [6]. Temporal decomposition is followed by horizontal-spatial filtering and vertical-spatial filtering using 9/7 biorthogonal filters [6].

Selection of optimum quantizer for different subbands based on their statistical characteristics and visual importance is the key factor for developing subband coder. Figure 4 shows the frequency map of the 11 video subbands which can be classified as below:

Band 1, the low temporal and spatial frequency band, is a blurred version of the original video frame. It has much higher energy compared to other subbands and has the most visual importance. However while all the subbands histogram follow well a generalized Gaussian distribution, this subband does not follow any fixed distribution [6].

Bands 2-7, the low temporal and high spatial frequency bands, include information of texture and sharpness of video signal in the spatial domain. Depends on amount of these information in scene, the energy of these bands could be higher or lower. Among these bands, bands 4 and 7 have much lower due to two times highpass filtering (vertical and horizontal).

Band 8, the high temporal and low spatial frequency band, has higher average energy compared to other high temporal bands, and it shows the major changes in consecutive video frames.

Bands 9-11, the high temporal and high spatial frequency bands (Bands 9-11) have low energy, but high variation in time. They represent sharp and fast objects movements in the video scene.

The amplitude distribution of Bands 1 and 8 does not follow any fixed probability distribution function (PDF) [6]. We use phase scrambling operation to change the PDF of these bands to a nearly Gaussian shape [6]. The added random phase could be an additional secret key between the transmitter and the registered receiver. Since the information of the sign language video signal are embedded in another video signal that might face various types of signal processing operation, we need to protect the key subbands properly. We use multiple

description PDF-optimized scalar quantizer for Bands 1 and 8, as they follow well a Gaussian distribution after phase-scrambling. Multiple Description Coding (MDC) is a joint source-channel coding technique where the source is encoded by multiple descriptions and there is some redundancy among descriptions [6]. These descriptions are transmitted over separate channels for error protection. Figure 5 shows the block diagram of MDC. At the receiver, the multiple description decoder combines the information of the two descriptions and reconstructs the original signal. It can decode only one channel, when data on the other channel is highly corrupted; otherwise it can combine the received information from both channels.
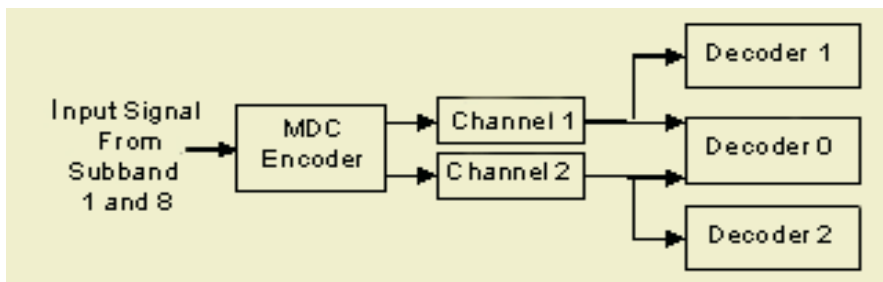
Fig. 5. Multiple description coding of subband 1 and 8.

The other high frequency subbands contain the texture information of the video scene which are not visually important in recognizing the sign language video. We drops Bands 4,7 and 11, as they have very low energy and only quantize the other six subbands (2,3,5,6,8,9) with PDF-optimized quantizer assuming Laplacian PDF for them. We split the quantizers indices into two groups and embed them in different spatio/temporal position or frequency bands of the host video to have higher protection against compression and noise addition. The splitting and scrambling of the encoded bit-stream reduce the effect of loss of information on the reconstructed video.

3   Data Embedding and Extraction

Various types of transform domain methods were suggested for data embedding in image and video [2]. The DCT and wavelet domains were more popularly selected due to their compatibility with popular compression methods.

We embed the sign language video with QCIF frame size of (176×144) pixel and rate of 15 frames per second into the host video with CIF format $(352 \times 288$ pixels and 30 frames per second). The sign language video is monochrome, but the host video is a colored video in YUV (4:2:0) format. We decompose the Y component of the host video using the same 3-D subband decomposition depicted in Figure 3. The chorma components are decomposed only in 8 subbands as depicted in Figure 6. The high temporal subbands usually have low energy, and they can be easily affected by various types of compression schemes, therefore,

Fig. 6. 3-Dimensional filter bank frequency map for chroma components.

we embed the date only in the medium frequency and low temporal subbands. The selected subbands are: 2, 3, 5 and 6 of luminance signal and bands 2, 3, 6 and 7 of chrominance signal.

The host video is in 4:2:0 format, it means for each group of four luminance pixels, there is only one pixel with two chrominance information. However the visual system is less sensitive to color distortion. We split and distribute the encoded bit-stream among the selected host video subbands blocks. Since the sign language video has much lower tempo/spatial resolution compared to the host video (1/6 in spatial and 1/2 temporal resolution), therefore we have enough flexibility in distributing the embedded bits.

We embed the bit-stream of signature information in the host video frames in the area with high texture content to be less visible. In order to evaluate the texture content of a 4×4 block, we define a normalized measure for the energy of high frequency bands by

$$\mu_k = \left| \frac{e_H}{e_L} \right| \tag{1}$$

where e_H is the average of the absolute value of the high frequency bands (2,3,5,6), and e_L is the absolute value of the lowest frequency band (Band 1) of the corresponding block. characterizes the given block texture energy. Higher value of μ_k shows the block has strong high frequency component or high texture. We consider these blocks good candidates for data embedding, and replace wavelet coefficients with embedded data [3].

We split and distribute the encoded bit-stream among the selected host video subbands blocks. Since the sign language video has much lower tempo/spatial resolution compared to the host video (1/6 in spatial and 1/2 temporal resolution), therefore we have enough flexibility in distributing the embedded bits. We embed the significant bits of quantization indices in higher frequency subbands and less significant bits in lower frequency subbands (Bands 5 and 6). This will increase the robustness of our data embedding scheme to compression methods which are usually affecting high frequency bands.

At the receiver, we first reconstruct the two important frequency subbands (Band 1 and 8) of the sign language video from the extracted indices in each portion of the host video, and recombine the indices. If the two indices are close together we can recombine them and use the multiple description scalar quantizer table to have higher resolution [8]. On the other hand if the two indices were far from each other, we assume that one of them is highly corrupted. In order to choose the less corrupted index, we compare the difference between the block containing those indices with former blocks in time-domain.

4   Experimental Results and Analysis

In our developed system, we do not want the host video to face any visual distortion due to data embedding. At the same time, we would like to recover sign language video with reasonable quality. In all of our experiments we set the embedding factor so that average PSNR of the host video sequence stays above 35 dB.

There are various ways to evaluate the recovered sign language video depends on the application [9]. We produce a sign language video sequence that speaker tells 40 independent sentences, and we embed the sign language interpretation of the sentence in a video sequence. Ten deaf persons, familiar with sign language, independently interpret each sentence after data recovery. We report the average percentage of correct sentence recovery (PCSR). In the normal situation that host video does not face any modification or signal processing operation, the PCSR is above 95. We test the system resistance to MPEG compression. The MPEG-2 compression scheme with various compression ratios (CR) is tested. Table 1 shows the percentage of correct sentence recovery (PCSR) for each case. In another experiments, additive noise with various standard deviations was added to the video sequences. The PCSR of recovered video based on PSNR of the host video are shown in Table 2.

Table 1. The PCSR after MPEG compression of the host video.

CR	3	6	12	24
PCSR	92.9	87.2	81.8	74.4

Table 2. The PCSR after addition of noise to the host video.

PSNR(db)	35	30	27	25
PCSR	94.2	88.0	75.6	69.4

Table 3. The PCSR down-sampling the host video.

Down-Sampling	1/2	1/4	1/8
PCSR	85.3	73.0	59.5

5   Conclusion

We have presented a new scheme for embedding sign language video sequence into a host video sequence. We used 3-D subband decomposition and multiple description coding of the key information for robust encoding of the sign language video. We also use subband decomposition for the host video and embed the information of the sign language signal with very low visible distortion in the

host video. The results show that the system is able to transmit the sign language video even when the host video face operations like compression, addition of noise and down-sampling. The system has very low complexity and it can be used for transmission of sign language video in an already established video communication system.

References

1. Letellier P. , Nadler M. and Abramatic J.F.: The telesign project. Proceeding of IEEE , vol.37, (**1985**)813–827
2. Petitcolas, F.A.P., Anderson R.J., and Kuhn, M.G.: Information Hiding-a Survey, Proceedings of the IEEE, Vol.87, No.7, (**1999**)1062–1078
3. Chae J.J. Manjunath, M.J. :A Technique for Image Data Hiding and Reconstruction without Host Image. in Proc. of SPIE, Security and Watermarking of Multimedia Contents, San Jose, California, Vol. 3657, San Jose, California,(**1999**)386–396
4. Mukherjee, J.J., et. al.: A Source and Channel-Coding Framework for Vector-Based Data Hiding in Video. IEEE Trans. on Circuits and System for Video Technology, Vol.10, No.6,(**2000**)630–645
5. Swanson, M.D., et. al.: Data Hiding for Video-in-Video. IEEE International Conference on Image Processing, Vol.2, (**1997**)676–679
6. Ashourian, M.: Embedding Sign Language Interpretation of a Video in the Wavelet Transform Domain. European Signal Processing Conference, (**2004**)469–473
7. Ashourian, M., Ho, Y.S.: Multiple Description Coding for Image Data Hiding in the Spatial Domain. Lecture Notes in Computer Science (LNCS), Vol. 2869, (**2003**)659–666
8. Vaishampayan, V.A.: Design of Multiple Description Scalar Quantizers. IEEE Trans. on Information Theory, Vol.39 , No.5, (**1993**)821–834
9. ITU. Supplement 1 to Series H - Application Profile - Sign language and lip-reading real-time conversation using low bit-rate video communication, (**1999**)821–834

A New Watermark Surviving After Re-shooting the Images Displayed on a Screen

Seiichi Gohshi[1], Haruyuki Nakamura[1], Hiroshi Ito[2], Ryousuke Fujii[2],
Mitsuyoshi Suzuki[2], Shigenori Takai[3], and Yukari Tani[3]

[1] NHK, 1-10-11 Kinuta Setegaya-ku Tokyo
[2] Mitsubishi Electric Corporation, Information Technology R&D Center,
Ofuna, Kamakura, Kanagawa
[3] Mitsubishi Electric Engineering Corporation,
Kamimachiya, Kamakura, Kanagawa, Japan

Abstract. Piracy is a major issue in the content business, and although digital technologies raise the efficiency of new content production, pirates also accelerate their piracy. Commercial cameras have become sufficiently good to re-shoot images displayed on screens, and TV monitors are becoming larger with better picture quality; this combination will enable content to be pirated in the home in the near future. Re-shooting with a commercial camera is the ultimate piracy method because there is no engineering way to stop it. We have developed a new watermark method that remains even after images displayed on a monitor have been re-shot. This paper describes the concept and experimental results.

1 Introduction

Digital technologies have advanced remarkably in the last decade, and most contents are now made and processed digitally, offering many methods to create new contents. However, the technologies also enable contents to be copied illegally and there is no difference between the originals and copies in the market. There are several ways of protecting contents against piracy, such as Digital Rights Management (DRM) that encrypts the content.

Nevertheless, there is no way to protect content against piracy using commercial video cameras to re-shoot screens or TV monitors, and these cameras are becoming more powerful. It is well known that some people go into movie theaters to video the showing, make copies and sell them. Indeed, the Hollywood film industry estimates that such illegal copying costs the industry US $3 billion a year[1].

In Japan, some sellers of these illegal copies have been arrested[2]. These crimes did not exist when analogue copy systems were used, but TV monitors in the home have become larger than those used ten years ago.

It will soon be possible to pirate pay-per-view TV programs in the home by re-shooting the TV screen, so some means of protection is urgently needed. DRM was developed to prevent contents from being seen by those other than

rights holders, but this is an exception because the pirates pay for tickets or are regular subscribers. DRM cannot protect contents against regular members.

Watermarking is a useful means of content protection. Although it is fragile and does not have any power by itself, it adheres to any copies that are made. Watermarks were developed to conceal information in contents that are copied digitally or by analog means, but most of them cannot prevent re-shooting images displayed on a screen or monitor. Only a few ideas were proposed for it[3][4] and these embed information on the time-axis.

Most manufacturers and developers do not reveal the algorithms used in their watermark technologies, since it would be easy to crack them if the algorithms were open. However, this prevents them from standardizing.

Macrovision technology[5] is very useful in terms of protecting against casual copying. Its algorithm is open, and some machines are sold that can delete the Macrovision signal from the contents. However, it is used all over the world to protect contents. We think that we will need some new technology to protect contents from duplication by re-shooting, which is a typical method of casual copying, since virtually anyone will be able to do it.

Learning from the experience of the Macrovision technology, we can open the algorithm if we use it just to protect against casual copying. Opening an algorithm enables standardization. The requirements for such watermarks are not the same as those today. Indeed, such requirements have yet to be discussed. As we are proposing a new watermark technology that is robust against re-shooting in this paper, we would like to initiate discussion on its requirements.

In this paper we propose a new watermark technology which remains even after re-shooting an image displayed on a monitor. Our idea is that using very low frequency compared with moving images and that embedding information on the special axis. The rest of this paper is organized as follows. Section 2 explains the picture quality of a commercial camera; it is remarkable that re-shot contents is of marketable quality. In section 3, we explain the concept of our watermark. The experiential conditions and results are shown in section 4. Finally, section 5 concludes the paper.

2   Characteristics of Commercial Cameras

As only broadcasters and a few people had good cameras thirty years ago, people could not use them to copy contents by re-shooting. However, the performance of commercial cameras has dramatically improved in the last ten years, and they are now small enough to use anywhere and are not expensive. TV monitors including PDP and LCD displays now also offer good picture quality. High-performance cameras and high-definition displays provide an excellent combination for piracy.

We conducted experiments to measure the characteristics of re-shot pictures with multi-burst signal. Fig. 1 shows the original multi-burst signal measured with an oscilloscope. The same signal is given to a CRT monitor and is re-shot with a commercial camera. Fig. 2 shows the re-shot one measured with the same oscilloscope. The multi-burst signal frequencies are 0.5MHz, 1MHz, 2MHz, 3MHz,4MHz and 5.75MHz, respectively.
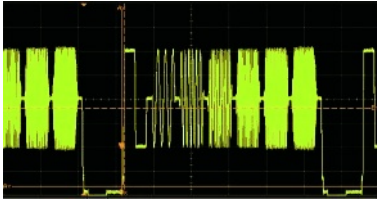
Fig. 1. Original Multi-burst Signal.
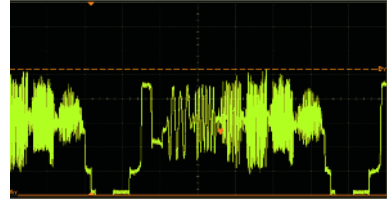
Fig. 2. Re-shot Multi-burst Signal.

The picture quality of the re-shot one is not the same as the original one. We will show the picture quality of normal moving images in section 4. We exhibited re-shot moving images in our Open House Exhibition in 2002. More than twenty thousand people came to our laboratories at that time. None of them said that they were bad and most of them said that it was worth paying.

This result enables piracy in movie theaters, but in the near future pirates will be able to copy contents in their own homes. Therefore, new technologies to protect contents against such piracy are needed. As mentioned in section 1, DRM and encryption systems do not offer sufficient security against these illegal copies, and so this is a pirate's paradise.

We propose a new watermark technology that survives after re-shooting the images displayed on a monitor. This approach is also useful for preventing piracy in movie theaters.

3   The Concept

3.1   Watermark Embedding

We assume video images $f(x, y, t)$ as three-dimensional signals, where x denotes the horizontal axis of the screen or monitor, y the vertical axis, and t the time axis. We also assume $g(x, y, t)$ as hiding information in the moving images, and x, y, t are the same as above. $g(x, y, t)$ must satisfy the following conditions.

- $g(x, y, t)$ has very low frequency signals compared with moving images $f(x, y, t)$
- $g(x, y, t)$ must satisfy the following equations:

$$g(x, y, t) = g_1(x, y)g_2(t) \tag{1}$$

$$\int_0^B \int_0^A g_1(x, y)dxdy = 0 \tag{2}$$

where, A denotes the width of the image and B the height of the image. These conditions are required to get equation(8) from equation(7). Examples of $g(x, y)$ functions are $\sin x$, wavelet functions and so on.

The moving picture which has embedded information can be written as follows:

$$f(x, y, t) + g(x, y, t) \tag{3}$$

where, $|g(x,y,t)| \leq 2$. In our approach, the watermark is embedded in only luminance signals. When $|g(x,y,t)| \geq 3$, degradation caused by the watermark becomes visible in some areas of images, which will be discussed in section 4.

3.2 Watermark Detection

Moving images $f(x,y,t)$ can be expressed as follows:

$$f(x,y,t) = f_{DC}(t) + f_{AC}(x,y,t) \tag{4}$$

where, $f_{DC}(t)$ denotes DC elements of images and $f_{AC}(x,y,t)$ denotes AC elements. This expression is used to derive equation(8) by equation(6) and equation(7).

We obtain the following expression from equation (3) for watermark embedded moving images:

$$f_{DC}(t) + f_{AC}(x,y,t) + g(x,y,t) \tag{5}$$

where, $f_{DC}(t)$ denotes the DC value of time t. We assume one more condition: that $f_{AC}(x,y,t)$ satisfies the following equation if the period of integration is sufficiently long:

$$\int_{-t_0}^{t_0} f_{AC}(x,y,t)g_2(t)dt \approx 0 \tag{6}$$

where, $-t_0$ and t_0 denote the period of time. Although there is not $g_2(t)$ that satisfies equation(6) for all moving images, we can choose several $g_2(t)$ candidates depending on images and we embed watermark only into images that satisfies this condition. According to our experiments, many moving images satisfy eqatuion(6) if we choose appropriate $g_2(t)$ and take the appropriate integral period. We need further research about it.

Here, we obtain the product of equation (5) and $g(x,y,t)$, and integrate them. The integral periods must be sufficiently long to satisfy equation (2) and (6).

$$\int_{-t_0}^{t_0} \int_0^B \int_0^A (f_{DC}(t) + f_{AC}(x,y,t) + g(x,y,t))g(x,y,t))dxdydt$$

$$= \int_{-t_0}^{t_0} f_{DC}(t)g_2(t)dt \int_0^B \int_0^A g_1(x,y)dxdy$$

$$+ \int_0^B \int_0^A g_1(x,y)dxdy \int_{-t_0}^{t_0} f_{AC}(x,y,t)g_2(t)dt$$

$$+ \int_{-t_0}^{t_0} \int_0^B \int_0^A g^2(x,y,t)dxdydt \tag{7}$$

Then, the first term and the second term of equation (7) equal zero; only the third term has some limited value as follows:

$$\int_{-t_0}^{t_0} \int_0^B \int_0^A g^2(x,y,t)dxdydt \neq 0 \tag{8}$$

If we choose an appropriate threshold value, we will be able to judge whether the moving images contain embedded information $g(x, y, t)$ or not.

In this paper we use a very simple two dimensional $g(x, y)$. Three dimensional $g(x, y, t)$ will be a further research. Here, we use a very simple function $g(x, y) = \sin x$ to explain a practical example. When we embed some information into moving images, $\sin x$ is used to embed $+1$ and $-\sin x$ for 0. The integration period of x should be from 0 to 2π. The third term of equation (7) becomes as follows when $\sin x$ is embedded:

$$\int_0^{2\pi} \sin x \sin x dx = \pi \tag{9}$$

If $-\sin x$ is embedded, The third term of equation (7) becomes as follows:

$$\int_0^{2\pi} (-\sin x) \sin x dx = -\pi \tag{10}$$

We can thus detect embedded information by these methods. Although there are several assumed conditions, we will give proofs that are valid in normal moving images in the next section.

4   Experimental Results

Several experiments were made to examine the validity of our approach. Information is embedded into moving images with the watermark explained in the previous section. The embedded moving images are stored on the computer hard disk, which is shown as the signal source in Fig. 3. They are supplied to the SDTV monitor and are re-shot by the camera. The camera video output is connected to the watermark detector.
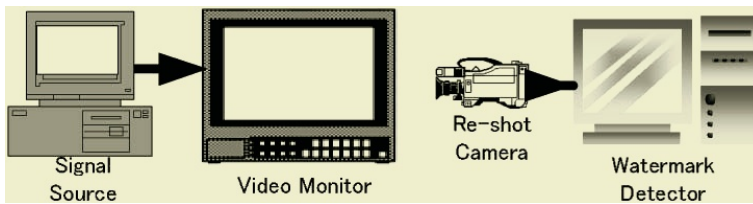
Fig. 3. Experimental System.

Information is embedded before the experiment in non-real time. We developed a real-time watermark detector with a computer equipped with a Pentium 4, 2.8 GHz CPU. We have to select $g(x, y, t)$ which satisfies equation (2). Although there are many candidates, we select three two-dimensional functions to investigate the basic characteristics of moving images.

$$g_1(x, y) = -\alpha \cos(\frac{2\pi x}{A}) + \alpha \cos(\frac{2\pi y}{B}) \tag{11}$$

$$g_1(x, y) = \alpha \sin(\frac{2\pi x}{A}) + \alpha \cos(\frac{2\pi y}{B}) \tag{12}$$

$$g_1(x, y) = \alpha \sin(\frac{2\pi x}{A}) \times \cos(\frac{2\pi y}{B}) \tag{13}$$

where, A denotes the width of the monitor and B denotes the height. The co-ordinates are defined as Fig. 4. x denotes the x axis and y denotes the y axis. α denotes the intensity of the embedded watermark. For example, if we assume $g_1(x, y) = -\alpha \cos(\frac{2\pi x}{A})$ and $\alpha = 2$, $f(x, y, t)$ is changed by $f(x, y, t) + 1$ while $\cos(\frac{2\pi x}{A}) \geq 0.5$ and is changed $f(x, y, t) - 1$ while $\cos(\frac{2\pi x}{A}) \leq -0.5$. If we want to embed binary information 1, α should be plus, i.e. $\alpha = +1$. If we want to embed binary information 0, α should be minus, i.e. $\alpha = -1$.

We used ITE (Institute of Image Information and Television Engineers, Japan) standard images and MPEG images as shown in Table 1. We edited 19 videos in series. Each image has 450 frames. We embedded two bits of information in every 450 frames.

Table 1. Images used in experiments.

Picture names	Source	Frame number	Picture names	Source	Frame number
Woman with Birdcage	ITE	$1 \sim 450$	Buddhist Images	ITE	$451 \sim 900$
Buildings along a Canal	ITE	$901 \sim 1350$	Character Pattern	ITE	$1351 \sim 1800$
Chromakey (flowers)	ITE	$1801 \sim 2250$	Church	ITE	$2251 \sim 2700$
Flash Photography	ITE	$2701 \sim 3150$	Ice Hockey	ITE	$3151 \sim 3600$
Horse Race	ITE	$3601 \sim 4050$	Group Game	ITE	$4051 \sim 4500$
Overlap	ITE	$4501 \sim 4950$	Skyscrapers	ITE	$4951 \sim 5400$
Whale Show	ITE	$5401 \sim 5850$	Bicycles	MPEG	$5851 \sim 6300$
Cheerleaders	MPEG	$6301 \sim 6750$	Flower Garden	MPEG	$6751 \sim 7200$
Football	MPEG	$7201 \sim 7650$	Carousel	MPEG	$7651 \sim 8100$
Susie	MPEG	$8101 \sim 8550$	Wool	MPEG	$8551 \sim 9000$
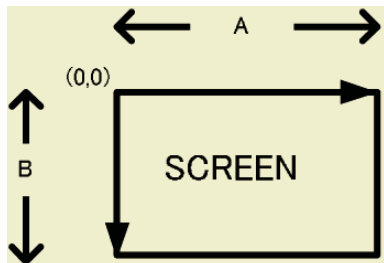
Fig. 4. Definition of Screen.

Fig. 5. WM Embedded Image.

Fig. 5 shows the watermark embedded picture and Fig. 6 shows the re-shot picture without watermark respectively. Fig. 7 shows the original picture quality. It should be emphasized that the picture quality of Fig. 6 is not so bad, and would

Fig. 6. Re-shot Image.

Fig. 7. Original Picture.
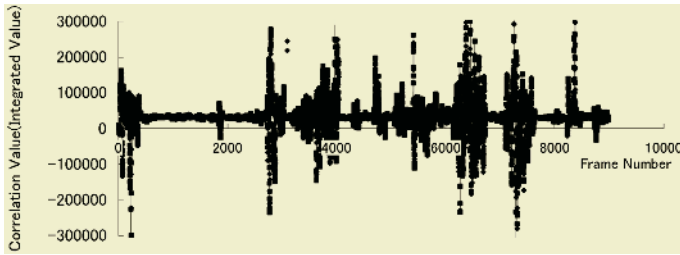
Fig. 8. Experimental Result.

be acceptable in the contents business. This is the cause of piracy in theaters. As Fig. 5 and Fig. 6 are taken by a video camera, they have analogue distortions. It is very difficult to adjust two pictures pixel by pixel to measure PSNR of these pictures. Instead of that, it is theoretically possible to calculate PSNR between the original picture and the embedded picture. Equation (13) gives 52.4dB under the condition of $\alpha = 2$.

One of the most important requirements for watermarks is that they do not degrade picture quality. Comparing Fig. 5 and Fig. 6 to check their qualities, it is almost impossible even for video experts to detect the degradation caused by the embedded watermark.

Fig. 8 shows an experimental result with equation(13). The x axis denotes the frame number and the y axis denotes the correlation values calculated with equation (7). Here, α equals 2, which means the image pixels change by +1 or -1 at most embedded pixels and change by +2 or -2 at most, and are very slightly changed in images. It is desirable that the correlation values are as large as possible, in which case we can set a high threshold value to detect whether a watermark is embedded in images or not. Here, if we set the threshold level as 25,000, most of them are above it. This means that the images shown in Table 1 contain an embedded watermark. Due to the resolution of the graph, it looks like as if many frames show the correlation values that are lower than 25,000. Though we checked all of them, 98.9% of them are higher than 25,000. If we change it to 20,000, 99% are higher than it.

Some frames have lower values than the threshold. However, this is raw data of watermark detection, and we usually use an error correction code in those

cases. The error rate of watermark technologies is higher than that of normal transmission lines, so it is not appropriate to use common error correction codes such as Reed Solomon and BCD. The major detectable decision error correction code is commonly used for watermarks; when this code was used, all of the two bits embedded in every 450 frames were detected by our method.

According to the result of this experiment, our approach is adequate for a re-shot watermark system. We conducted experiments with three functions that are equation(11), (12) and (13). We also changed α from2 to 4. Due to lack of space, other results are reported only briefly. All three functions showed almost the same detection rate except images which contained high-speed moving areas, such as Cheerleaders or Football in the MPEG reference images. Stationary images and images which have low-speed moving areas did not show any difference. The embedding intensity of the watermark, α, should be lower than 3, which means image pixels change by +2 or -2 in areas and change by +3 or -3 at most. If we set it higher than 4, it becomes visible in many images.

There are several approaches for watermarking, such as analogue copies, geometric distortions, rotation and scaling. Although we have not checked all of these approaches, there are several experimental results. We made fourth generation copies of a VHS tape in the economy playback mode (EP/3x) and successfully detected embedded information from them. We moved the re-shot camera position until the re-shot images were visibly distorted. However, embedded information is detected with the assistance of major detectable decision error correction code. Of course it is necessary to continue further study, but our method offers some tolerance to geometric distortion.

5   Conclusion

In this paper we described a new watermark scheme which survives re-shooting the images displayed on a screen, and explained the concept and experimental results of the scheme. Although this approach meets several requirements, such as robustness and picture quality, further discussion is required about this type of watermark.

We used only two-dimensional functions in this paper. Three-dimensional functions offer wider possibilities to develop better watermarking systems. Further studies will be needed in order to use these in practical systems.

A watermark itself cannot protect contents, but is the last defense to prevent illegal copying. Re-shooting of TV monitors to copy contents will soon become possible. We hope that these technologies will be studied and improved in order to protect copyrighted material so that the content business will have a bright future.

References

1. http://www.sarnoff.com/news/
2. http://www.nikkei.co.jp/news/shakai/20050221AT1G2101521022005.html
   (in Japanese)

3. Jaap Haitsma, Ton Kalker,"A Watermarking Scheme for Digital Cenema", Proceedings of IEEE International Conference on Image Processing (ICIP), (2001) 487–489
4. Jaap Haitsma, Ton Kalker,"On Digital Cenema and Watermarking", Security and Watermarking of Multimedia Contents, Proceedings of SPIE-IS& T Electronic Imaging, SPIE Vol.5020(2003), pp526-353
5. http://www.macrovision.com/

Semi-fragile Watermarking Based on Zernike Moments and Integer Wavelet Transform

Xiaoyun Wu[1], Hongmei Liu[1,2], and Jiwu Huang[1]

[1] School of Information Science and Technology, Sun Yat-Sen University
Guangzhou, Guangdong 510275, China
`isshjw@zsu.edu.cn`
[2] The state key laboratory of information security, Beijing, 100039, P.R. China

Abstract. A semi-fragile watermarking algorithm for digital image content authentication is proposed in this paper. The watermark is generated from the Zernike moments of the low frequency subband in IWT (integer wavelet transform) domain of the original image. This sort of image-dependent watermark can resist forgery attack and improve the watermark security. Watermark is embedded in the IWT domain of an image by exploiting the features of human visual system (HVS). It makes the high-fidelity and high-quality watermarked image. Moreover, we can distinguish mild distortion from severe distortion according to the difference of Zernike moments before and after attack. Experimental results show that the proposed algorithm is fragile to malicious tamper and can locate the tampered area accurately while is robust to content-preserved manipulation, such as JPEG lossy compression.

1 Introduction

Owing to the rapid development of internet, digital image has spread all over the world. It is easily manipulated by digital image editing software. Such manipulations make no visual clue in the modified images [1]. Thus, people may suspect their authenticity once they are evidence submitted to the court. As a result, verification of the integrity and authenticity of digital image has been drawn more attention.

In the literature, the techniques for authentication of digital images are classified into digital signature-based and fragile watermarking-based. Although it has higher security, digital signature method has the following shortcomings: i) Signature appends to the image, it can be easily removed. ii) It can not resist content-preserved operation, such as noise, JPEG lossy compression. As an effective supplement, fragile watermarking approach makes up the above limitations.

Fragile watermarking can be roughly classified as exact fragile watermarking and semi-fragile watermarking [2]. Images that are embedded an exact fragile watermark can detect even a single pixel's change. It is applicable to exact authentication. Nowadays, digital images inevitably suffer from normal signal processing, such as JPEG lossy compression, blurring or sharpening etc. It demands an authentication system should give an authentic result if the image underwent incidental distortions. Furthermore, it should be fragile to malicious manipulation. A feasible solution is to embed a semi-fragile watermark in images. It is a tradeoff between robustness and fragility.

As a practical scheme for fragile watermarking, we think it should have the following characteristics: i) Perceptual invisibility. ii) Higher security. It can resist forgery attack [3]. iii) Blindness. Watermark extraction requires no original image. iv) Ability to tolerate JPEG lossy compression to a large extent. v) Ability to detect and locate malicious tamper with high probability. vi) Lower computational complexity in watermark embedding and extraction.

Although many fragile watermarking algorithms have been proposed in the literature [4-9], they can not completely realize all the features above. Hence, we propose a semi-fragile watermarking algorithm based on Zernike moments and integer wavelet transform. We construct an image-dependent watermark via Zernike moments of an image. It can prevent watermark from forgery. Imperceptibility and lower computational load are guaranteed by embedding the watermark in the IWT domain of an image based on the characteristics of HVS. We can extract watermark without the original image. Experimental results demonstrate that the proposed scheme is fragile to malicious attack and locate modified region accurately while is tolerant of JPEG lossy compression to a large extent.

The paper is organized as follows. In Section 2, we describes the proposed algorithm, including the construction of IWT of 9-7 biorthogonal wavelet, the analysis of the semi-fragile characteristic of Zernike moments and watermark generation, watermark embedding, watermark extraction and tamper detection. Performance analysis on security and computational complexity is given in Section 3. Experimental results and conclusions are given in Section 4 and 5, respectively.

2   The Proposed Algorithm

2.1   Integer Wavelet Transform

We use lifting scheme to realize integer wavelet transform. Generally speaking, lifting scheme includes three steps: splitting, prediction and update [10]. Daubechies gives an example of the lifting of CDF 9-7 biorthogonal wavelet [10]. To one dimensional signal $\{x_l\}_{l \in Z}$, the lifting steps are described as follows.

$$\begin{cases} s_l^{(0)} = x_{2l} \\ d_l^{(0)} = x_{2l+1} \end{cases} \quad \begin{cases} d_l^{(1)} = d_l^{(0)} + \alpha(s_l^{(0)} + s_{l+1}^{(0)}) \\ s_l^{(1)} = s_l^{(0)} + \beta(d_l^{(1)} + d_{l-1}^{(0)}) \end{cases} \quad \begin{cases} d_l^{(2)} = d_l^{(1)} + \gamma(s_l^{(1)} + s_{l+1}^{(1)}) \\ s_l^{(2)} = s_l^{(1)} + \delta(d_l^{(2)} + d_{l-1}^{(2)}) \end{cases} \tag{1}$$

$$\begin{cases} s_l = \zeta s_l^{(2)} \\ d_l = d_l^{(2)} / \zeta \end{cases} \tag{2}$$

$$\alpha = -1.586134342; \ \beta = -0.05298011854; \ \gamma = 0.8829110762;$$

$$\delta = 0.4435068522; \ \zeta = 1.149604398 \tag{3}$$

where s_l and d_l are generally referred to as lower frequency and detail coefficients, respectively. $s_l^{(i)}$, $d_l^{(i)}$ (i=0, 1, 2) are mid-outputs.

According to integer wavelet theory [11], we construct integer wavelet transform based on the framework mentioned above. That is:

$$\begin{cases} s_l^{(0)} = x_{2l} \\ d_l^{(0)} = x_{2l+1} \end{cases} \begin{cases} d_l^{(1)} = d_l^{(0)} + Int(\alpha(s_l^{(0)} + s_{l+1}^{(0)})) \\ s_l^{(1)} = s_l^{(0)} + Int(\beta(d_l^{(1)} + d_{l-1}^{(1)})) \end{cases} \begin{cases} d_l^{(2)} = d_l^{(1)} + Int(\gamma(s_l^{(1)} + s_{l+1}^{(1)})) \\ s_l^{(2)} = s_l^{(1)} + Int(\delta(d_l^{(2)} + d_{l-1}^{(2)})) \end{cases} \tag{4}$$

$$\begin{cases} d_l^{(3)} = d_l^{(2)} + Int((\zeta - \zeta^2)s_l^{(2)}) \\ s_l^{(3)} = s_l^{(2)} + Int((-1/\zeta)d_l^{(3)}) \end{cases} \begin{cases} d_l^{(4)} = d_l^{(3)} + Int((\zeta - 1)s_l^{(3)}) \\ s_l^{(4)} = s_l^{(3)} + d_l^{(4)} \end{cases} \begin{cases} s_l = s_l^{(4)} \\ d_l = d_l^{(4)} \end{cases} \tag{5}$$

where $Int(x)$ means integer part of x. The values of parameters α, β, γ, δ, ζ are given in formula (3).

2.2 Zernike Moments and Watermark Generation

Zernike moments are the feature of a digital image and widely applied to pattern recognition, target classification, target identification and scene analysis. The Zernike moments of order n with repetition m for a digital image may be referred to [12].

Zernike moments magnitudes (ZMMs) have semi-fragile characteristic. That is, they are robust to content preserved manipulations and fragile to content altered operations. Such characteristic can be evaluated by computing the differences of ZMMs between the original image and the processed image. We exhibit this characteristic here by experiments on Lena, Baboon, Boat, and Goldhill images.

Considering the high computational load in computing Zernike moments of the whole image, we choose LL_3 subband in IWT domain, which is low-pass approximation of the original image and includes most of the image energy, to compute 49 ZMMs. The difference of ZMMs of an image before and after processed is defined as:

$$\Delta = \sum_{i=1}^{49} (M_i - M_i')^2 \tag{6}$$

where $M_1 \sim M_{49}$ denote 49 moment values before processed, $M_1' \sim M_{49}'$ denote 49 moment values after processed and Δ denote difference of ZMMs before and after processed.

Table 1 demonstrates experimental results. We can see that if an image is suffered from incidental modification (JPEG), Δ is small. Otherwise Δ is large in the case of malicious attack (cut and replacement). It indicates the semi-fragile characteristic of Zernike moments in LL_3 subband. Furthermore, it is clear that Δ of non-malicious attack is far smaller than Δ of malicious modification. Using this feature, we can distinguish incidental distortion from severe distortion.

Table 1. Difference of ZMMs in LL_3 subband of an image befor and after processed

Attack Type	Lena	Baboon	Boat	Goldhill
JPEG Q=80	0.000165	0.000201	0.000544	0.000351
JPEG Q=60	0.000624	0.000705	0.001241	0.000913
JPEG Q=50	0.001012	0.001668	0.002117	0.001963
JPEG Q=40	0.002972	0.003661	0.006721	0.005975
JPEG Q=30	0.007684	0.009218	0.014281	0.011984
Cut (16×16)	0.049359	0.086567	0.460912	0.224589
Replace(16×16)	0.035813	0.061542	0.416027	0.170327

We use 49 ZMMs in LL_3 subband of the image to construct watermark *W*. Denote these ZMMs as $M_i(org)$, $1 \leq i \leq 49$. Each moment is quantized uniformly first and then coded to a 16 bits binary number. Thus, watermark length is 784 bits. If the length is less than the number of wavelet coefficients that will be used to embed watermark, we can simply fill it using "0" or "1". To enhance the security of watermark data, we may further encrypt or scramble these data.

2.3 Watermark Embedding and Extraction

Watson et al. [13] proposed a wavelet-based HVS. We adopt this perceptual model to embed watermark for the purpose of satisfying transparency of a watermarked image. HL_3 subband is used to embed the watermark. To embed a watermark bit, we first calculate

$$Q(i,j) = \begin{cases} 0 & \lfloor f(i,j)/JND(i,j) \rfloor \text{ is even} \\ 1 & \lfloor f(i,j)/JND(i,j) \rfloor \text{ is odd} \end{cases} \tag{7}$$

where *f(i,j)* denotes an IWT coefficient in HL_3 subband, $JND(\cdot)$ is the corresponding Watson's quantization matrix. $\lfloor \cdot \rfloor$ is the floor function. Let $w(i,j)$ denote the current watermark bit to embed. We alter the wavelet coefficient as:

$$\tilde{f}(i,j) = \begin{cases} f(i,j) & \text{if } Q(i,j)=w(i,j) \\ (\lfloor f(i,j)/JND(i,j) \rfloor \pm 1) \cdot JND(i,j) & \text{if } Q(i,j) \neq w(i,j) \end{cases} \tag{8}$$

where if $f(i,j)$ is positive, we choose "+", and otherwise, "−". With inverse integer wavelet transform, we get the watermarked image.

Denote an IWT coefficient in HL_3 subband of the to-be-tested image as $f'(i,j)$. The extraction formula is expressed as:

$$w'(i,j) = \begin{cases} 0 & \lfloor f'(i,j)/JND(i,j) \rfloor \text{ is even} \\ 1 & \lfloor f'(i,j)/JND(i,j) \rfloor \text{ is odd} \end{cases} \tag{9}$$

2.4 Tamper Detection

To a suspect image, we compute its 49 ZMMs in LL_3 subband. Denote them as $M_i(new)$, $1 \leq i \leq 49$. Using formula (10), we compute the difference of ZMMs.

$$\Delta' = \sum_{i=1}^{49} \left(M_i(org) - M_i(new) \right)^2 \tag{10}$$

In Section 2.2, we point out that the difference of ZMMs can tell the malicious modifications from incidental distortions. Here, we use a threshold τ to distinguish them. According to a lot of experiment, we find that if Δ' is smaller than 0.003 when the watermarked image goes through non-malicious attack and Δ' is bigger than 0.003 when it is malicious attacked. We set τ to 0.003.

Once the image is regarded as malicious modification, we use the difference of the watermark defined below to locate tampered region.

$$D = |W - W'| \tag{11}$$

where W' is the extracted watermark.

3   Performance Analysis

3.1   Security

To a fragile watermarking scheme, the deadly attack is forgery attack. In a forgery attack, the attacker modifies image content without altering the watermark. Although the image is modified, the authentication system gives authentic result because the watermark is unchanged. Our proposed scheme can discern this kind of attack. The reason is as follows.

When the watermarked image undergoes forgery attack, the image content is changed accompanied by the variation of $M_i(new), 1 \le i \le 49$. According to formula (10), we can conclude that the image is maliciously attacked. However, when we further locate the tampered regions, the difference of the watermark shows that the watermarked image is not modified. Two contradictive results indicate the watermarked image go through forgery attack. Hence, we resist this attack.

3.2   Computation Complexity

Given a signal with the length n, the computational load of conventional DWT and DCT may be expressed as $O(n)$ and $O(n \log(n))$, respectively. The conventional DWT offers lower computational complexity than DCT. According to Equation (4)~(5), to one dimensional signal, the 9-7 IWT based on lifting scheme needs 12 integer additions, 7 floating-point multiplications and 7 round-off operations for two wavelet coefficients, and for conventional 9-7 DWT, 14 floating-point additions and 16 floating-point multiplications should be used. The 9-7 IWT will save almost half of the computational cost as compared with the conventional 9-7 DWT.

4   Experimental Results

We use "Lena" and "Baboon" images (256×256×8 bits) to evaluate our scheme. Fig. 1 shows the results of imperceptibility of the watermark. The PSNRs are 43.62dB for "Lena" and 43.76dB for "Baboon", respectively. We can see no difference between the original images and the watermarked images. Fig. 2 shows the results of fragility to malicious tamper such as cut, replacement. It is obvious that our algorithm can detect malicious modifications and locate the tampered areas with high probability. Table 2 shows robustness to JPEG compression. It is demonstrated that JPEG-50 is classified as non-malicious processing, implying that the proposed watermarking method is robust to JPEG at quality of 50.
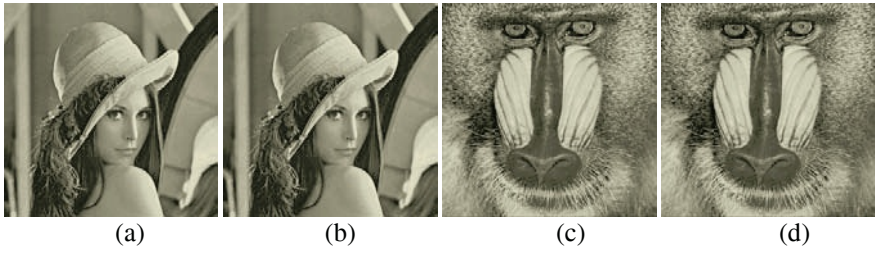
Fig. 1. Imperceptibility. (a), (c) Original images. (b), (d) Watermarked images.
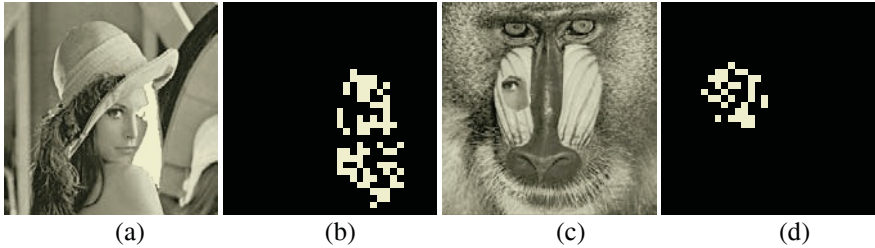
Fig. 2. Tamper Detection. (a), (c) Tampered images. (b) , (d) Difference images of extracted watermarks.

Table 2. Robustness to JPEG lossy compression

JPEG quality	Lena	Baboon
60	0.000802	0.000907
50	0.001472	0.002284
40	0.003683	0.004986

5 Conclusions

In this paper, we propose a semi-fragile watermarking method based on Zernike moments and integer wavelet transform. We choose Zernike moments of an image to construct the watermark. It can resist forgery attack efficiently. Using Watson's quantization matrix and integer wavelet transform, we embed the watermark in the image. It ensures perceptual invisibility and lower computational load. Experimental results show the proposed scheme can tolerate JPEG lossy compression as low as quality of 50 and rejects malicious modifications while tampered area is located accurately.

Acknowledgement

References

1. Fridrich, J.: Security of fragile authentication watermarks with localization. SPIE Security and Watermarking of Multimedia Contents IV, San Jose (2002) 691-700

2. Fridrich, J.: Methods for tamper detection in digital images. Proc. ACM Workshop on Multimedia and Security, Orlando (1999) 19-23
3. Lin, E.T., Delp, E.J.: A review of fragile image watermarks. Proc of the Multimedia and Security Workshop at ACM Multimedia (1999) 36-39
4. Yeung, M., Mintzer, F.: An invisible watermarking technique for image verification. Proc. of ICIP, Vol. 2 (1997) 680-683
5. Wu, M., Liu, B.: Watermarking for image authentication. Proc. of ICIP, Vol. 2 (1998) 437-441
6. Kundur, D., Hatzinakos, D.: Towards a telltale watermark techniques for tamper-proofing. Proc. of ICIP, Vol. 2 (1998) 409-413
7. Lin, C.Y., Chang, S.F.: Semi-fragile watermarking for authentication jpeg visual content. SPIE Security and Watermarking of Multimedia Content II, San Jose (2000) 140-151
8. Xie, L., Arce, G.: Joint wavelet compression and authentication watermarking. Proc. of ICIP, Vol. 2 (1998) 427-431
9. Han, S.J., Chang, I.S., Park, R.H.: Semi-fragile watermarking for tamper proofing and authentication of still images. Proc. of IWDW (2003) 347-358
10. Daubechies, I., Sweldens, W.: Factoring wavelet transforms into lifting steps. Journal of Fourier Analysis, Vol. 4 (1998) 245-267
11. Calderbank, R., Daubechies, I., Sweldens, W., Yeo, B.L.: Wavelet transforms that map integers to integers. Journal of Applied and Computational Harmonic Analysis, Vol. 5 (1998) 332-369
12. Khotanzad, A., Hong, Y.H.: Invariant Image Recognition by Zernike moments. IEEE Trans on Pattern Analysis and Machine Intelligence Vol. 12 (1990) 489-497
13. Watson, B., Yang, G.Y., Solomon, A., Villasenor, J.: Visibility of wavelet quantization noise. IEEE Trans on Image Processing Vol. 6 (1997) 1164-1175

Shadow Watermark Extraction System

Feng-Hsing Wang [1], Kang K. Yen [2], Lakhmi C. Jain[1], and Jeng-Shyang Pan [3]

[1] School of Electrical and Information Engineering, University of South Australia, Australia
lakhmi.jain@unisa.edu.au
http://www.kes.unisa.edu.au/tiki-view_articles
[2] Department of Electrical and Computer Engineering, Florida International University, USA
[3] Department of Electronic Engineering
National Kaohsiung University of Applied Sciences, Taiwan

Abstract. A new watermarking scheme which provides the ability of sharing secret with multi-users is proposed. It splits the original watermark into two shares and embeds one share into the cover image to increase the security. A polarization procedure is performed to establish a polarity stream from the cover image. The second share and the polarity stream are used to generate a master key and several normal keys. In our system, only the super user can reveal the genuine watermark directly. Other users possess the normal keys can obtain shadow watermarks merely. By combining the shadow watermarks together, the real watermark can be recovered.

1 Introduction

Today the storage of item in digital formats such as documents, images, video, and audio is very common. As is well known, due to the digital nature of the information, it is easy to make unlimited lossless copies from the original digital source, to modify the content, and to transfer the copies rapidly over the Internet. Therefore, the demands of copyright protection, ownership demonstration, and tampering verification for digital data are becoming more and more urgent. Among the solutions for these problems, digital watermarking [1]–[6] is the most popular one. Researchers have given consideration to this in the past decade.

A variety of watermarking methods have been proposed. Some of them provide better imperceptibility and some of them focus on other benchmarks such as robustness or embedding capacity. In this paper, we focus on the security of the watermark to be hidden and propose a new watermarking system for sharing the watermark, which is regarded as a secret, with multi-users [7]–[9]. The proposed system splits the original watermark into two shares, embeds one share into the cover image, and uses another share for the proposed key generating procedure. The proposed system therefore enables the user which possesses the unique master key to reveal the genuine watermark directly. Otherwise, a shadow watermark can be obtained merely. We illustrate how to implement this system and give the experimental results to show its effectiveness.

2   Proposed Watermarking System

The structure of the proposed system is depicted in Fig. 1. The brief steps are listed below. Details of the methods for implementation are given in the following sections.
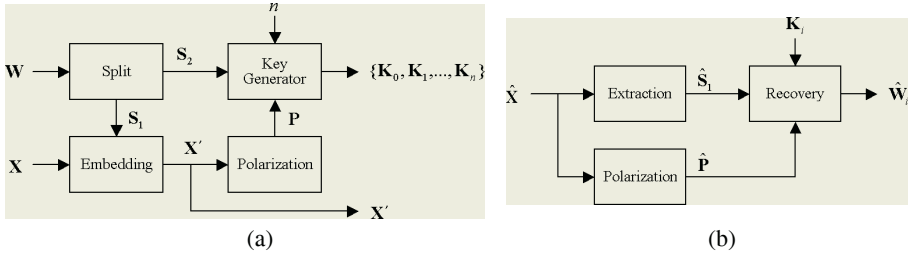
Fig. 1. Block diagrams of the proposed watermarking system: (a) the embedding procedure and (b) the extraction procedure

The steps of the embedding procedure are as follows:

Step 1:   Split the original watermark $\mathbf{W}$ into two shadow watermarks $\mathbf{S}_1$ and $\mathbf{S}_2$ by applying the method presented in Sect. 3.1.

Step 2:   Embed the shadow watermark $\mathbf{S}_1$ into the cover image $\mathbf{X}$ so that a watermarked image $\mathbf{X}'$ can be obtained.

Step 3:   Establish a polarity stream $\mathbf{P}$ from $\mathbf{X}'$. The methods presented in Sect. 3.3 illustrate this step.

Step 4:   Generate the unique master key $\mathbf{K}_0$ and n normal keys $\{\mathbf{K}_1, \mathbf{K}_2, ..., \mathbf{K}_n\}$ from $\mathbf{S}_2$ and $\mathbf{P}$. Section 3.4 describes the details.

To obtain a watermark, which may be the original watermark or a shadow watermark, according to the key used, from a watermarked image $\hat{\mathbf{X}}$, the steps below are used:

Step 1:   Extract a shadow watermark $\hat{\mathbf{S}}_1$ from $\hat{\mathbf{X}}$.

Step 2:   Establish a polarity stream $\hat{\mathbf{P}}$ from $\hat{\mathbf{X}}$ by applying the method described in Sect. 3.3.

Step 3:   Generate a watermark $\hat{\mathbf{W}}_i$, where $0 \le i \le n$, from $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{P}}$, according to the key $\mathbf{K}_i$ used. If $\mathbf{K}_i$ is a master key, i.e., $i = 0$, then $\hat{\mathbf{W}}_i$ is the original watermark. Otherwise, $\hat{\mathbf{W}}_i$ is a shadow watermark. Step 4 is then used to recover the original watermark.

Step 4:   Combine all the shadow watermarks extracted to recover the original watermark. The method proposed in Sect. 3.5 explains this step.

3   Methods for The Proposed Watermarking System

Generally speaking, to implement the proposed system, any method available can be employed. In the following sections, some methods are given as examples for implementing each part of this system.

3.1   Watermark Splitting

To split the original watermark into two shares, a method such as visual cryptography [7] can be applied. Below two other methods are given for achieving this goal.

The most easy and effective method is to apply the exclusive-or (XOR) operator. We first generate the share $\mathbf{S}_1$ randomly. For example, if the length of $\mathbf{S}_1$ is 5, then $\mathbf{S}_1 = \{0,1,0,0,1\}$ may be used. The XOR operator is then applied to generate the second share $\mathbf{S}_2$. That is, $\mathbf{S}_2 = \mathbf{W} \oplus \mathbf{S}_1$.

To increase the security of the system, using a fake watermark to generate a share [8] can be considered. For the original watermark $\mathbf{W}$ and a fake watermark $\mathbf{W}_F$, a temporary image $\mathbf{S}_{tmp}$ is generated from $\mathbf{W}$ and $\mathbf{W}_F$ using the XOR operator. Then, according to a predefined seed for the pseudo-random-number generator, the pixel-permutation procedure is performed on $\mathbf{S}_{tmp}$ in order to disturb its spatial relationship. The permutated result and the fake watermark are regarded as $\mathbf{S}_2$ and $\mathbf{S}_1$ respectively. An example of applying the above procedure is given in Fig. 2.

(a)   (b)   (c)
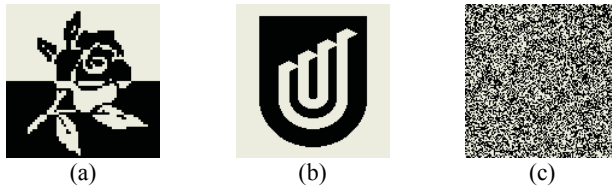
Fig. 2. The example of generating a shadow watermark using a fake watermark. (a) The original watermark. (b) The fake watermark. (c) The generated result

3.2   Watermark Embedding and Extraction

In this system, the choice of the watermarking algorithm is not limited. Any suitable watermarking method proposed in the literature can be used. Details of the embedding and the extraction schemes are not included here.

3.3   Polarization

To establish a polarity stream from an image, the method proposed in the literature such as [9] can be employed. Another method based on vector quantisation (VQ) [10] is given here.

Let **X** be a given image and **C** be a set of l predefined codewords. The size of any codeword is d pixels. And **C** can be a public or private codeword set, according to the concern of security. The steps below are used:

Step 1: Decompose **X** into k non-overlapping blocks.
Step 2: Execute the nearest-codeword-search procedure [10] to obtain a nearest codeword from **C** for each of the blocks. The indices $\{I_1, I_2,..., I_k \mid I_i \in [0,l), 1 \le i \le k\}$ of the codewords obtained are collected as the index set **I**.
Step 3: Establish the polarity stream **P** from **I** using Eq. (1):

$$\mathbf{P} = \bigcup_{i=1}^{k} (I_i \ \text{MOD} \ 2). \tag{1}$$

Two examples of generating the polarity streams from the image of LENA are shown in Fig. 3.

(a)   (b)

Fig. 3. The polarity streams established from the image of LENA using the proposed method with different values of l: (a) $l = 128$ and (b) $l = 256$

3.4 Key Generating

To generate one master key and n normal keys from $\mathbf{S}_2$ and **P**, we have:

Step 1: Generate the master key $\mathbf{K}_0$ by applying the XOR operation to $\mathbf{S}_2$ and **P**.
Step 2: Generate the $(n-1)$ normal keys $\{\mathbf{K}_1, \mathbf{K}_2,..., \mathbf{K}_{n-1}\}$ randomly.
Step 3: Generate the last normal key $\mathbf{K}_n$ by applying the XOR operation:

$$\mathbf{K}_n = \mathbf{K}_0 \oplus \mathbf{K}_1 \oplus ... \oplus \mathbf{K}_{n-1}. \tag{2}$$

The master key is then delivered to the super user and the other n normal keys are delivered to the related users to share the hidden secret **W**.

3.5 Watermark Recovery

As shown in Fig. 1(b), after obtaining the polarity stream $\hat{\mathbf{P}}$ and the embedded share $\hat{\mathbf{S}}_1$ from the watermarked image $\hat{\mathbf{X}}$, a watermark $\hat{\mathbf{W}}_i$, where $0 \le i \le n$, can be determined according to the key $\mathbf{K}_i$ used. That is, $\hat{\mathbf{W}}_i = \hat{\mathbf{S}}_1 \oplus \hat{\mathbf{P}} \oplus \mathbf{K}_i$. Here if $i = 0$, $\hat{\mathbf{W}}_i$ is the original watermark. Otherwise, $\hat{\mathbf{W}}_i$ is a shadow watermark. When using a

normal key for watermark extraction, the additional step below has to be considered for revealing the real watermark from the shadow watermarks:

$$\hat{\mathbf{W}} = \hat{\mathbf{W}}_1 \oplus \hat{\mathbf{W}}_2 \oplus ... \oplus \hat{\mathbf{W}}_n. \tag{3}$$

4   Simulation Results

In our experiments, the image of LENA and the image of ROSE (Fig. 2(a)) were used as the cover image and the original watermark respectively. The sizes of them are 512×512×8 bit/pixel and 128×128×1 bit/pixel, respectively. The codeword set used in the polarization procedure were generated using the LBG algorithm [10] with a threshold of 0.0001. The DCT-based watermarking scheme proposed in [5] was employed to implement the embedding and extraction methods of the proposed system, where we hid 4 bits in the 14th, the 15th, the 16th, and the 27th DCT coefficients of each 8×8 DCT block. Other settings used were $n = 3$, $m = 4096$, and $k = 16384$. The peak-signal-to-noise ratio (PSNR) and the bit-correct ratio (BCR) were used to evaluate imperceptibility and robustness respectively.

After splitting the original watermark into two shares and embedding the first share into the image of LENA, the PSNR value between the original cover image and the watermarked image is 33.87 dB. When no attack is applied on the watermarked image, Fig. 4(a) to 4(d) display the watermarks extracted using the master key and other three normal keys, respectively. Figure 4(e) shows the watermark recovered from the shadow watermarks listed in Fig. 4(b) to Fig. 4(d).

To test the robustness of the system, some common image-processing procedures were employed to attack the watermarked images. The attacks and the test results are listed in Table 1, where QF denotes the quality factor used in the JPEG compression. The watermarks extracted from the attacked images when $l = 32$ are displayed in Fig.5.

| (a) | (a) | (c) | (d) | (e) |
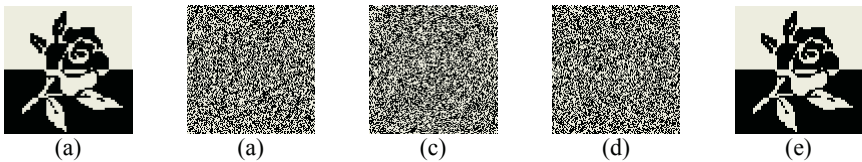
Fig. 4. The watermarks extracted and recovered from the watermarked image when $l = 128$

5   Discussion

Some issues which affect the performance or relate to the scheme are listed below:

(i)   From Table 1, it can be seen that a smaller value of l, which means a larger degree of quantisation, leads to stronger robustness.

(ii) In addition to the master key and normal keys used, other parameters related to the security of this system are the codebook **C** and the parameters used in the employed watermarking algorithm. These parameters can be public or private according to the concern for security.

Table 1. The robustness test results (BCR in %) under some common attacks

Attack	BCR (%)			
	$l=32$	$l=64$	$l=128$	$l=256$
JPEG (QF=20%)	74.25	72.95	67.52	60.32
JPEG (QF=30%)	90.39	88.61	83.47	78.69
JPEG (QF=80%)	97.80	97.27	95.70	94.42
Median filtering	84.47	84.08	81.74	79.11
Low pass filtering	80.30	79.47	77.31	73.65
Cropping	86.91	87.81	88.05	87.65
Shifting 1 line downwards	87.28	85.43	81.41	76.56

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)
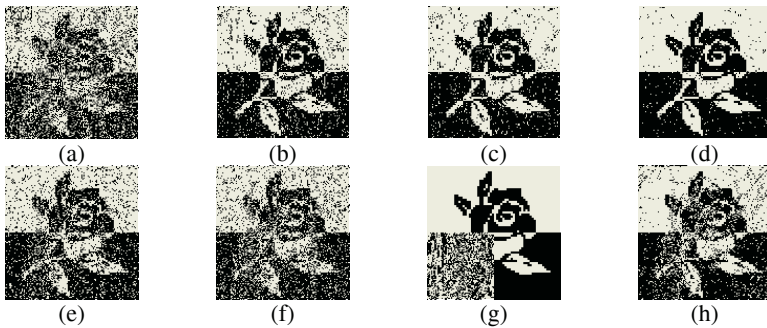
Fig. 5. The watermarks extracted from the attacked images when $l=32$: (a) JPEG with QF=20%, (b) JPEG with QF=30%, (c) JPEG with QF=40%, (d) JPEG with QF=80%, (e) median filtering with window size = 3, (f) low-pass filtering with window size =3, (g) cropping the left-bottom corner, and (h) shifting 1 line downwards

(iii) The size of **C**, which is *l*, also affects the security. That is, a larger *l* results in stronger security.

(iv) Due to the method used for establishing a polarity stream, the size of **C** affects the coding time. The coding time increases with the size of **C**.

(v) In some cases the question of key delivery is of concern, but it is beyond the focus of this paper.

(vi) The proposed system recovers the genuine watermark from the shadow watermarks and the polarity stream. Using a watermarking algorithm provides better imperceptibility also helps to improve the extracted results.

(vii) The architecture of the proposed watermarking system is open. Either part of the system can be replaced by other parts which provide similar functions, for example, replacing the watermarking procedure by other available methods, as mentioned in (vi). This means the system is easy to implement.

6 Conclusion

An open-frame watermarking system for sharing secrets with multi-users has been presented. The watermark extracted may be the genuine watermark or a shadow wa-

termark according to the key used for extraction. This means the owner of the unique master key can reveal the hidden secret directly. The normal users have to presente their shadow watermarks extracted to reveal the secret. The robustness against some attacks is good enough even with a high JPEG compression rate (QF=20%). In conclusion, the proposed system is novel, effective, as the experimental results have shown.

References

1. Barni M., Bartolini F.: Watermarking Systems Engineering: Enabling Digital Assets Security and Other Applications, Marcel Dekker, Inc., New York (2004)
2. Katzenbeisser, F. Petitcolas (eds.): Information Hiding: Techniques for Steganography and Digital Watermarking, Artech House, Norwood, MA (2000)
3. Huang J., Shi Y.Q., Shi Y.: Embedding Image Watermarks in DC Components, IEEE Trans. Circuits and Systems for Video Technology, 10 (6) (2000) 974–979
4. Hsu C.T., Wu J.L.: Hidden Digital Watermarks in Images, IEEE Trans. Image Proc., 8 (1) (1999) 58–68
5. Shieh C.S., Huang H.C., Wang F.H., Pan J.S.: Genetic Watermarking Based on Transform-Domain Techniques, Pattern Recognition, 37 (3) (2004) 555–565
6. Niu X.M., Lu Z.M., Sun S.H.: Digital Image Watermarking Based on Multiresolution Decomposition, Electronics Letters, 36 (13) (2000) 1108–1110
7. Noar M., Shamir A.: Visual Cryptography, Lecture Notes in Computer Science, Vol. 950, Springer-Verlag, (1994) 1–12
8. Pan J.S., Wang F.H., Jain L.C., Ichalkaranje N.: A Multistage VQ Based Watermarking Technique with Fake Watermarks, Lecture Notes in Computer Science, Vol. 2613, Springer-Verlag, (2003) 81–90
9. Wang F.H., Jain L.C., Pan J.S.: A Multi-user Based Watermarking System with Two-Security-Level keys, Lecture Notes in Computer Science, Vol. 2613, Springer-Verlag, (2003) 40–50
10. Gersho A., Gray R.M.: Vector Quantization and Signal Compression, Kluwer Academic Publisher, London (1992)

Development of Nearly Lossless Embedding Technology of Contactless Sensible Watermarks for Audio Signals

Toshio Modegi

Media Technology Research Center, Dai Nippon Printing Co., Ltd.
Ichigaya-Daitoh Building, 6-3, Ichigaya-Daimachi, Shinjuku-ku
Tokyo 162-0066, Japan
Modegi-T@mail.dnp.co.jp
http://www.dnp.co.jp/index.html

Abstract. We have proposed a novel audio watermarking technology, which embeds a set of bitstream data by changing two-channel stereo locations of lower frequency components in an embedding target audio signal. This method features nearly lossless embedding, robustness against lossy data compression or analogue conversion, and enables contactless asynchronous detection of embedded watermarks through speaker and microphone devices without the original audio signals. In this paper, we propose an extended monaural audio watermark embedding method, which supports both monaural and stereo audio signals, and enables watermark detection by a single monaural microphone such as cell phone voice receiver devices. We describe an abstract of our proposed watermark embedding and extracting algorithm, and experimental results of watermark extraction precision.

1 Introduction

There are several problems on the current watermark technology such as quality degradation by embedding and fragility against miscellaneous content modification or format conversion. Some research works try to support robustness against MPEG lossy compression [1], and robustness against analogue conversion [2]. However, quality degradation problems still exist, and general purpose robust watermark embedding methods for audio signals have not been proposed. In particular, a contactless watermark detection method in an acoustic space has not been developed, whereas contactless watermark detection technology for image data has been already proposed using by a cell phone camera [3].

In order to solve these problems, we have proposed a novel audio watermarking technology, which embeds a set of bitstream data by changing two-channel stereo locations of lower frequency components in an embedding target audio signal [4]. As this embedding method modifies locations of only lower frequency components, whose locations cannot be recognized definitely, this enables nearly lossless watermark embedding. In this paper, we propose an extended monaural audio watermark embedding method, which supports both monaural and stereo audio signals, and enables watermark detection by a single monaural microphone. We describe an abstract of our proposed watermark embedding and extracting algorithm, and its evaluation results.

2   Stereo Audio Watermark Method

Figure 1 shows our proposed watermark embedding method for stereo audio signals. This embeds a set of given bitstream data by changing two-channel stereo locations of low frequency components in an embedding given audio signal. In each separated audio frame, a single tri-state code (-1, 0, +1) is embedded. The panning location of lower frequency components (around less than 200Hz) will be changed to the left in case of the embedded code −1, changed to the right in case of the code +1, and changed to the center in case of the code 0. This method will not modify any frequency components except for the lower-frequency sound source direction, which we (human beings) can not clearly recognize. Even if we modify the quality of each channel component, the embedded codes will not disappeared as far as the volume balance of two channels is maintained. Therefore, this method features nearly lossless embedding, robustness against lossy data compression or analogue conversion. And we have confirmed a contactless detection capability of embedded watermarks through speaker and microphone devices.
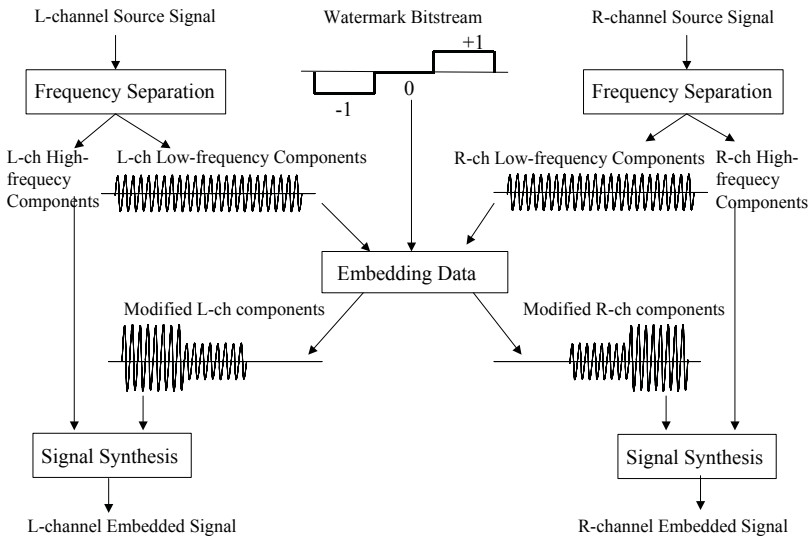
Fig. 1. Watermark embedding method for stereo audio signal.

3   Monaural Supported Method

In order to contactlessly extract the embedded watermarks on the previously proposed method, we have to use stereo dual-directional microphone like shown in the left side of Fig.2. This restricts strictly stereo directional microphone position and requires enough sound source output energy. In this section, we propose an extended method, by which we can detect watermarks with a monaural microphone as shown in the right side of Fig.2. This proposing method enables detection by non-directional monaural miniature microphone units embedded in such as mobile voice recorder or cell phone devices.

[Stereo Embedding Method]   [Monaural Embedding Method]
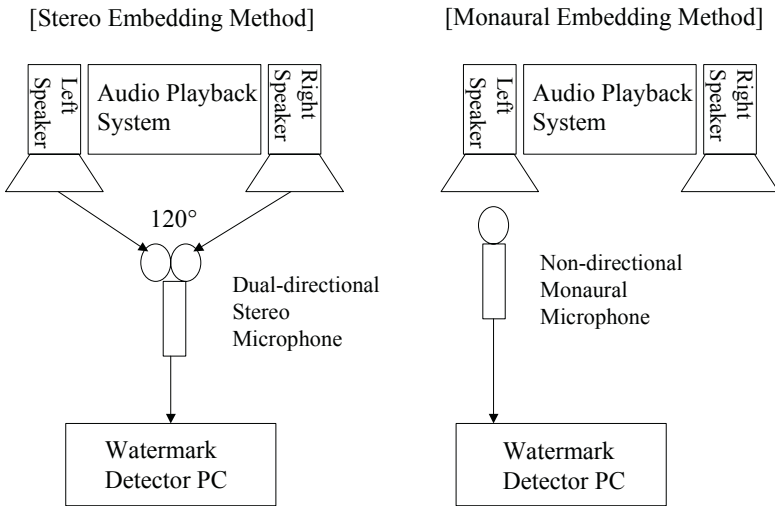
Fig. 2. Watermark contactless detection method using microphones.

Figure 3 shows concepts of our proposing two kinds of watermark embedding method for both stereo and monaural audio signals. The upper one is our previously proposed stereo embedding method, which we had described. In this method, lower frequency sound source components are directed within two-channel horizontal position by the embedded codes. And in order to extract embedded data, we have to compare the level balance between two-channel lower frequency signals. Whereas, the lower one is our newly proposed monaural embedding method, by which we can extract embedded data only form left-side monaural channel. In this method, lower frequency sound sources are dynamically directed to two reverse locations, either from left to right, or from right to left in cases of embedded code -1 and +1.
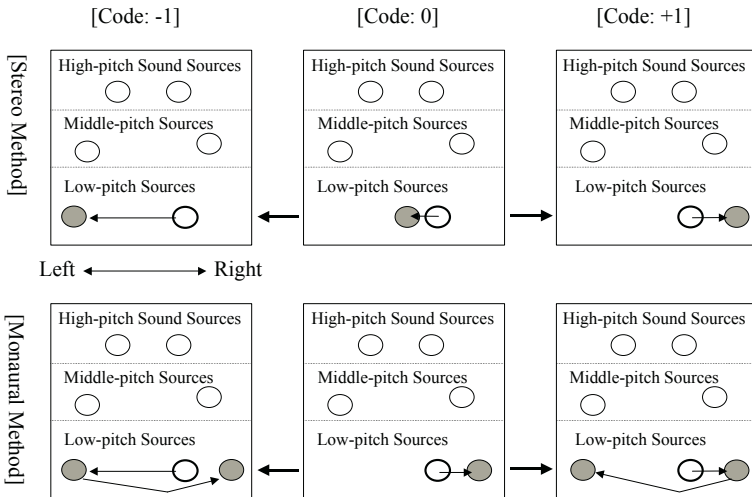
Fig. 3. Concept of two kinds of watermark embedding methods

Figure 4 shows this extended proposal of watermark embedding method based on the lower concept shown in Fig.3. This embeds a set of given bitstream data by changing the level balance between two temporally divided sub-frame low frequency components, in a monaural signal or stereo left-channel signal. The component of the first sub-frame will be removed in case the embedded code is +1, the second sub-frame will be removed in case of the code −1, and the both sub-frame will be removed in case of the code 0. The removed components will be transferred to the corresponding frequency position of the right channel if exists.
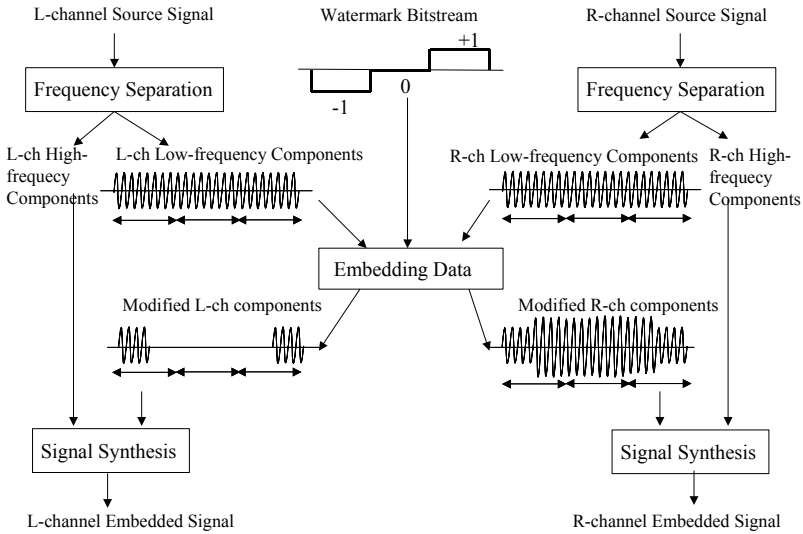
Fig. 4. Monaural supported watermark embedding method.

If the embedded audio signal is played back by stereo, the right-channel can be used for compensating the components removed from the left channel. Therefore, this extended method supports also nearly lossless watermark embedding if the source is stereo, and we can detect the embedded watermark only from the left-channel audio signal using a monaural microphone.

4 Experimental Results

The following is an experimental result based on our proposed watermark embedding method for both stereo and monaural (stereo left-channel) signal. The bit-stream sequence of embedding 8-bit ASCII or binary characters by the stereo embedding method is shown in Figure 5. If the lower frequency component level of embedding target audio frame is not sufficient, bit data embedding will be skipped, instead space code is embedded. Therefore, we have designed the sequence as each byte data would be divided to variable bit-length words and embedded between space codes.
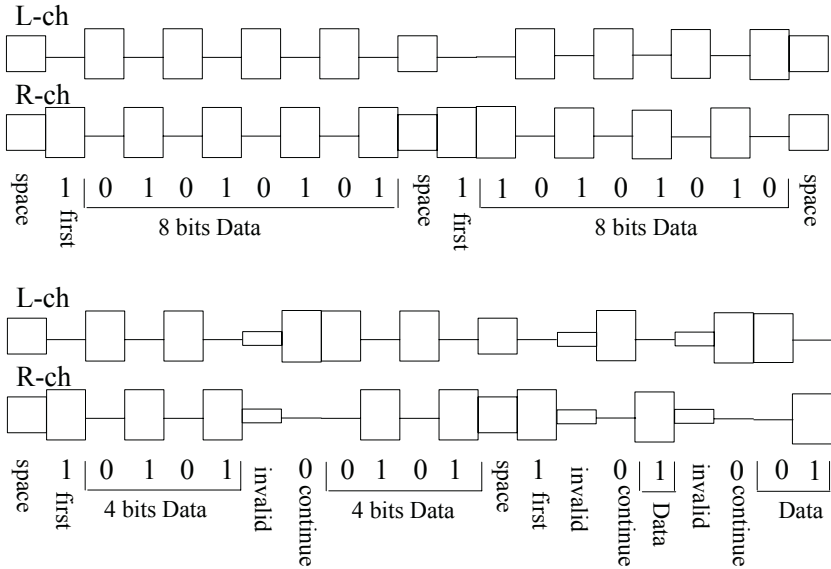
Fig. 5. Variable bit-length word data embedded sequence by stereo embedding method.

4.1 Experiment by Stereo Method

[Experiment Conditions]
- Source Music Materials:
     DNP Image Song "Sharing a Dream", 4min.21sec.
             44.kHz, 16bits, 2-channels, 43.9Mbytes
- Signal Analysis: 4096-point FFT
  Embedding Data at Lower 1-20 point area (< 200Hz).
- Embedding Data: ASCII characters 191 bytes for stereo two channels
                             (repeat the following 80-bytes strings)
     DNPgroup_ImageSong,Sharing-a-Dream,MusicByTakaoKisugi,
     SongByEPO,(C)DNP2002-2005
[Experiment Results]
Appropriately Extracted Characters:
- extracted from embedded digital uncompressed data: 191
- extracted from MP3 (56kbps) compressed digital data: 191
- extracted from the signal resampled through analogue audio lines: 185
- extracted from the signal resampled through speaker and microphone: 158

4.2 Experiment by Manural Method

[Experiment Conditions]
- Source Music Materials:
   DNP Image Song "Sharing a Dream", 4min.21sec.
             44.kHz, 16bits, 2-channels, 43.9Mbytes

- Signal Analysis: 4096-point FFT
   Embedding Data to two divided 2048-point sub-frames.
   Embedding Data at Lower 1-20 point areas (<200Hz).
- Embedding Data: ASCII characters 115 bytes for stereo left channel only
                              (repeat the following 80-bytes strings)
   DNPgroup_ImageSong,Sharing-a-Dream,MusicByTakaoKisugi,
   SongByEPO,(C)DNP2002-2005

[Experiment Results]
The specific extracted string examples are shown in Table 1.
Appropriately Extracted Characters:
- extracted from embedded digital uncompressed data: 115
- extracted from MP3 (56kbps) compressed digital data: 112
- extracted from the signal resampled through analogue audio lines: 112
- extracted form the signal resampled through speaker and microphone: 93

Table 1. Watermark detection experiment results by monaural method.

Extraction from embedded digital uncompressed data: 115 bytes DNPgroup_ImageSong,Sharing-a-Dream,MusicByTakaoKisugi, SongByEPO,(C)DNP2002-2005DNPgroup_ImageSong,Sharing-a-Dream,M
Extraction from MP3 (56kbps) compressed digital data: 112 bytes DNPgroup_ImageSong,Sharing-a-Dream,MusicByTakaoKisugi, SongByEPO,(C)DNP2002-3005DNPgrop_ImageSng,Shaing-a-Dream,M
Extraction from resampled data through analogue audio lines: 112 bytes DNPgroup_ImageSong,Sharing-a-Dream,MusicByTakaoKisgi, SongByEPO,(C)DNP2002-2005DNPgroup_ImageSon,Sharing-a-Drem,M
Extraction from resampled data through speaker and microphone: 93 bytes ..Pgu_Imageog,Sharig-a-Dem,MicByTkaKig,SongByP,()DNP2002- 2005DNPgroup_mageSon,Sharing-a-Dream,M

5   Conlusions

In this paper, we have confirmed contactless detection capability using either stereo
or monaural microphones by our developed two kinds of watermark embedding
methods. Particularly, the extraction precision of our newly proposed monaural
method has been about 90% and superior to that of stereo embedding method. How-
ever, this monaural method has drawbacks of giving distortion to the right channel
signal in some cases and destructing panning positions in wider band range, which we
have to further evaluate and improve. And we are going to improve extraction preci-
sions with inserting error correction codes.

   In the future, we will implement our developed real-time watermark extraction
software in PDA or cell phone devices, and apply to a meta-deta retrieval system of
audio contents distributed by CD/DVD packages or network media. Furthermore, we

will develop an audio material tracability system, which will embed all of the audio production parameters into a mixed down signal and can identify all of the used sound material IDs from a mixed down music piece signal.

References

1. Ching-Te WANG,Tung-Shou CHEN, and Zhen-Ming XU: A Robust Watermarking System Based on the Properties of Low Frequency in Perceptual Audio Coding, IEICE Transactions, Vol.E85-A, No.6, pp. (Jun. 2002) 1257-1264.
2. Kirovski, D. and Malvar, H.S., Spread-spectrum watermarking of audio signals, Signal Processing, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.51, No.4, (2003) 1020-1033,.
3. Nakamura T., Katayama J., Yamamuro M. and Sonehara N.: Fast Watermark Detection Method from Analogue Images Using Cell Phone Camera, IEICE Transactions, Vol.J87-D2, No.12, (December, 2004) 2145-2155 (in Japanese).
4. Modegi T., Development of Contactless Sensible Watermark Embedding Technology for Audio Signals, Technical Report of IEICE, Vol.105, No.9, (April 2005) 19-24 (in Japanese).

SVG-Based Countermeasure to Geometric Attack

Longjiang Yu, Xiamu Niu, and Sheng-He Sun

Information countermeasure Technique Research Institute
Harbin Institute of Technology
`xiamu.niu@hit.edu.cn`

Abstract. In this paper, a novel application of SVG (Scalable Vector Graphics) is proposed in watermarking. Vector graphic is used to detect geometric transform and control geometric distortion. Experimental results show that the proposed method is efficient to solve geometric attack to watermarking.

1 Introduction

The success of the Internet introduces a new set of challenging problems regarding security. One of many issues that have arisen is the problem of copyright protection of electronic information. Digital watermarking has been proposed as a way to claim the ownership of the source and owner [1]. However, there are some open problems in digital watermarking. Geometric distortion is one of the most difficult problems. Possible solutions [2][3][4][5] have tried to solve this problem recently. But limited precision in geometric estimation is their significant deficiency. Alghoniemy and Tewfik introduced a RST (Rotation, Scaling and Translation) synchronization scheme based on the wavelet decomposition of an image [6]. Zhang and Sam used geometric moments to estimate geometric transform parameters and converted the distortion version to its original size and orientation [7]. The latter is more precise than the former in geometric estimation (error can be limited in 10^{-5}), which is the most precise method so far. But after all the best method can not give true transform parameters yet and unreliable when noise or crop exists. In this paper a novel method is proposed to incorporate watermarked image into SVG, a vector graph format. Parameters of geometric transform are easily acquired by indexing in SVG content. Since those parameters are acquired without any error, graphic distortion can be efficiently detected and controlled based on the method.

2 The Concept of SVG

SVG, abbr. Scalable Vector Graphics, is a new language for describing two-dimensional graphics in XML [8]. It has been an open industry standard in vector file format [9]. It could consist of images, text and vector graphics. The Document Object Model (DOM) of SVG inherits full XML DOM features, which allows for effective vector animation via scripting independent of platform. Common advantages of vector graphics over raster image are small capacity and no deformation in geometric transform. SVG has its own features as adaptability to interactive and dynamic opera-

tions. The image can be edited via scripting. Image content index is made easy. So SVG has rich potential for multimedia transmission and communication on the web. GIS (Geographic Information System) adopts SVG technology recently [10]. PDA (Personal Digital Assistant) perhaps is a next market to SVG [11]. The composition of SVG is shown in Fig.1.
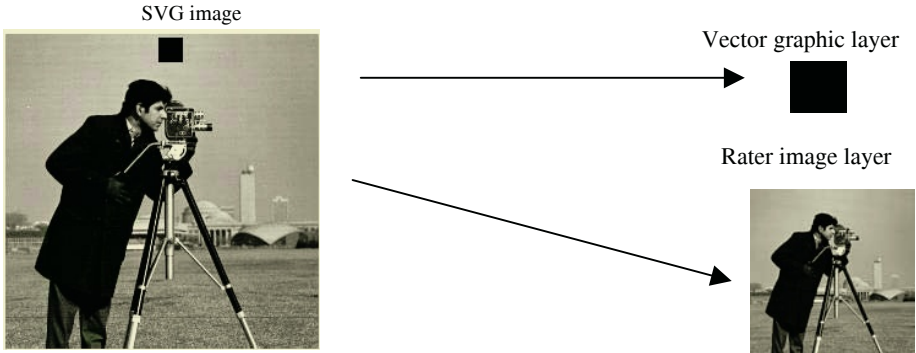
Fig. 1. SVG image composition

3   Geometric Distortion Detection and Control Method

Basic graphic pattern in raster image is searched as vector graphic layer in SVG. A connectivity region segmented algorithm is applied to the pattern searching which is quite simple. When grayscale of a pixel is different from four adjacent pixels respectively in upside, downside, left and right, the pixel is detected and removed. Connectivity region in the image is left after the operation. Then basic graphic pattern, such as rectangular, polygon, circle and line, is to be defined by a certain geometric shape template match respectively. In this paper rectangular pattern is selected and placed in vector graphic layer which represents as basic graphic pattern. A rectangular template of 2x2 scale is used for the pattern matching. The result image is shown in Fig.2.

After pattern is selected, raster image removes the corresponding region, which is to avoid pattern removing. If pattern in vector graphic layer is removed, the corresponding region of SVG image is empty causing rejection. When geometric transform happens, the parameters can be easily found in *transform* attribute.

For example:

<rect x="196" y="1" width="263" height="263" *transform*="scale(1.2,0.8), rotate(30), translate(-75,-60)" fill="none" stroke="black" *stroke-width*="1"/>

The above sentence paints a rectangular with width 263 and height 263, then scale as 1.2 in x direction, 0.8 in y direction, rotate 30 degree clockwise and translate –75 in x direction, -60 in y direction. Those transform parameters are the value of attribute *transform*, which is printed in italic form. So geometric distortion is easily detected without any error. Difference between vector graphic and raster image leads to displacement in pixel position, which is apparent to eyes and image quality degrades. So *stroke-width* of the rectangular is set to 1, seen in the above example statement which is printed in italic form this page, in order to compensate the difference. Rotate before

compensating is allowed as Fig.3(a). Perceptual distortion can be seen apparently. Rotate after compensating is shown as Fig.3(b). Part enlarged image is shown in Fig.3(c) and Fig.3(d) respectively to compare. By comparing it can be concluded that no distortion exists and geometric distortion in SVG is controlled.

raster image layer   vector graphic layer   selected pattern

pattern removed   SVG image
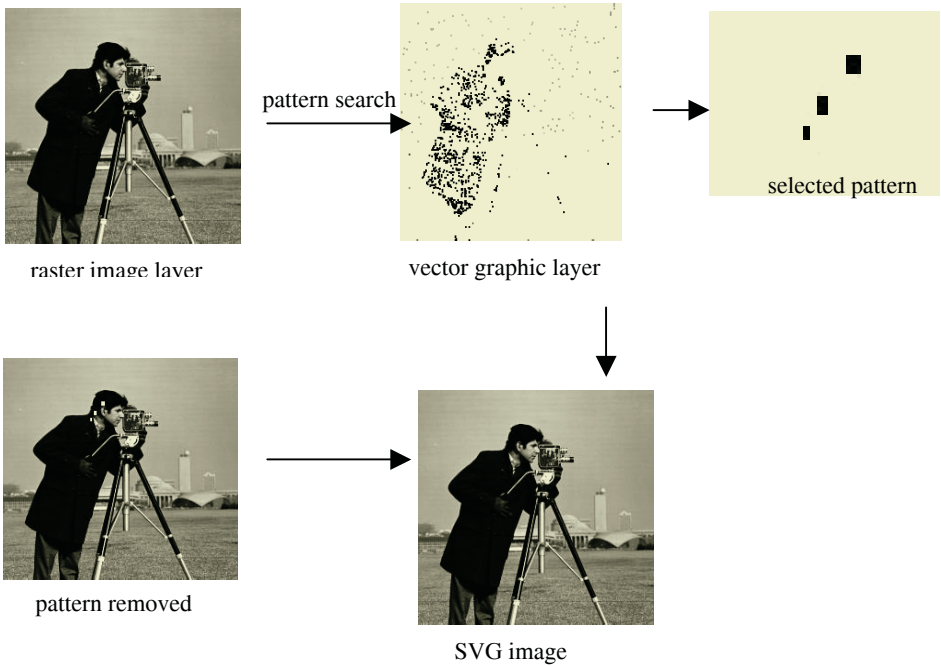
Fig. 2. Pattern searching process

4   Experimental Results

The 256x256 *cameraman* is chosen in experiments. Experimental result images can be viewed in Adobe SVG viewer 3.0 or Adobe Illustrator 10.0 and transformed in bitmap to be shown in this paper. Results verify this method can acquire the parameters of geometric transform and control geometric distortion successfully. A block-based DCT watermarking [12] is embedded in a rater image. Stirmark is used to perform geometric attack. The watermark is extracted successfully using the method proposed in this paper.

Acknowledgments

(a) rotate before compensating,

(b) rotate after compensating

(c) enlargement of (a)

(d) enlargement of (b)

Fig. 3. Geometric distortion control

References

1. Wolfgang, R. B. , Podilchuk, C. I. and Delp, E. J.: Perceptual Watermarks for Digital Images and Video, Proceedings of the IEEE, Vol. 87 (7), (1999) 1108-1126
2. Kutter, M.: Watermarking resisting to translation, rotation and scaling, Proc. SPIE, vol. 3528, (1998) 423–431
3. Pereira, S. and Pun, T.: Fast robust template matching for affine resistant image watermarking, International Workshop on Information Hiding, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, Vol. LNCS 1768, (1999) 200–210
4. Lin, C.-Y., Bloom, J. A., Cox, I. J., Miller, M. L. and Lui, Y. M.: Rotation-,scale-, and translation-resilient public watermarking of images, Proc.SPIE, (2000) 90–98
5. Davoine, F., Bas, P., Hebert, P.-A., and Chassery, J.-M.: Watermarking et résistance aux déformations géométriques, Coresa99. Sophia Antipolis, France, (1999)
6. Alghoniemy, M., Tewfik, A. H.: Geometric distortion correction in image watermarking, Proc. SPIE, (2000) 82–89

7. Zhang, Li, Kwong, S., Gang, Wei: Geometric moment in image watermarking**,** ISCAS '03, Proceedings of the 2003 International Symposium, Vol. 2, (2003) 932 –935.
8. Vaughan-Nichols, S.J.: Will vector graphics finally make it on the Web?, Computer, Volume: 34 Issue: 12 , (2001) 22 –24.
9. http://www.w3.org/TR/2003/REC-SVG11-20030114/. Scalable Vector Graphics (SVG) 1.1 Specification. W3C Recommendation (2003)
10. Baru, C., Behere, A., and Cowart, C.: Representation and display of geospatial information: a comparison of ArcXML and SVG, Web Information Systems Engineering, Proceedings of the Second International Conference on , Vol. 2, (2001) 48 -53
11. http://www.w3.org/TR/2003/REC-SVGMobile-20030114/. Mobile SVG Profiles: SVG Tiny and SVG Basic. W3C Recommendation (2003)
12. Koch, E., and Zhao, J.: Towards Robust and Hidden Image Copyright labeling, Proc. IEEE Workshop on Nonlinear Siganl and Image Processing, Los Alamitos, CA, 452-455

Watermarking Protocol Compatible with Secret Algorithms for Resisting Invertibility Attack

Xinpeng Zhang and Shuozhong Wang

School of Communication and Information Engineering, Shanghai University
Shanghai 200072, P.R. China
{xzhang,shuowang}@staff.shu.edu.cn

Abstract. Invertibility attack is a hostile measure to breach watermarking systems. In this paper, a novel watermarking protocol using a one-way hash function and a check of random watermarks is proposed in order to combat invertibility attacks. The described technique can be used in conjunction with any watermarking algorithm, no matter it is kept secret or made public, without resorting to a third party jury as required by some previous approaches. By introducing a set of reference sequences, segmentation of the digital information and iterative computation of watermarks, the protocol is further enhanced so that it can resist more sophisticated types of attack based on forging an illegitimate detector.

1   Introduction

Since mid-1990s digital watermarking as an effective means for intellectual property rights protection has attracted a great deal of research interests. Watermark is a digital code embedded imperceptibly and robustly into a multimedia signal, typically containing information about the owner, origin, status, and/or destination of the host material [1, 2]. With the development of watermarking techniques, various hostile attacks have emerged, such as geometric distortion attack, implementation attack, and protocol attack, which attempt to destroy watermarking systems in order to make illegal profits [3].

Craver et al. [4] proposed another smart protocol attack, invertibility attack. It does not aim at eliminating the embedded information or destroying synchronization, but demolish the very concept of the watermarking application itself. In invertibility attack, an attacker can create a counterfeit watermark in a digital product containing another legal mark to cause confusion therefore prevent rightful assertion of ownership when the attacker's watermarking algorithm is invertible. To combat this attack, non-invertible watermarking algorithms [5, 6] may be used since they cannot be abused as tools in invertibility attacks. Using these techniques, the structure of watermarking algorithms must be made public by submitting to an authoritative organization to check its non-invertibility. An alternative countermeasure is to design a special non-invertible watermarking protocol in conjunction with a traditional algorithm to defeat this attack even if the watermarking algorithm is invertible [7, 8]. However, a loophole still exists when a non-invertible algorithm or protocol is used. Having obtained any

particular watermarked multimedia data, an attacker can always tailor-make a counterfeit inserter and a corresponding detector, as well as a fake original and a workable key. In order to defeat such an attack, a secure watermarking protocol has been proposed in which correlation between the detecting algorithm and the embedded mark is imperative [9].

In the above-mentioned secure frameworks for resisting invertibility attack, a watermark hider must reveal his embedding/detecting scheme to an authorized jury to confirm its eligibility, i.e., the watermarking algorithm should be non-invertible or based on some normal mechanism, for example, QIM or spread spectrum method. As such, security of the embedded watermark will rely solely on the confidentiality of the key as stipulated by the Kerckhoff's principle universally observed in cryptography. In watermarking applications, however, this authorization procedure is rather cumbersome therefore lacks practicality. Without a general consensus within the watermarking community, a watermark hider may not wish to reveal his embedding/detecting algorithm to any third party, whether being an independent jury or the general public. The Kerckhoff's principle requires that an embedded watermark should be secure or cannot be removed even if the watermarking scheme is not a secret. It does not mean, however, that a watermarking algorithm must be revealed. In other words, watermarking is not necessary bound to the requirement of publicizing the marking mechanism. So, a watermarking technique capable of resisting invertibility attacks that is compatible with secret algorithms is certainly desirable.

However, a problem may occur in case the marking algorithm is kept secret. An attacker can forge an arbitrary mark-detector to declare presence of his fake watermark in a digital product since no one can open the black box to disclose the illegitimate mechanism. To resolve this problem, a secure watermarking protocol is proposed in this paper to prevent attackers from performing an invertibility attack by fabricating a fake original, a counterfeit key and a constructed secret algorithm. In this protocol, correlation between the detecting algorithm and the embedded mark is also necessary as in [9]. In addition to this common feature, the present framework also checks the detected results of random watermarks, and uses a set of reference sequences, segmentation of the digital product and iterative computation of marks. In this way, arbitrary forge of secret mark-detecting mechanisms by the attackers can be prevented.

2   Previous Works on Invertibility Attack

2.1   Invertibility Attack

Suppose that the copyright owner A embeds his watermark $\mathbf{W}_A$ into the host medium $\mathbf{I}$ with an inserter $\mathbf{E}_A$ and a key $\mathbf{K}_A$. A watermarked medium $\mathbf{I}_A$ is thus produced that is available to both the legitimate user and an attacker named B. Assume that B possesses a watermarking tool consisting of an inserter $\mathbf{E}_B$ and a detector $\mathbf{D}_B$. If B is able to construct a watermark $\mathbf{W}_B$ as well as a counterfeit key $\mathbf{K}_B$ and a fake data $\mathbf{I}'$ from $\mathbf{I}_A$ such that
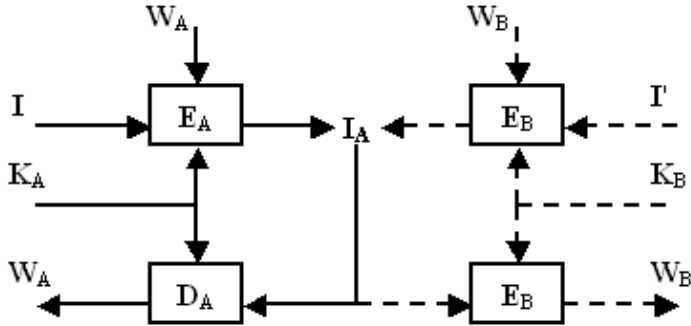
Fig. 1. Sketch of invertibility attack.

1. $\mathbf{E_B}(\mathbf{I'}, \mathbf{W_B}, \mathbf{K_B}) = \mathbf{I_A}$,
2. $\mathbf{D_B}(\mathbf{I_A}, \mathbf{W_B}, \mathbf{K_B}) = 1$, (here 1 means "watermarked", and 0 "not watermarked") or $\mathbf{D_B}(\mathbf{I_A}, \mathbf{K_B}) = \mathbf{W_B}$, and
3. $\mathbf{I'}$ is similar to $\mathbf{I_A}$,

the watermarking system $(\mathbf{E_B}, \mathbf{D_B})$ is said to be invertible. In the above operation, both the authentic watermark $\mathbf{W_A}$ and the counterfeit mark $\mathbf{W_B}$ can be detected from $\mathbf{I_A}$ using $\mathbf{D_A}$ and $\mathbf{D_B}$, respectively. Therefore, the rightful owner of $\mathbf{I_A}$ cannot be identified based on the extracted watermark. Figure 1 is a sketch of invertibility attack, in which the two sides of $\mathbf{I_A}$ are symmetric, the left half being a legal watermarking procedure and the right half an illegal procedure counterfeited by the attacker.

In addition, a similar term "quasi-invertibility" can also be defined [4]. Quasi-invertibility is less stringent than full-invertibility therefore more harmful to legitimate watermarking system users. In the following discussion, quasi-invertibility and full-invertibility are not distinguished and will simply be referred to as "invertibility".

2.2   Non-invertible Algorithm and Protocol

To defeat invertibility attack, Craver et al. suggested that non-invertible watermarking techniques be used. In their approach, all legitimate watermarking schemes must be designed to be noninvertible. A one-way hash function or public-key cryptology is employed for this purpose. For example, let the watermark signal be a hash output of the host data [5], $\mathbf{W} = h(\mathbf{I})$. The host data are needed in detection. This way, the attacker can no longer forge a counterfeit watermark and counterfeit host data. In another method [6], the original medium is not needed in detection. The watermark is a hash of pseudorandom data selected by the digital product owner. As a key, the pseudorandom data is used to compute the watermark in detection. Since it is almost impossible to derive a counterfeit watermark $\mathbf{W_B}$ and counterfeit host data $\mathbf{I'}$ from the watermarked data $\mathbf{I_A}$, invertibility attack is virtually impossible. In this case, the mechanism of watermarking algorithm should be made public or submitted to an authorized organization to prove its non-invertibility.

An alternative method for resisting invertibility attack is to design some special watermarking protocol. The protocol presented in [7] can be used in conjunction with traditional algorithms based on addition of a pseudo-random sequence. Another protocol [8] is applicable to any watermarking algorithm regardless of being non-invertible or not, with which a watermark hider must embed, in addition to the watermark itself, the digital signatures of the original product and the mark into the digital content. Given a claimed watermarked copy $\mathbf{I}_A$, the signatures should be verified after all the embedded data are extracted. Note that the asymmetry is derived from the watermarking protocol, which uses a signature mechanism, but not from the embedding/detecting algorithms. If the attacker's watermarking algorithm $(\mathbf{E}_B, \mathbf{D}_B)$ is given, it will be very difficult to find a fake original $\mathbf{I}'$, a fake mark $\mathbf{W}_B$ and a fake key $\mathbf{K}_B$ to satisfy every aspect of this protocol, though the watermarking algorithm may be invertible. In this case, the watermarking mechanism must also be revealed to show it is based on a normal technique, such as QIM or spread spectrum method. Otherwise, an attacker can arbitrarily create a detected result that serves his purpose since the marking tool is a black box.

2.3 Invertibility Attack Based on Forged Algorithm and a Countermeasure

Although a non-invertible protocol or a rule requiring that all legitimate watermarking tools must be non-invertible is used, a pirate, having obtained a single watermarked product, can tailor-make an inserter $\mathbf{E}_B$, a corresponding detector $\mathbf{D}_B$, as well as a fake original, a counterfeit mark and a workable key to perform an invertibility attack [9]. When the invertibility of marking algorithm is required, the attacker can publish the mechanism of his forged watermarking algorithm $(\mathbf{E}_B, \mathbf{D}_B)$ to show its non-invertibility, therefore declare the legitimacy of the fake mark $\mathbf{W}_B$ in $\mathbf{I}_A$. When a non-invertible protocol is used, the attacker can also obey the protocol and publish $(\mathbf{E}_B, \mathbf{D}_B)$ like a normal algorithm. So, this invertibility attack based on forged algorithm is more powerful than the basic invertibility attack.

In order to combat this attack, a secure protocol has been proposed in which the embedded mark signal is derived from the detecting software $\mathbf{D}$ and an embedding key using a one-way hash function [9]. By establishing correlation between the embedded signal and the watermarking algorithm, an attacker cannot obtain a suitable mark and a workable key when he constructs a counterfeit watermarking algorithm. Thus the forged-algorithm attack can no longer succeed. Similarly, the mechanism of marking algorithm should also be publicized to show its embedding technique as normal.

3   A Basic Protocol Compatible with Secret Algorithms

In the previous anti-invertibility-attack strategies, the requirement is too stringent for most watermarking applications since all watermarking algorithms must

be checked by an authorized organization to prove its non-invertibility or normality. In practice, watermarking system users are usually reluctant to reveal the marking procedure to any outsider. Furthermore, it is difficult to define "normality" of a watermarking algorithm. It is also possible that an attacker can design an invertible or abnormal algorithm that may be mistakenly judged as non-invertible or normal to cheat the third-party jury.

Thus, a more convenient and user-friendly watermarking framework capable of resisting invertibility attacks without the need of revealing the embedding and detecting algorithms is desired. With this system, although revelation of the marking algorithm is not compulsory, the watermark hider, to conform to the Kerckhoff's principle, should still ensure that an embedded mark cannot be removed even if the watermarking mechanism is known to attackers.

Without an authorized organization, an attacker, B, can arbitrarily construct not only a counterfeit mark $\mathbf{W_B}$ and a fake key $\mathbf{K_B}$, but also his own watermarking algorithm $(\mathbf{E_B}, \mathbf{D_B})$, in which the detector $\mathbf{D_B}$ can be considered a black box. Specifically, the forged detector $\mathbf{D_B}$ may have the same input-output relation as an authentic detector but with an entirely different internal structure. Assuming that inputs to the detector $\mathbf{D_B}$ are a received cover $\mathbf{I_I}$, a key $\mathbf{K_I}$ and a watermark $\mathbf{W_I}$, and output from $\mathbf{D_B}$ is "1" or "0", the attacker can forge an detecting algorithm as follows to confuse the copyright of a digital media $\mathbf{I_A}$:

If $\mathbf{I_I}$ is similar to $\mathbf{I_A}$, $\mathbf{W_I} = \mathbf{W_B}$ and $\mathbf{K_I} = \mathbf{K_B}$
$\qquad \mathbf{D_B}(\mathbf{I_I}, \mathbf{K_I}, \mathbf{W_I}) = 1$
Else
$\qquad \mathbf{D_B}(\mathbf{I_I}, \mathbf{K_I}, \mathbf{W_I}) = 0$

Although the output result is directly derived from the inputs, no one can find the illegal mechanism.

To prevent anyone from carrying out such an attack, we propose a novel watermarking protocol, which specifies the way of watermark generation, but not the internal structure of watermarking algorithm. Here we view any mark-detecting software as a binary sequence, although its actual structure is unknown. The proposed watermarking protocol employs a one-way hash function h and a concept of *key space*. The key space includes a large number of candidate keys, say, 2^{128} binary sequences, each having a length of 128. After selecting a key from the key space according to the protocol, the rightful owner, A, produces a mark signal by using h, his key, and a binary sequence corresponding to his detector, and embeds them into the cover data. On the detection side, the mark signals derived from both the key provided by A and other keys randomly selected from the key space, are used as the detector inputs for the judgment as to whether or not A is a rightful owner.

The key of this protocol is to establish correlation between the watermarking algorithm, the embedding key, and the mark signal, and to check the detection results of marks derived both from a given key and from randomly selected keys, so that any attacker cannot find a workable key, a counterfeit mark and a constructed secret algorithm simultaneously. Details of the proposed protocol are described as follows:

1. The protocol employs a one-way hash function h that is made public.
2. Embedding procedure: The detecting software $\mathbf{D}$ can be considered and treated as a binary sequence $\mathbf{S_D}$. The embedded mark signal $\mathbf{W}$ must be determined by $\mathbf{S_D}$ and $\mathbf{K}$, that is, $\mathbf{W} = h(\mathbf{S_D}, \mathbf{K})$, where the key $\mathbf{K}$ is selected by the copyright owner from the key space. Thus, the watermarked medium $\mathbf{I_A}$ is generated: $\mathbf{I_A} = \mathbf{E_A}$ $(\mathbf{I}, \mathbf{K_A}, \mathbf{W_A}) = \mathbf{E_A}$ $[\mathbf{I}, \mathbf{K_A}, h(\mathbf{S_{D_A}}, \mathbf{K_A})]$.
3. Detecting procedure: Suppose a consumer, C, has obtained the digital product $\mathbf{I_A}$, and wants to know who is the rightful owner of $\mathbf{I_A}$. If someone claims that he is the rightful owner, he must provide his detector $\mathbf{D}$, but no its mechanism, and his key $\mathbf{K}$ for the one-way hash function h. After obtaining $\mathbf{W}$ using h, he must also provide his key $\mathbf{K}$ and the signal $\mathbf{W}$ for his detector $\mathbf{D}$ to show that $\mathbf{D}(\mathbf{I_A}, \mathbf{K}, \mathbf{W}) = 1$. The procedure should be under C's supervision. The claimer may keep his key secret but must prove that the keys entered into the hash function and the detector $\mathbf{D}$ are identical. For example, he may store his key in an IC card and insert the card into a special device to provide the content of his key. Furthermore, he must randomly select many, say, 100, other keys $\mathbf{K}'$ in the key space to compute the derived signals $\mathbf{W}'$ using h, and checks if $\mathbf{D}[\mathbf{I_A}, \mathbf{K}, h(\mathbf{S_D}, \mathbf{K}')] = 0$. The claimer must also prove that the keys $\mathbf{K}$ entered into the detector $\mathbf{D}$ now are the same as the previous one. If the above procedure is successfully completed, the claim of copyright is accepted.

Here, the purpose of checking whether or not $\mathbf{D}[\mathbf{I_A}, \mathbf{K}, h(\mathbf{S_D}, \mathbf{K}')]$ equals zero is to prevent anyone from forging a detector in which the output is always 1 when using his key. The purpose of proving identity of the two entered keys is to verify the prescribed correlation between the key, the algorithm, and the mark signal.

Using this protocol, the rightful owner A can provide his legitimate key $\mathbf{K_A}$ and the detector software $\mathbf{D_A}$ to satisfy the above requirements. On the other hand, even if the attacker B can construct a watermarking algorithm $(\mathbf{E_B}, \mathbf{D_B})$ and a forged mark $\mathbf{W_B}$ with $\mathbf{D_B}(\mathbf{I_A}, \mathbf{K_B}, \mathbf{W_B}) = 1$, he cannot satisfy $\mathbf{W_B} = h(\mathbf{S_{D_B}}, \mathbf{K_B})$ so that fails to pass himself off as a copyright owner.

However, since the internal structure of the detecting software $\mathbf{D}$ can be arbitrary, that is, $\mathbf{D}$ is treated as a black box, the attacker B can define an arbitrary mapping between the input and output at his disposal using any possible method. This leaves some flaws for more sophisticated attacks. Therefore the protocol needs to be further enhanced as described in the following sections.

4 Scenarios of Possible Attacks

Although performing invertibility attacks by constructing counterfeit watermark and watermarking algorithm is difficult with the above-described protocol, more sophisticated attacks are still possible. Two different kinds of possible attacks are described in the following.

4.1   First Type of Possible Attack: Hash-Based Detector Attack

The first type of attack employs the identical one-way hash function h in the forged detector $\mathbf{D_B}$ so that correlation between a fake key, a counterfeit mark, and a workable secret algorithm can be established. Suppose that inputs to $\mathbf{D_B}$ are a received cover $\mathbf{I_I}$, a key $\mathbf{K_I}$, and a watermark signal $\mathbf{W_I}$, and the output is 1 or 0. An attacker B selects a key $\mathbf{K_B}$ from the key space as his counterfeit key and constructs the inside mechanisms of his forged detector $\mathbf{D_B}$ as follows.

   If $\mathbf{I_I}$ is similar to $\mathbf{I_A}$, and $\mathbf{W_I} = h(\mathbf{S_{D_B}}, \mathbf{K_I})$

   $\quad\mathbf{D_B}(\mathbf{I_I}, \mathbf{K_I}, \mathbf{W_I}) = 1$

   Else

   $\quad\mathbf{D_B}(\mathbf{I_I}, \mathbf{K_I}, \mathbf{W_I}) = 0$

Here, the detector treats itself as a binary sequence $\mathbf{S_{D_B}}$ and computes the hash of $\mathbf{S_{D_B}}$ and $\mathbf{K_B}$. So, the attacker can claim that he satisfies the protocol as described in the previous section. After providing his counterfeit key $\mathbf{K_B}$ and the forged detector $\mathbf{D_B}$ for the one-way hash function h to yield $\mathbf{W_B}$, he input $\mathbf{K_B}$ and $\mathbf{W_B}$ into the forged detector $\mathbf{D_B}$, output of the detector will be 1, while other mark signals derived from the keys randomly selected in the key space will produce 0. Thus the basic protocol is defeated.

4.2   Second Type of Possible Attack:
Forged Detector Attack Based on Watermark Subset

With the protocol proposed in Section 3, for the given legitimate key $\mathbf{K_A}$, $\mathbf{D_A}[\mathbf{I_A}, \mathbf{K_A}, h(\mathbf{S_{D_A}}, \mathbf{K_A})]$ must be 1, and for any other randomly selected key $\mathbf{K'}$, $\mathbf{D_A}[\mathbf{I_A}, \mathbf{K_A}, h(\mathbf{S_{D_A}}, \mathbf{K'})]$ must be 0. Assume that $\mathcal{W}$ is the entire set of $\mathbf{W}$. The attacker can arbitrarily define a set $\mathcal{W}_1$ that is a small subset of $\mathcal{W}$, and construct the internal structure of a fake detector $\mathbf{D_B}$ as follows.

   If $\mathbf{I_I}$ is similar to $\mathbf{I_A}$, and $\mathbf{W_I} \in \mathcal{W}_1$

   $\quad\mathbf{D_B}(\mathbf{I_I}, \mathbf{K_I}, \mathbf{W_I}) = 1$

   Else

   $\quad\mathbf{D_B}(\mathbf{I_I}, \mathbf{K_I}, \mathbf{W_I}) = 0$

A counterfeit key $\mathbf{K_B}$ can be found using a method of exhaustive enumeration so that $h(\mathbf{S_D}, \mathbf{K_B})$ belongs to $\mathcal{W}_1$. In this way, the attacker can cheat the customer by providing a fake key $\mathbf{K_B}$.

   Cost of this type of attack is estimated as follows. Assume that the ratio between the sizes of $\mathcal{W}_1$ and $\mathcal{W}$ is p, the average number of attempts for finding the first useful counterfeit key $\mathbf{K_B}$ is $1/p$ because

$$\sum_{j=1}^{\infty} p\,(1-p)^{(j-1)}\,j = \frac{1}{p} \tag{1}$$

On the detection side, a consumer randomly selects other N keys to check for $\mathbf{D_B}[\mathbf{I_A}, \mathbf{K_B}, h(\mathbf{S_{D_B}}, \mathbf{K'})] = 0$. The probability of satisfying the condition, Q, can be obtained:

$$Q = (1 - p)^N \tag{2}$$

This indicates that the attacker has constructed a workable detector and found a fake key, leading to a successful attack.

When p is very close to zero,

$$p \approx \frac{1 - Q}{N} \tag{3}$$

For example, if $p=0.0001$, a fake key can be found after, in average, 10000 attempts. In this case, if the consumer makes $N=100$ checks, he will accept the illegal claim of the attacker with a probability $Q=99\%$. Since a large N is unrealistic, the cost of attack is not a problem for the attacker.

5   Protocol Enhancement Against Sophisticated Attacks

Since the two types of attack as described in the previous section form a threat to the basic protocol proposed in Section 3, we introduce further enhancement in order to provide additional resistance against possible attacks by using a forged detector.

5.1   Enhancement Against Hash-Based Attack

The core of the hash-based attack is that the fake detector can be treated as a binary sequence by itself, and the one-way hash function h can be used to compute a watermark signal in a forged detector. In order to prevent this, a set of reference sequences with different lengths is introduced into the protocol. The reference sequences are made of random numbers containing little redundant information, and therefore are hard to be compressed. And these reference sequences are made available to the public. In the enhanced protocol, the embedded watermark should be the hash of the sequence $\mathbf{S_D}$, the key, and in addition, the reference. In other words, $\mathbf{W} = h(\mathbf{S_D}, \mathbf{K}, \mathbf{R})$. The reference sequence $\mathbf{R}$ is the shortest among all reference sequences that are longer than the binary sequence $\mathbf{S_D}$ corresponding to $\mathbf{D}$.

When someone claims his copyright, he must obtain $\mathbf{W}$ by providing the key $\mathbf{K}$, the detector $\mathbf{D}$, and the reference $\mathbf{R}$ to h, then enter $\mathbf{W}$, together with $\mathbf{K}$ and $\mathbf{I_A}$, to the detector $\mathbf{D}$ in an environment isolated from the reference, as shown in Figure 2. The consumer should check an output "1" from $\mathbf{D}$. With other mark signal derived from any arbitrarily selected key, however, the result is always "0". It should be noted that the reference $\mathbf{R}$ cannot be included in the detector $\mathbf{D}$, since it is longer than $\mathbf{S_D}$. Therefore, a hash $h(\mathbf{S_D}, \mathbf{K}, \mathbf{R})$ cannot be computed in a forged detector when the detector is isolated from $\mathbf{R}$. Thus, the first type of the detector attack is defeated.

5.2   Enhancement Against Watermark Subset Based Attack

To deal with the above-described second type of attack, an additional mechanism is introduced to make the cost of the attack too high to be feasible. The host
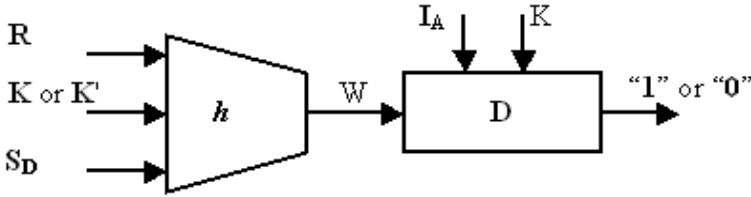
Fig. 2. Structure of the proposed watermarking protocol in which a reference **R** is added to prevent construction of a fake detector with the one-way hash.

medium **I** is first segmented into M sections, $\mathbf{I}_1$, $\mathbf{I}_2$, ..., $\mathbf{I}_M$ in a way prescribed in the enhanced protocol. And the enhanced protocol also prescribes $\mathbf{W}_{A,1} = h(\mathbf{S}_{\mathbf{D}_A}, \mathbf{K}_A)$, and $\mathbf{W}_{A,i} = h(\mathbf{S}_{\mathbf{D}_A}, \mathbf{W}_{A,i-1})$ where $i= 2, 3, \ldots, M$, so that all $\mathbf{W}_{A,i}$ form a chain. The marked sections are generated according to $\mathbf{I}_{A,i} = \mathbf{E}_A(\mathbf{I}_i, \mathbf{K}_A, \mathbf{W}_{A,i})$ where $i = 1, 2, \ldots, M$. Finally, $\mathbf{I}_{A,i}$ are combined to yield the watermarked result $\mathbf{I}_A$.

When someone claims his copyright, he should produce $\mathbf{W}_1 = h(\mathbf{S}_{\mathbf{D}}, \mathbf{K})$ and $\mathbf{W}_j = h(\mathbf{S}_{\mathbf{D}}, \mathbf{W}_{j-1})$ $(j > 1)$, and show $\mathbf{D}(\mathbf{I}_{A,i}, \mathbf{K}, \mathbf{W}_i) = 1$ for each section of $\mathbf{I}_A$ under supervision $(i= 1, 2, \ldots, M)$. With other $\mathbf{W}'_i$ derived from randomly selected keys, $\mathbf{D}(\mathbf{I}_{A,i}, \mathbf{K}, \mathbf{W}'_i)$ should be 0.

In this case the attacker B may still forge a detector $\mathbf{D}_B$ to carry out the second type of fake-detector based attack as follows.

If $\mathbf{I}_I$ is similar to $\mathbf{I}_{A,i}$, and $\mathbf{W}_I \in \mathcal{W}_i$
$\mathbf{D}_B(\mathbf{I}_I, \mathbf{K}_I, \mathbf{W}_I) = 1$
Else
$\mathbf{D}_B(\mathbf{I}_I, \mathbf{K}_I, \mathbf{W}_I) = 0$

Here, $\mathcal{W}_i$ are subsets of $\mathcal{W}$. As previously explained, the attacker must find a suitable $\mathbf{K}_B$ satisfying $\mathbf{W}_{B,1} = h(\mathbf{S}_{\mathbf{D}_B}, \mathbf{K}_B) \in \mathcal{W}_1$ and $\mathbf{W}_{B,i} = h(\mathbf{S}_{\mathbf{D}_B}, \mathbf{W}_{B,i-1}) \in \mathcal{W}_i$ where $i= 2, 3, \ldots, M$. Assume that all $\mathcal{W}_i$ have the same size, and the ratio between this size and that of $\mathcal{W}$ is p. Similar to Equation 1, the average number of attempts in finding the first useful counterfeit key $\mathbf{K}_B$ is p^{-M}. If a consumer uses N randomly selected keys in the key space to derive the mark chains and check each $\mathbf{I}_{A,i}$, the probability of satisfying the conditions $\mathbf{D}(\mathbf{I}_{A,i}, \mathbf{K}_B, \mathbf{W}'_i) = 0$ $(i= 1, 2, \ldots, M)$ will be:

$$Q = (1-p)^{MN} \tag{4}$$

Then, Q is the probability of being cheated by the attacker. The parameter p must be very close to 0 for a sufficiently large Q. Thus,

$$p \approx \frac{1-Q}{MN} \tag{5}$$

In this case, the cost of attack increases exponentially with M, therefore the attack scheme becomes infeasible. For example, when $Q = 50\%$, $N = 10$, $M = 10$, the average number of attempts is greater than 10^{21}.

5.3   Enhanced Protocol

The integrated watermarking protocol against possible invertibility attacks is summarized as follows.

1. The protocol employs a one-way hash function h and a set including many reference sequences of different lengths that are made public. These reference sequences contain very little redundancy.
2. The legitimate owner of a digital product selects a key $\mathbf{K_A}$ and a reference sequence $\mathbf{R}$ that possesses the smallest length among the existing sequences with a length greater than the binary array $\mathbf{S_{D_A}}$ corresponding to $\mathbf{D_A}$. $\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_M$ are calculated according to $\mathbf{W}_1 = h(\mathbf{S_{D_A}}, \mathbf{K_A}, \mathbf{R})$, $\mathbf{W}_i = h(\mathbf{S_{D_A}}, \mathbf{W}_{i-1})$ where $i= 2, 3, \ldots, M$.
3. The host medium $\mathbf{I}$ is segmented into M sections $\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_M$ in a prescribed way. Watermarked sections $\mathbf{I}_{A,i} = \mathbf{E_A} \, (\mathbf{I}_i, \mathbf{K_A}, \mathbf{W}_i)$ are generated accordingly, and then combined to produce $\mathbf{I_A}$.
4. When someone claims that he is the rightful owner of $\mathbf{I_A}$, he must compute $\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_M$ using his detector $\mathbf{D}$, his key $\mathbf{K}$, a reference sequence $\mathbf{R}$ and the function h. For each $\mathbf{I}_{A,i}$, the results $\mathbf{D}(\mathbf{I}_{A,i}, \mathbf{K}, \mathbf{W}_i) = 1$ and $\mathbf{D}(\mathbf{I}_{A,i}, \mathbf{K}, \mathbf{W}'_i) = 0$ should be checked where $\mathbf{W}'_i$ is derived from one of other N randomly selected keys. Here, the detecting software must be run in an environment isolated from the reference to ensure that it cannot make use of the reference sequences. The procedure of Step 4 should be under supervision and the copyright claimer may keep his key secret but must prove that the keys entered into the hash function and the detector $\mathbf{D}$ are identical.

6   Conclusion

A novel watermarking protocol capable of resisting invertibility attacks is introduced. Unlike the other methods subject to non-invertibility limitations, the proposed approach is applicable to any watermarking algorithm, whether the embedding scheme is kept secret or made public, and whether the algorithm is invertible or noninvertible. Using this protocol, a third party jury is not needed. The only additional requirements are that the embedded watermark signal should be generated in line with the protocol, and a series of detected results are used to decide the ownership of a digital product.

In the described protocol, the watermark signal is produced according to a key, a detector, as well as a set of reference sequences using a specified one-way hash function. Consequently, it is virtually impossible to generate an illegal watermark in a forged detector with a counterfeit key. In addition, by embedding several iteratively generated watermark signals into segmented medium data, the price for producing illegal detectors and keys is prohibitively high. In this way, invertibility attacks and possible attacks based on forged detector are defeated.

Acknowledgments

References

1. Petitcolas, F.A.P., Anderson, R.J., Kuhn M.G.: Information Hiding–A Survey. Proc. IEEE. **87** (1999) 1062–1078
2. Hartung, F., Kutter, M.: Multimedia Watermarking Techniques. Proc. IEEE. **87** (1999) 1079–1107
3. Voloshynovskiy, S., Pereira, S., Pun, T., Eggers, J., Su, J.: Attacks on Digital Watermarking: Classification, Estimation-Based Attacks, and Benchmarks. IEEE Communications Magazine. (Aug. 2001) 118–126
4. Craver, S., Menon, N., Yeo, B. L., Yeung, M.: Resolving Rightful Ownerships with Invisible Watermarking Techniques: Limitations, Attacks, and Implications. IEEE Journal on Selected Areas of Communications. **16** (1998) 573–586
5. Swanson, M. D., Zhu, B., Tewfik, A. H., Boney. L.: Robust Audio Watermarking Using Perceptual Masking. Signal Processing. **66** (1998) 337–355
6. Hwang, M., Chang, C., Hwang, K.: A Watermarking Technique Based on One-Way Hash Functions. IEEE trans. on Consumer Electronics. **45** (1999) 286–294
7. Li, Q., Chang, E.: On the Possibility of Non-Invertible Watermarking Schemes. 6th International Workshop on Information Hiding, Lecture Notes in Computer Science, **3200** (2004) 13–24
8. Katzenbeisser, S., Veith, H.: Securing Symmetric Watermarking Schemes against Protocol Attacks. Security and Watermarking of Multimedia Contents IV, Proc. SPIE. **4675** (2002) 260–268
9. Zhang, X., Wang, S.: Invertibility Attack against Watermarking Based on Forged Algorithm and a Countermeasure. Pattern Recognition Letters. **25** (2004) 967–973

System Architecture Analysis
of a Hybrid Watermarking Method

Chaw-Seng Woo[1], Jiang Du[1], Binh Pham[2], and Hamud Ali Abdulkadir[2]

[1] Information Security Institute, Faculty of Information Technology
Queensland University of Technology, GPO Box 2434, Brisbane, QLD4001, Australia
`cs.woo@student.qut.edu.au, j2.du@qut.edu.au`
[2] Faculty of Information Technology, Queensland University of Technology
GPO Box 2434, Brisbane, QLD4001, Australia
`b.pham@qut.edu.au, h.abdulkadir@student.qut.edu.au`

Abstract. A hybrid watermark that consists of a robust part and a fragile part can be used to serve multiple purposes. The robust part can protect copyright information, the fragile part can detect tampering, and their combination enables identification of attacks encountered. This paper analyses an overlap and a non-overlap implementation of the robust and fragile parts in a hybrid system. The difference between the two implementation methods lies in the robust and fragile watermarks embedding positions. Embedding capacity, computational costs, watermark robustness, and tamper detection localization of the two implementations are analyzed. In addition, optimization issues of block size in the hybrid system are discussed.

1   Introduction

To date, hybrid watermarking methods published are limited compared to other multimedia content watermarking methods. Some of the recent works are [1], [2], and [3]. A hybrid watermarking system that consists of a robust part and a fragile part can be used to serve multiple purposes. For example, robust watermarks are suitable for copyright protection because they remain intact with the protected content under various manipulative attacks. Fragile watermarks are good for tamper detection since it is sensitive to changes. The combination of robust and fragile watermarks offers some advantages over single watermarks. For example, it can identify copy attack and substitution attack [2]. In addition, watermark detection results of a hybrid system can be used to deduce whether a malicious tampering or a common image processing operation has taken place [1].

The performance factors of a watermarking method are mutually exclusive. For instance, increasing watermark robustness normally degrades its imperceptibility and limits its embedding capacity. To achieve a desirable balance among the performance factors, a designer must understand the influence of one factor on another. Therefore, it is important to evaluate the effects of system architecture on a hybrid system.

We analyzed the system architecture of a hybrid digital image watermarking system, and compared two implementation methods of its robust and fragile parts. The first method ensures that robust and fragile watermarks are embedded in non-

overlapping positions [2], and will be called "non-overlap" implementation. The second method overlaps both watermarks, and will be called "overlap" implementation. The overlap implementation has the advantage of full embedding capacity and higher localization in tamper detection. However, the compromise in its robustness and computational cost need to be investigated. The comparison include watermark embedding capacity, computational cost, robustness of the robust part, and tamper detection effectiveness of the fragile part. In addition, the effects of block size on the hybrid system's performance are also studied.

2   Overview of the Hybrid Watermarking System

The hybrid system chosen in our analysis [2] embeds a periodic robust watermark pattern in the discrete wavelet transform (DWT) domain. It uses robust estimation technique with superior performance to enable watermark detection. Thus, it is suitable for real life application where the cover image may not be available during watermark detection. It also has a state-of-the-art fragile part that embeds watermark blocks in the least significant bits (LSB) of pixels. The following two paragraphs describe the robust and fragile parts. Detail steps can be found in [2].

The robust part uses a self-reference method to recover from geometrical distortions. Firstly, the watermark message is encoded using an error correction code (ECC) for reliable decoding, and encrypted for confidentiality. Secondly, the message bits are spread in a symmetric pattern to cover the whole image size. This provides regularly-spaced peaks in geometrical re-synchronization for watermark detection. Finally, the watermark is embedded in DWT domain for robustness. Instead of employing human visual system (HVS) masking for imperceptibility, we simplified it with constant energy embedding. To detect a watermark in an attacked image, it exploits the periodic peaks of magnitude spectrum for image re-synchronization. These steps are detailed in [4]. We applied thresholding on the magnitude spectrum to extract the peaks for simplicity. Fig. 1 (Left) depicts peaks obtained from the magnitude spectrum of the embedded watermark. Fig. 1 (Right) shows peaks extracted from a rotated stego image. Assuming local distortions are restricted by the acceptable image quality change, local and non-linear transformations can be recovered using the same approach at the local level. Details of such approaches are described in [5]. After that,
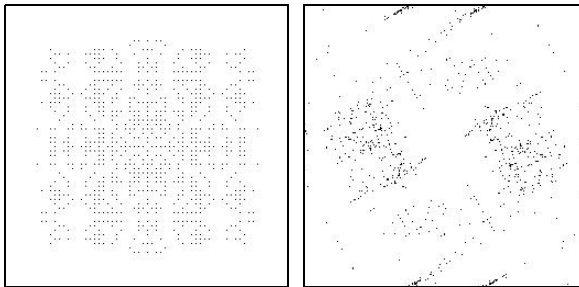
Fig. 1. (Left) Peaks obtained from the magnitude spectrum of the embedded watermark. (Right) Peaks extracted from a stego image with 30 degree rotation and auto-crop

a watermark estimation based on *Maximum a Posteriori* (MAP) probability is applied on the re-synchronized image. Then, a correlator detector is used in watermark decoding with a threshold value.

The fragile part uses a block-wise scheme to locate tampered regions. It computes a key-dependant hash value for overlapping blocks of an image and embeds the value into the LSB of pixels inside that block. By comparing the estimated signatures of the fragile blocks, tampered regions can be highlighted.

3   Non-overlap Implementation of the Hybrid System

In the non-overlap implementation, the embedding of its robust and fragile parts is performed simultaneously as described in Section 2 above. The robust positions do not overlap with the fragile positions within each block. Hence it was named "orthogonal" in [2]. The detection of watermarks in the non-overlap implementation is the same as its embedding part where the robust and fragile parts are processed independently. In our implementation, the robust blocks and fragile blocks are chosen to have the same size.

The block-wise hashing of fragile part takes the current block with its eight neighboring blocks as input. The computed hash code is then embedded into the current block. This provides local contextual dependency. However, this approach not only detects modifications within the block but also modifications in its neighboring blocks. Compensation steps mentioned in [2] are not implemented at this stage.

4   Overlap Implementation of the Hybrid System

In this implementation, the robust part is embedded prior to the fragile part. By definition, the robust part must survive distortions caused by the fragile part. Therefore, we proposed to embed the fragile part in all positions, overwriting the LSBs of the robust stego image. As a result, both the robust and fragile parts can be embedded into all positions, achieving maximum watermarking capacity. Furthermore, it gives the highest possible localization for tamper detection. In addition, it also reduces computation by eliminating position tracking of the robust and fragile parts.

The watermark detection in the overlap implementation is similar to those in the non-overlap implementation. However, all pixel locations in overlap implementation are processed because both the robust and fragile parts are embedded in every position. This requires an examination of the compromise in computational cost.

5   Experimental Results Analysis

To compare the overlap and non-overlap implementations in a hybrid system, the parameters listed in Table 1 were applied. Two test images with 256 gray levels were used. They are *Lena* and *Cameraman* of 256×256 pixels. A set of general image manipulation operations listed in Table 2 was used to evaluate the performance of the robust watermark. Three types of attacks were used in fragile watermark evaluation:

local tampering modify the pixel values of a small area, copy attack copies a small region from an image and pastes it onto the same image whereas collage attack pastes it onto another image.

Table 1. Parameter values for non-overlap and overlap implementations

Parameter	Non-overlap implementation	Overlap implementation
Block size $t_1 \times t_2$	16×16 = 256 pixels	16×16 = 256 pixels
Robust positions	178	256
Fragile positions	42	256
Empty positions	36	0

Table 2. Attacks for robustness evaluation

Attack	Descriptions
Rotation followed by cropping	Rotate 30 degree with auto-crop
Scaling followed by shearing	Uniform scaling factor 0.98; Shear 2%
JPEG compression	Quality factors at 50%
Gaussian noise insertion	Zero mean; variance 0.02
Contrast adjustment	Gamma value 0.6
Median filtering	2×2 smoothening kernel

5.1   Analysis of Robust Part Results

Using block size of 32×32, the stego images of non-overlap and overlap implementations give PSNR of 37.02dB and 37.29dB respectively. This indicates image qualities of both implementations are very close. Such observation can be explained by the small difference of un-marked positions between the two implementations, i.e. 256 − 178 = 78 bits ≈ 30.47% in each block.

Regularly-spaced peaks can be observed after thresholding the magnitude spectrum of the non-attacked stego images. These patterns are very similar to those of the embedded watermark. Therefore, the peak patterns can be used in geometrical re-synchronization, and the robust watermark message can be extracted successfully in both non-overlap and overlap implementations. The non-attacked stego image of overlap implementation gives better peak patterns compared to those of the non-overlap implementation because it has full embedding capacity. To improve the robustness of the implementation modes, the watermark embedding energy can be increased to warrant better peak patterns, but it will degrade the visual qualities. In the non-overlap implementation, compromise must be made between the densities of robust part and fragile part. Increasing fragile watermark positions to enhance its localization in tamper detection will reduce those of the robust watermark, thus degrade its robustness.

To evaluate the robustness of the watermark, the attacks listed in Table 2 were carried out. With the obvious axes in the peak patterns, distortions can be compensated with a re-synchronization step to enable successful watermark detection. This is done using Hough transform to estimate the rotation angle, and Maximum Likelihood (ML) to estimate peak periods. Details of the recovery steps can be found in [4].

There are two items worth description here. Firstly, the estimation outcome of Hough transform may deviate one degree. Therefore, brute force search need to be applied in finding the correct parallel lines for period estimation. Secondly, a predefined period range must be specified in the estimation of period between peaks as mentioned in [2]. Overall, both of the implementation modes are equally robust to the attacks. The robust watermark was detected in both non-overlap and overlap implementations after re-synchronization.

Computational costs for the implementation methods are listed in Table 3 for block size 16×16 pixels. The overlap implementation requires more processing time because it embeds robust watermark into every pixel in each block whereas the non-overlap implementation only need to process about 70% of the pixels in each block. The savings of not tracking robust and fragile watermark positions in an overlap implementation does not offset the overall computational costs.

Table 3. Comparison of hybrid watermarking time (seconds)

	Non-overlap implementation	Overlap implementation
Robust embedding	3.10	3.12
Robust detection	0.49	0.50
Fragile embedding	5.70	6.35
Fragile detection	5.60	6.34

5.2   Analysis of Fragile Part Results

The fragile watermark evaluation for both implementation modes were done using local tampering, copy attack, and collage attack. Local tampering was easily detected and highlighted as shown in Fig. 2.

A copy attack on *Cameraman* stego image and its fragile watermark detection results are given in Fig. 3. In the test, a dark color region is copied and pasted onto another region on the cloth of the same image. A similar operation is performed on a textured region, i.e. the lawn. The results of a collage attack involving *Lena* and *Cameraman* stego images gave similar results. The fragile watermark in both implementations highlighted tampered regions correctly.

Since the overlap implementation employs full capacity embedding, it was able to highlight modifications at each pixel. Conversely, the non-overlap implementation embedded its fragile watermark in about 30% pixels of each block. As a result, it was not as accurate as the overlap implementation.

Fig. 2. (Left) Tampered Lena stego image. (Right) Tampered regions highlighted by fragile watermark detection
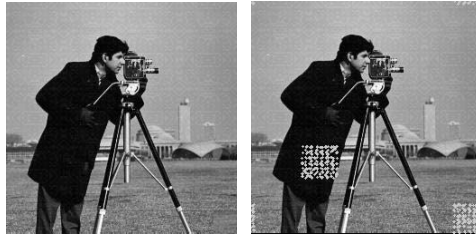
Fig. 3. (Left) Copy-attacked Cameraman stego image. (Right) Tampered regions highlighted by fragile watermark detection

Besides the three types of attacks above, the effects of block size on the fragile watermark are also examined on the non-overlap implementation. As tabulated in Table 4, larger block size requires less processing time. This is due to the convolution operation in hashing neighboring blocks. Also, large block size allows high security with long signatures. On the other hand, the smaller the block size, the more blocks are involved. Thus, the more computing cycles are needed. Nevertheless, smaller block size offers better localization in tamper detection.

Table 4. Effects of block size on fragile watermarking time (seconds)

Block size	Embed time	Detect time
4×4	15.60	15.30
8×8	7.53	7.51
16×16	5.70	5.60
32×32	5.00	4.90

6 Conclusions

We have analyzed and compared the overlap and non-overlap implementations of a hybrid system and have found that both implementations generally produce similar results. This is due to the fact that the robust part in the overlap implementation resisted distortions introduced by the fragile part. Although the overlap implementation reduces computational by not tracking robust and fragile watermark positions, its embedding time and detection time is slightly longer compared to those of the non-overlap implementation. This is caused by the extra processing load of embedding fragile and robust watermarks in all pixel positions. The overlap implementation offers higher watermark capacity for both the robust and fragile watermarks compared to the non-overlap implementation. Hence, the overlap implementation gives better peak patterns than the non-overlap implementation in robust watermark extraction. Due to the same reason, the overlap implementation has better localization in tamper detection compared to the non-overlap implementation. Finally, a balance between tamper detection localization and computational cost must be determined when selecting an optimum block size for both implementation modes. In summary, the overlap implementation can meet high integrity requirements in digital contents while the non-overlap implementation is suitable for commercial applications where processing speed is a preference.

Acknowledgement

References

1. Jiri Fridrich, "A hybrid watermark for tamper detection in digital images", Proceedings of the Fifth International Symposium on Signal Processing and Its Applications (ISSPA '99), 22-25 Aug. 1999. 1 (1999) 301–304
2. F. Deguillaume, S. Voloshynovskiy, T. Pun, "Secure hybrid robust watermarking resistant against tampering and copy attack", Signal Processing, Elsevier. 83 (2003) 2133–2170
3. Chaw-Seng Woo, Jiang Du, Binh Pham, "Multiple Watermark Method for Privacy Control and Tamper Detection in Medical Images", APRS Workshop on Digital Image Computing (WDIC2005), Brisbane, Australia, 21 Feb. 2005 (2005) 43–48
4. F. Deguillaume, S. Voloshynovskiy, T. Pun, "Method for the estimation and recovering of general affine transforms in digital watermarking applications", IS& T/SPIE's 14th Annual Symposium, Electronic Imaging 2002: Security and Watermarking of Multimedia Contents IV, San Jose, CA, USA, 20–25 January 2002. 4675 (2002) 313–322
5. S. Voloshynovskiy, F. Deguillaume, T. Pun, "Multibit digital watermarking robust against local nonlinear geometrical distortions", IEEE International Conference on Image Processing (ICIP2001), Thessaloniki, Greece, October 2001 (2001) 999–1002

Audio Secret Sharing for 1-Bit Audio

Ryouichi Nishimura, Norihiro Fujita, and Yôiti Suzuki

Research Institute of Electrical Communication
Graduate School of Information Sciences, Tohoku University,
2-1-1 Katahira Aoba-ku Sendai 980-8577, Japan
ryou@ais.riec.tohoku.ac.jp
http://ais.riec.tohoku.ac.jp/

Abstract. In this paper, we propose a new secret sharing scheme (SSS)
[1] for audio signals, called as "Binary audio secret sharing (BASS)." SSS
is an encryption method and produces n shared data from an original
data to hide useful information. Applying SSS to audio communications
on the Internet can help to make it more robust against theft and the
tapping of information. Thus, we focused on the 1-bit audio format and
applied SSS to 1-bit audio signals to realize audio secret sharing. More-
over, we propose a method to make each shared data heard as its intended
sound.

1 The BASS Algorithm

1.1 Visual Secret Sharing

The BASS algorithm is based on visual secret sharing (VSS). Of several proposals
regarding VSS, "k out of k VSS [2]" is well known as a method for binary images
and shares an original image into k random dot images. By stacking all of the
shared images, the original information of the image can be easily obtained,
but cannot be with any $k - 1$ shared data. Figure 1 is a sample of 3 out of 3
VSS. The process of encryption of 3 out of 3 VSS is depicted schematically in
Fig. 2, where sub-pixels are produced from each pixel of the original image. In
our proposed method, each pixel of an image corresponds to each sample of a
1-bit audio signal.
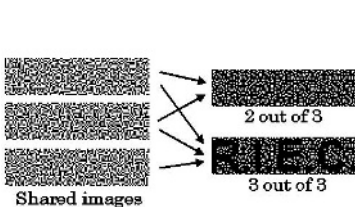
Fig. 1. A sample of 3 out of 3 VSS.
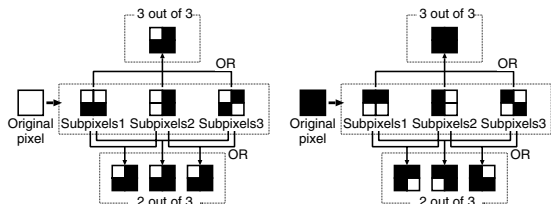
Fig. 2. A sample pattern of sub-pixels.
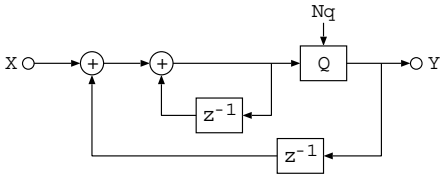
Fig. 3. Block diagram of a first-order $\Delta\Sigma$ converter.
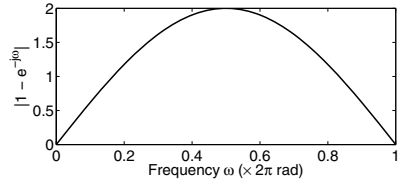
Fig. 4. Frequency response of the first-order $\Delta\Sigma$ converter.

1.2   1-Bit Audio

1-bit audio is a high quality digital audio format employed in the super audio CD (SACD). It has a sampling frequency of 2,822.4 kHz and a resolution of 1-bit. Quantization noise caused by 1-bit resolution can be reduced by the use of the $\Delta\Sigma$ converter. A first-order $\Delta\Sigma$ converter may be expressed by the block diagram depicted in Fig. 3, where Q denotes a quantizer of 1-bit resolution and N_q is the quantization noise. The input X is a discrete time signal sampled with a very high sampling rate. The output signal Y is expressed as

$$Y = X + (1 - z^{-1})N_q. \tag{1}$$

The second term on the right-hand side of this equation consists of the quantization noise N_q, which has a uniformly distributed frequency characteristic, multiplied by a filter having the frequency response depicted in Fig. 4. Consequently, 1-bit audio can realize its large dynamic range despite its low quantizing resolution because the quantization noise is concentrated on the region of high frequency.

1.3   Algorithm of Encryption and Decryption

BASS shares the original signal according to sharing tables $t_b(b = 0, 1)$ for each sample, which have a value of '0' or '1' because of the 1-bit audio. These sharing tables t_b are $k \times n$ boolean matrices, where k is the sharing number and $n = 2^{k-1}$. Defining $A(m, b)$ as the number of '1' in the logical sum of rows of $m(1 \le m \le k) \times n$ sub-matrices of t_b, t_b should be constructed so that $A(m, b)$ satisfies

$$A(m, b) = \begin{cases} \sum_{i=1}^{m} \dfrac{n}{2^i} & (m < k) \\ n - 1 & (m = k, b = 0) \\ n & (m = k, b = 1). \end{cases} \tag{2}$$

Defining T_b as collections of matrices obtained by exchanging any rows of t_b with other rows, shared signals $s_i(1 \le i \le k)$ are obtained as

$$\begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_k \end{pmatrix} = \begin{pmatrix} t_{o_1} & t_{o_2} & \dots & t_{o_L} \end{pmatrix} \quad t_{o_j} \in \{T_0, T_1\} \tag{3}$$

from the original signal o represented by

$$o = \begin{pmatrix} o_1 & o_2 & \dots & o_j & \dots & o_L \end{pmatrix} \quad o_j \in \{0, 1\}. \tag{4}$$

When the sharing number k is $3(k=3)$, for example, t_0 and t_1 may be 3×4 matrices like as

$$\begin{pmatrix} 1\,0\,0\,1 \\ 0\,1\,0\,1 \\ 1\,1\,0\,0 \end{pmatrix} \in T_0 \quad \begin{pmatrix} 1\,0\,0\,1 \\ 0\,1\,0\,1 \\ 0\,0\,1\,1 \end{pmatrix} \in T_1. \tag{5}$$

Decryption of the original signal in BASS requires all the k shared signals. The $m \times (n \cdot L)$ matrix represented in Eq. (3) is separated at every n columns, resulting in L sets of $m \times n$ matrices. When $A(m, t)$ is calculated for each $m \times n$ matrices,

$$A(m, t, j) = \begin{cases} \sum_{i=1}^{m} \dfrac{n}{2^i} & (m < k) \\ n - 1 + o_j & (m = k) \end{cases} \tag{6}$$

is obtained for each $j(1 \leq j \leq L)$th $m \times n$ matrix. Consequently, we can decrypt the original signal $o_j(1 \leq j \leq L)$ from all the k shared signals using Eq. (6), since the information of the original signal o_j can be obtained only if $m = k$.

2   Shared Signals Heard as Intended Decoy Sound Signals

BASS based on the original k out of k VSS makes all shared signals heard as random noise. However, we can encrypt an original signal into $k - 1$ shared signals heard as distinct and intended decoy sounds and one random noise-like signal by properly constructing sharing tables. These $k - 1$ shared signals can be any audio signals. Note that making shared signals heard as intended sounds does not lessen any security of the BASS.

Defining $1 \times k$ matrix $p_i(1 \leq i \leq k)$ as

$$\begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_k \end{pmatrix} = t_0 \in T_0 \tag{7}$$

using any sharing table t_0. We also define $n \cdot L$ samples of arbitrarily selected audio signal as d_i. Define $d_i'(1 \leq i \leq k - 1)$ as d_i decimated by $1/n$, resulting d_i' in a signal of L samples. We further define shared signals $s_i(1 \leq i \leq k - 1)$,

which should be heard as the intended sounds, and a random noise-like shared signal s_k as

$$\begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_i \\ \vdots \\ s_k \end{pmatrix} = \begin{pmatrix} p_1^{(d'_{11})} & \cdots & p_1^{(d'_{1j})} & \cdots & p_1^{(d'_{1L})} \\ p_2^{(d'_{21})} & \cdots & p_2^{(d'_{2j})} & \cdots & p_2^{(d'_{2L})} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ p_i^{(d'_{i1})} & \cdots & p_i^{(d'_{ij})} & \cdots & p_i^{(d'_{iL})} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ p_k^{(d'_{k1})} & \cdots & p_k^{(d'_{kj})} & \cdots & p_k^{(d'_{kL})} \end{pmatrix}, \tag{8}$$

where

$$d'_i = \begin{pmatrix} d'_{i1} & d'_{i2} & \cdots & d'_{iL} \end{pmatrix} \tag{9}$$

$$d'_{kj} = o_j + \sum_{i=1}^{k-1} d'_{ij} \tag{10}$$

$$p_i^{(0)} = p_i \tag{11}$$

$$p_i^{(1)} = \overline{p_i}. \tag{12}$$

In Eq. (12), $\overline{p_i}$ denotes the logical negation of p_i.

Finally, we verify that it is possible to decrypt the original signal from all of $s_i (1 \leq i \leq k)$. We define any column of Eq. (8) as

$$t'_j = \begin{pmatrix} p_1^{(d'_{1j})} \\ p_2^{(d'_{2j})} \\ \vdots \\ p_i^{(d'_{kj})} \end{pmatrix}. \tag{13}$$

Here, t'_j is a matrix obtained by exchanging columns of any matrix included in T_0 and taking logical negation $\sum_{i=1}^{k} d'_{ij}$ times. Sharing tables T_0 and T_1 have the property that when we exchange any column of any sharing table included in T_0 with its logical negation, it becomes one included in T_1, and vice versa. This property is expressed as

$$t_b^{(I)} \in T_{((I+b) \bmod 2)} \quad (b = 0, 1). \tag{14}$$

Thus,

$$t'_j \in T_{(\sum_{i=1}^{k} d'_{ij}) \bmod 2} = T_{(o_j + 2 \cdot \sum_{i=1}^{k-1} d'_{ij}) \bmod 2}$$
$$= T_{o_j}. \tag{15}$$

is obtained.
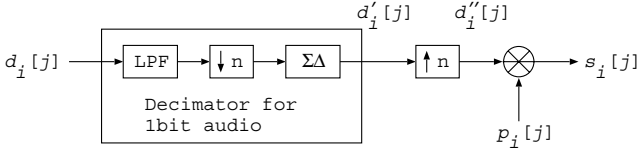
Fig. 5. Block diagram of the relationship between d_i and s_i.

3   Verification of Decoy Sound

Figure 5 shows a block diagram to illustrate the method described in the previous section enabling $k - 1$ shared signals to be heard as decoy sounds. In this figure, the input signal $d_i[j]$ and the output signal $s_i[j]$ are an intended decoy signal and an obtained shared signal, respectively. Here, we show that these two signals have close frequency characteristics. Notations of the variables used in this verification are summarized in Table 1.

Table 1. Notations of variables.

Variable	Description
o	1-bit audio signal of the original signal $\{0, 1\}$
s_i	1-bit audio signal of a shared signal $\{0, 1\}$
$s_i[j]$	Scalar representation of s_i ($\{0\}$ is replaced with $\{-1\}$)
$S_i[k]$	Frequency spectrum of $s_i[j]$
$S_i'[k]$	Frequency spectrum of $s_i[j]$ decimated by $1/n$
p_i	Decomposition pattern $\{0, 1\}$
$p_i[j]$	Scalar representation of p_i ($\{0\}$ is replaced with $\{-1\}$)
$P_i[k]$	Frequency spectrum of $p_i[j]$ after zero-padding
d_i	1-bit audio signal of a decoy signal $\{0, 1\}$
$d_i[j]$	Scalar representation of d_i ($\{0\}$ is replaced with $\{-1\}$)
$D_i[k]$	Frequency spectrum of $d_i[j]$
d_i'	1-bit audio signal of d_i decimated by $1/n$ $\{0, 1\}$
$d_i'[j]$	Scalar representation of d_i' ($\{0\}$ is replaced with $\{-1\}$)
$D_i'[k]$	Frequency spectrum of $d_i'[j]$
$d_i''[j]$	$d_i'[j]$ up-sampled by n
$D_i''[k]$	Frequency spectrum of $d_i''[j]$

3.1   Analysis of the Time Domain

In Fig. 5, $d_i[j](j = 1, 2, \ldots, nL)$ is a signal obtained from a decoy signal $d_i \in \{0, 1\}$ by replacing $\{0\}$ with $\{-1\}$. The output signal $s_i[j](j = 1, 2, \ldots, nL)$ is then represented using d_i' and $p_i[j]$ as follows:

$$s_i[j] = (d_i'[1] \cdot p_i' \quad d_i'[2] \cdot p_i' \quad \ldots \quad d_i'[L] \cdot p_i') \tag{16}$$
$$p_i' = (p_i[1] \quad p_i[2] \quad \ldots \quad p_i[n]).$$
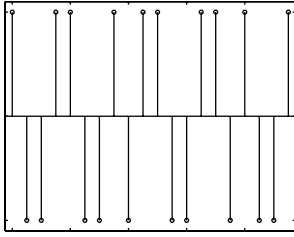
Fig. 6. An example of $s_i[j]$ actually obtained with the proposed method.
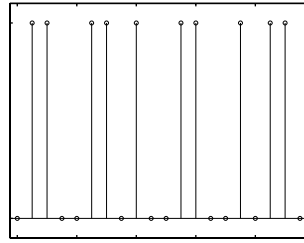
Fig. 7. Theoretically derived signal corresponding to Fig. 6.

For all components where $d_i'[j] = -1$, $\{1\}$ of $\boldsymbol{p}_i'$ is replaced with $\{-1\}$, and vice versa. Accordingly, $s_i[j]$ obtained by Eq.(16) is equivalent to $\boldsymbol{s}_i$ represented by Eq.(8). Note that the sign of the AC component of $s_i[j]$ is reciprocal of that of $\boldsymbol{s}_i$ because the sign of each component is inverted when $d_i[j]$ is -1, namely, d_{ij}' is 0.

Figure 6 shows an example of the waveform of $s_i[j]$ when $k = 3$, $L = 5$, $d'[j] = (1\ 1\ -1\ -1\ 1)$, and $p_i = (1\ 0\ 0\ 1)$. Accordingly, a $1 \times nL$ matrix $\boldsymbol{s}_i$ obtained with the parameters mentioned above is represented by

$$\boldsymbol{s}_i = (\bar{\boldsymbol{p}}_i\ \ \bar{\boldsymbol{p}}_i\ \ \boldsymbol{p}_i\ \ \boldsymbol{p}_i\ \ \bar{\boldsymbol{p}}_i)$$
$$= (0\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0). \qquad (17)$$

Equation (17) can be depicted as in Fig. 7. Comparing these two figures, it is clear that the AC components of the signals are the same.

3.2 Analysis of the Frequency Domain

In Fig. 5, a decoy signal $d_i[j]$ with a length of $n \cdot L$ is fed to an LPF, a down-sampler by n and a $\Sigma\Delta$ converter. The role of the LPF is to remove high frequency components with the cut-off frequency of $f_c = f_s/2n$, not to introduce aliasing at the following down-sampler. In this operation, some audible sound components in a decoy signal might also be removed when the value of n is large. It is then down-sampled and requantized into a 1-bit sequence by the $\Delta\Sigma$ converter. Consequently, $d'[j]$ becomes a 1-bit audio signal of decoy sound $d_i[j]$ decimated by $1/n$.

Next, $d'[j]$ is up-sampled by n. When the signal obtained by this operation is defined as $d''[j]$, the frequency spectrum of $d''[j]$ becomes that of $d'[j]$ shrunk by $1/n$ on the frequency axis due to the up-sampling by n, and includes aliasing components.

Define the frequency spectrum of the signal obtained by zero-padding into $p_i[j]$ so as to have a length of $n \cdot L$ as in $P_i[k](k = 0, 1, \ldots, nL - 1)$, that of $d_i[j]$ as $D_i[k](k = 1, 2, \ldots, nL - 1)$, and that of $s_i[j]$ as $S_i[k]$. From Fig. 5, $S_i[k]$ can be derived by Eq. (18) as the product of $D_i''[k]$ and $P_i'[k]$.

$$S_i[k] = D_i''[k] \cdot P_i[k] \qquad (k = 0, 1, \ldots, L - 1) \qquad (18)$$

Since $D_i''[k](k = 0, 1, \ldots, nL - 1)$ can be expressed by $D_i'[k](k = 0, 1, \ldots, L - 1)$ with a cycle period of L, they are exactly the same within $(k = 0, 1, \ldots, L - 1)$. Since the audible sound components are retained after the sampling frequency becomes $f_s/2n$ because of the very high sampling frequency, $s_i'[j]$, which is obtained by decimating $s_i[j]$ by $1/n$, can be heard as the same as $s_i[j]$. Its frequency spectrum is represented as

$$\begin{aligned} S_i'[k] &= D_i''[k] \cdot P_i[k] \\ &= D'[k] \cdot P_i[k] \quad \text{for } (k = 0, 1, \ldots, L - 1). \end{aligned} \tag{19}$$

Since $S_i'[k]$ and $D_i'[k]$ in Eq. (19) are the frequency spectra of $s_i[j]$ and $d_i[j]$ decimated by $1/n$, respectively, Eq. (19) can be represented using a LPF, $H_{LP}[k](k = 0, 1, \ldots, nL - 1)$, with a cut-off frequency of $f_s/2n$ as follows:

$$S_i[k] \cdot H_{LP}[k] = D_i[k] \cdot P_i[k] \cdot H_{LP}[k] \quad (k = 0, 1, \ldots, nL - 1). \tag{20}$$

Equation (20) states that a shared signal is heard as the decoy signal filtered by $P_i[k]$. Moreover, $p_i[k] = (1 \ \ -1 \ \ -1 \ \ 1)$, which is employed in the previous section as an example, should provide sound without distortion because of its linear phase characteristics. However, a signal actually obtained by this method includes not only the decoy sound but also an additive noise. The reason for this may be the quantization noise brought about by the $\Delta\Sigma$ converter in Fig. 5 to requantize the multi-bit signal into a 1-bit audio signal. Namely, $D_i'[j]$ is $D_i[j]$ decimated by $1/n$ and added to a quantization noise. This effect is not considered in the previous discussion, and thus, further investigation is needed to verify it.

4   Conclusions

We propose a binary audio secret sharing method, BASS, which can share an original audio signal into k shared audio signals. A method to compose the $k - 1$ shared signals heard as different and intended decoy sounds is also proposed. We have verified these sounds as the intended decoy sounds.

Acknowledgments

References

1. Adi Shamir: How to share a secret. Communications of the ACM **22** (1979) 612–613
2. Moni Naor: Visual Cryptography. Advances in Cryptology - EUROCRYPT '94 LNCS 950 (1995) 1–12

Watermarking with Association Rules Alignment

Jau-Ji Shen and Po-Wei Hsu

Institute of Information Management, National Formosa University,
Huwei, Yunlin County, Taiwan, R.O.C.
`jj.shen@msa.hinet.net`

Abstract. A new watermarking technique called associative watermarking is proposed in this paper. The watermark in our new scheme will be embedded by applying the concept of association rules borrowed from the field of data mining. We implant the watermark's rules into the original image by changing a very small portion of the image's pixels during the rules alignment process. As a result, the robustness can be enhanced by emphasizing the specific rules mined from critical areas in the original image. Instead of transforming the watermark into its bit pattern, we just mine the association rules from the watermark and embed them into a colorful image. Our experimental results reveal that the PSNR value of the image after watermark implanting exceeds 40, and more than 90% of the watermarked images can be accurately recognized after similarity checking.

Keywords: Association rule, Digital Watermarking, Image, Similarity

1 Introduction

Due to the shockingly high speed at which the popularity of the Internet has been growing, computers and the networks they form have already become a bigger and bigger essentially important part of modern people's everyday lives. Upon the basis of the Internet, our planet has been built into a global village where any user of the Internet can easily make her/his own original work public or offer various kinds of services through the Net, so that resources of all sorts can be conveniently spread out, shared and made best use of. Of course, to look at it another way, through the links of computer networks around the world, digital data of all kinds, including other people's creations, can be readily accessed and passed around. However, these enormous numbers of original works accessed on the Internet might get destroyed, changed, altered, modified, or copied and distributed without proper authorization by the copyright owner, and these intellectual-property-right-law-breaking actions can do serious damage to the value and integrity of the works and thus harm the copyright owner. To keep such illegal behaviors from taking effect, quite a number of researchers have been devoted to the development of a wide variety of protection mechanisms, and many watermarking techniques have been offered [2, 5, 6, 7, 8, 9, 11, 16, 18].

In this paper, we shall propose a new digital watermarking technique that combines the existing watermark bit pattern embedding extracting technique with the concept of association rules mining developed in the field of data mining. Distinct from traditional watermarking techniques, our new scheme functions by extracting useful association rules from the watermark image and embedding these association rules as the hidden

data into the cover image instead of the whole watermark bit pattern. One very important advantage our new method gains is that the strategy to hide the association rules other than the bit pattern frees the embedding process of most of its original limits on the image format, offering great flexibility, allowing the association rules to be about the brightness, contrast, color, item association sets, or the text. Besides the association rule mining technique, the discrete cosine transform (DCT) technique is also employed in our new scheme, so that the mid-low frequency coefficients can be taken as proof-reading items for the association rules, because the mid-low frequency coefficients are relatively more robust against image processing. Thus, we can not only effectively reduce the damage done to the image quality but also increase the robustness of the hidden data.

The rest of this paper is organized as follows. In Section 2 below, we shall briefly introduce the two techniques, namely association rule mining and discrete cosine transform, upon which our new scheme is based. Then, in Section 3, we shall present our new scheme to every detail, followed by the experimental results in Section 4 that will demonstrate the practicability of our new method. Finally, the conclusion will be in Section 5.

2   Background

■   Association Rules

Data mining [3, 4, 15, 17, 19] is a way to dig out useful information from huge data-bases, and the mining of association rules is one of the techniques developed in the field of data mining. At the very beginning, association rules are defined upon a transaction database. Suppose there is a set called *set I* that represents the whole product database. In this *set I*, each product is now called an *item*. Any set that is composed of one or more of the items from *set I* is called a *transaction* and is given a one-and-only identity code called *TID*. Putting all these transactions together, we obtain a database of transactions called *set D*. Suppose we have items X and Y that satisfy the following terms $X \subseteq I$, $Y \subseteq I$, $X \cap Y = \Phi$, and $|X| + |Y| = K$, and then we claim that $X \rightarrow Y$ is the association rule defined upon K-itemset. In the transaction database *D*, if c% of the transactions where item X is involved also involve item Y, then we say that the *confidence* of the association rule $X \rightarrow Y$ is c. On the other hand, in the transaction database *D*, if both item X and item Y show up in s% of the transactions, then we say that the *support* of the association rule $X \rightarrow Y$ is s. Therefore, the so-called association rules are those rules in *D* that satisfy a predefined combination of minimum confidence and minimum support.

■   Discrete Cosine Transformation (DCT)

The discrete cosine transform technique [1, 12] is a way to transform data from the spatial domain to the frequency domain or from the frequency domain to the spatial domain, where forward discrete cosine transformation (FDCT) is used to transform data from its spatial domain to its frequency domain, and inverse discrete cosine transformation (IDCT) is used to do exactly the opposite [13]. DCT is widely used in

various kinds of image processing applications because it is an effective way to describe data in the spatial domain by means of cosine waves, and it can be further divided into zones of low frequency, mid frequency, and high frequency, where the absolute values of the coefficients stand for the magnitude of the energy. Coefficients from different zones have different degrees of influence on the spatial data of the image, and the degree of influence goes down as the frequency value increases. As a result, in general, DCT-related watermarking techniques tend to hide the watermark by manipulating the mid/low frequency coefficients so that the image quality will not get degraded too much while the robustness of the hidden data can be ensured.

3   Association Rules Alignment

The following are the definitions of some notations to be used in our scheme:

I: the original image.

D: the test image.

W: the watermark. (which is also an image)

I_G, I_B: the green pixel value and the blue pixel value of the original image.

D_G, D_B: the green pixel value and the blue pixel value of the test image.

W_G, W_B: the green pixel value and the blue pixel value of the watermark image.

$B1_I(b)$, $B1_D(b)$, $B1_W(b)$: the 1st block value of any block b in I, D or W with respect to green pixel plane.

R_I, R_D, R_W: the set of association rules mined from the original, test and watermark images respectively.

I_B': the blue pixel value of the original image after the alignment.

R_A: the set of association rules for the original image after the alignment.

R_M: the set of association rules mined from R_W, where $R_M=\{(A\ B) \mid$ For all distinct A, Rule (A→B) has maximum confidence in rule (A→Y$_1$), (A→Y$_2$),…, where B∈ {Y$_1$, Y$_2$,…}}.

$R_{I,Count(i)}$: the times of occurrence of each association rule i (i is an element of R_M) in R_I.

$R_{D,Count(i)}$: the times of occurrence of each association rule i (i is an element of R_M) in R_D.

3.1   Association Rules of Image

To begin with, we define 2-itemsets on both the original image and the watermark image, and then we do alignment to the association rules formed among these 2-itemsets so that the association rules derived from the watermark image can be hidden among those derived from the original image. In order to define 2-itemsets, both the original image and the watermark image have to be divided into blocks with 8x8 pixels each, and the 1st item and 2nd item have to be defined for each block. The following Eq.(1) is set to generate the 1st block's value which will use to produce the 1st item value of association rule. For any block b,

$$B1_I(b) = \left| I_G(b) - \frac{1}{n} \sum_{\substack{\forall b'\ adjacent \\ to\ block\ b}} I_G(b') \right| \tag{1}$$

Where n=3 if block b locates at 4 corners, n=5 if block b locates at 4 boundary edges, n=8 if otherwise. It is a similar calculation of $B1_D(b)$ and $B1_W(b)$ by replacing "I_G" in Eq.(1) with "D_G" and "W_G" respectively.

Subsequently, the following Eq.(2) is set to generate the 2nd block's value which will use to produce the 2nd item value of association rule. For any block b,

$$B2_I(b) = \frac{1}{L} \sum_{\substack{L \text{ coefficients in low} \\ \text{frequency locations}}} \left| DCT \ (I_B(b)) \right| \tag{2}$$

Where $DCT(I_B(b))$ is the DCT transformation of block b. There are L DCT's values in low frequency area summation. In order to explain our new idea easily, L is set to 2 in this paper. It is a similar calculation of $B2_D$ and $B2_W$ by replacing "I_B" in Eq.(2) with "D_B" and "W_B" respectively.

Let $B1(b)$ and $B2(b)$ represent any block b's 1st and 2nd block value relevant to I, D or W, then the association rule's 1st and 2nd items of block b are defined as the following Q_1 and Q_2 respectively.

$$Q_1(B1(b)) = \left\lfloor \frac{|B1(b)|}{256 / \Omega} \right\rfloor \tag{3}$$

$$Q_2(B2(b)) = \begin{cases} 0 & |B2(b)| < \delta \\ \left\lfloor \frac{|B2(b)|}{\delta} \right\rfloor & \delta \le |B2(b)| < \delta(\Omega - 1) \\ \Omega - 1 & \delta(\Omega - 1) \le |B2(b)| \end{cases} \tag{4}$$

By using Eq.(3) and Eq.(4), the values of the 1st item and 2nd item of each block are restricted within the range from 0 to Ω-1. In addition, both Ω and δ have to be integers within the range from 1 to 256.

Since, the alignment process will tune the 2nd item value. Due to the less sensitivity of blue pixel change, the picture quality can keep better when the value change made by alignment process. Any block b in I, D or G will be assigned one association rule with two items. In the later explanation of our rule alignment process, the 1st item will be used to locate any I's block which has the same 1st item value to a chosen W's block. Then, we tune the 2nd item value of the located I's block to agree the 2nd item value of that chosen W's block. From the YUV to RGB transformation formula [9], the green color is more sensitive in pixel value changes whereas the blue color is not. We choose the green pixel to derive the 1st item of block's association rule because its sensitivity of change can keep the location alignment between I and W from destroying after image processing on I.

3.2 Watermark Embedding

There are two phases in the watermark embedding procedure. (see Figure 1)

In Phase I, both the original image and the watermark image are divided into image blocks, and two representative values are derived for each block, namely $B1_W$(or $B1_I$) and $DCT(B2_W)$(or $DCT(B2_I)$). Then, in Phase , we pick out the association rules (R_A) with the same 1st item value from R_I and R_M. This way, after the rule alignment, the $Q_2(DCT(B2_I))$ of each rule in R_A can be adjusted to achieve the effect of watermark rule hiding.
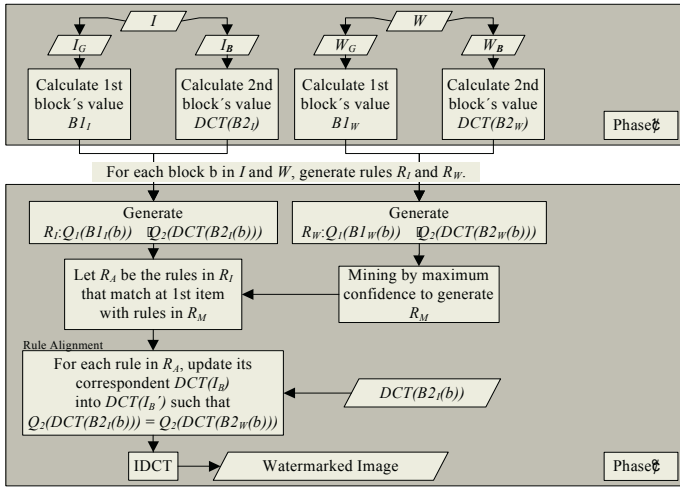
Fig. 1. Watermark embedding scheme.

In the association rules we define, the 1st item contains information about the degree of change of the image pixel value. However, for the majority of images, smooth areas take up most of the space, and modifications done to such smooth areas are unfortunately more easily detectable. Therefore, in the rule alignment process, we can skip the rules in R_M whose 1st item values are small. This way, not only the over-modification problem can be avoided, but the association rules can be hidden into more critical edges on the images.

3.3 Watermark Detection

Like the watermark embedding procedure, the watermark detection procedure in our new scheme is also composed of two phases. (see Figure 2)

In Phase , the original image, the test image, and the watermark image are all divided into blocks, and two representative values $B1_I$ and $DCT(B2_I)$ are derived for each block from the original image, two representative values $B1_D$ and $DCT(B2_D)$ are derived for each block from the test image, and two representative values $B1_W$ and $DCT(B2_W)$ are derived for each block from the watermark image. In Phase II , the times of occurrence of each association rule i from R_M in R_I and R_D are respectively counted and taken down as $R_{I,Count(i)}$ and $R_{D,Count(i)}$. With these figures, we define similarity S as follows:

$$S = \frac{\sum_{i=1}^{N} C_i R_{D,Count(i)}}{\sum_{i=1}^{N} C_i R_{I,Count(i)}} \tag{5}$$

In Eq.(5), N stands for the number of association rules there are in R_M, and C_i represents the 1st item value of association rule i in R_M. We set a threshold value T so that when $S \geq T$, it is judged that there is a watermark (or more) hidden in image D, and no watermark is said to be hidden in the image when $S < T$.
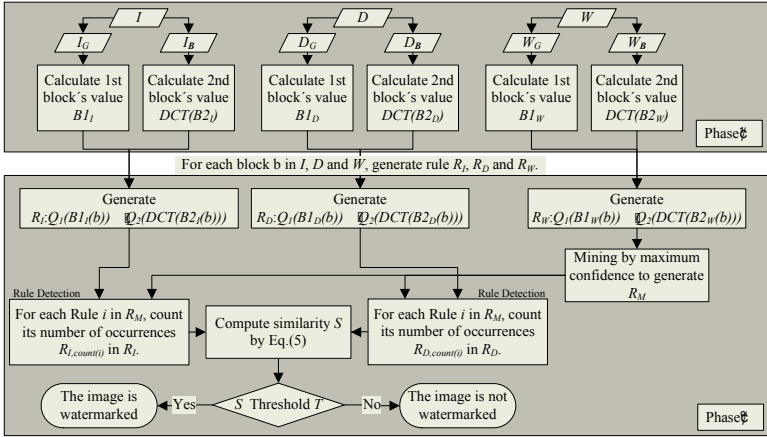
Fig. 2. Watermark detecting scheme.

4   Experimental Results

In this section, we shall present the results of some experiments we have conducted to demonstrate the effectiveness and practicability of our new scheme. The watermark image we used in our first experiment is a 128×128-pixel colored copy of the Emblem of National Formosa University, Taiwan (see Figure 3(a)). On the other hand, the original image we used in this experiment is a 256×256-pixel color image called "Lena," which is shown in Figure 3(b). In our first experiment, the parameters Ω and δ were respectively preset as 8 and 30. After embedding the watermark image into the original image, we obtained the resultant image shown in Figure 3(c), whose PSNR(Peak Signal-to-Noise Ratio) value we have derived by the following formula below:

$$MSE = \frac{1}{3MN}\sum_{i=1}^{M}\sum_{j=1}^{N}((f_R(i,j)-\hat{f}_R(i,j))^2 + (f_G(i,j)-\hat{f}_G(i,j))^2 + (f_B(i,j)-\hat{f}_B(i,j))^2)$$

$$PSNR = 10\log_{10}\frac{255^2}{MSE}$$

(6)

In Eq.(6), the variable N stands for the length of the original image, and M represents the width of it. In addition, f is the value of the original image pixel at position (i, j), and $\hat{f}$, on the other hand, is the value of the image pixel at position (i, j) with the watermark image already embedded in. The subscripts R, G, and B indicate red, green, and blue, respectively. The resultant PSNR value is 42.17dB.

As for the robustness tests against image processing, we applied PhotoShop5.0 and modified the embedded image by doing "More blurring," "More sharpening," "Histogram," "Brightness adjustment (+20,+30,+40)," "Gaussion noise (σ=10,15,20)," and "JPEG compression (50%,40%,30%)." After the above image processing, the resultant embedded images are those shown in Figure 4(a~f). Comparing these modified images with the original embedded image, we obtained the similarity value S, and we empirically set the threshold value for it to be 1.2. As a result, as Table 1 reveals, our

system is indeed capable of identifying that these Lena images have the watermark embedded in them even though they have gone under image processing and been modified.

Fig. 3. (a) Color watermark. (b) Color Lena image. (c) Watermarked color Lena image.

Fig. 4. (a) More blurring. (b) More sharpening. (c) Histogram. (d) Brightness adjustment (+40). (e) Gaussion noise(σ =20). (f) JPEG compression(30%).

Table 1. Check results for Lena image.

Image processing type	Similarity value	Using *T*=1.2	PSNR (dB)
No-embedded	1	No	-
Attack-Free	1.6713	Yes	42.17
More blurring	1.5888	Yes	31.22
More sharpening	1.3294	Yes	24.21
Histogram	2.1786	Yes	19.35
Brightness adjustment (+20, +30, +40)	1.6713 , 16713 , 16713	Yes , Yes , Yes	22.12 , 18.71 , 16.32
Gaussion noise (σ =10, 15, 20)	1.6596 , 1.6244 , 1.4240	Yes , Yes , Yes	28.06 , 24.75 , 22.3
JPEG compress (50%, 40%, 30%)	1.636 , 1.3535 , 1.2005	Yes , Yes , Yes	34.61 , 34.52 , 33.27

In our other experiment, we used our new scheme to embed the same watermark image (Figure 3(a)) in as many as 1,000 color images (256x384 pixels each) of people, scenery, architecture, flowers, animals, etc. After the same image processing procedures we did in the previous experiment, we tested the watermark detection ability of our new scheme, and the results are those shown in Table 2. The accuracy rate is higher than 90%.

Table 2. Check errors for 1000 color images (using T=1.2).

Image processing type	Check errors
No-embedded	0/1000 (0%)
Attack-Free	6/1000 (0.6%)
More burring	8/1000 (0.8%)
More sharpening	26/1000 (2.6%)
Histogram	0/1000 (0%)
Brightness adjustment (+20, +30, +40)	6,6,6/1000 (0.6%, 0.6%, 0.6%)
Gaussion noise (σ =10, 15, 20)	8,8,18/1000 (0.8%, 0.8%, 1.8%)
JPEG compress (50%, 40%, 30%)	8,23,96/1000 (0.8%, 2.3%, 9.6%)

5   Conclusion

In this paper, a new digital watermarking technique based on the concept of association rules is proposed. By means of data mining, we can obtain association rules for both the original image and the watermark image, and some specific association rules of the watermark image can be taken as the secret information and hidden among the association rules of the original image. By doing so, we can break through many of the limitations that trouble traditional watermarking techniques. In addition, by exploring different types of attributes in the watermark, we can come by a wider variety of sets of association rules, which gives greater flexibility to the choice of association rules to use when hiding the watermark. Besides, association rule mining done to the original image can help pick out more suitable areas to embed secret data in and can thereby provide the hidden watermark with better robustness without degrading the cover image quality. Finally, according to our experimental results, 90% of the hidden watermarks can be correctly detected, and our PSNR value can go beyond 40dB. As the results demonstrate, our new method is indeed effective and practical.

References

1. Barni M., Bartolini F., Cappellini V. and Piva A.: A DCT-domain system for robust image watermarking: Signal Processing, Vol. 66 (1998) 257-372.
2. Chen Liang-Hua and Lin Jyh-Jiun: Mean quantization based image watermarking: Image and Vision Computing, Vol. 21 (2003) 717-727.
3. Chen G, Wei Q., Liu D. and Wets G.: Simple association rules (SAR) and the SAR-based rule discovery: Computer and Industrial Engineering, Vol. 43, No. 4 (2002) 721-733.
4. Feng L., Dillon T. and Liu J.: Inter-transactional association rules for multi-dimensional contexts for prediction and their application to studying meteorological data: Data and Knowledge Engineering, Vol. 37, No. 1 (2001) 85-115.

5. Fan Y. C. and Tsao H. W.: Watermarking for intellectual property protection: Electronics Letters, Vol. 39, No. 18 (2003) 1316-1318.
6. Hsu C. T., Wu and J. L.: Hidden digital watermarks in images: IEEE Transactions on Image Processing, Vol. 8 (1999) 58-68.
7. Hsu C. T. and Wu J. L.: Multiresolution watermarking for digital images: IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, Vol. 45, No. 8 (1998) 1097-1101.
8. Ingemar J. Cox, Senior Member, IEEE, Joe Kilian, F. Thomson Leighton and Talal Shamoon, Member, IEEE: Secure Spread Spectrum Watermarking for Multimedia: IEEE Transactions on Image Processing, Vol. 6 (1997) No. 12.
9. Kim H. Y. and Afif A.: Secure authentication watermarking for binary images: Computer Graphics and Image Processing, (2003) 199-206.
10. Kutter M., Jordan F. and Bossen F.: Digital Watermarking of Color Images Using Amplitude Modulation: Journal of Electronic Imaging, Vol. 7, No. 2 (1998) 326-332.
11. Kim Young-Won and Oh Il-Seok: Watermarking text document images using edge direction histograms: Pattern Recognition Letters, Vol. 25 (2004) 1243-1251.
12. Kii H., Onishi J. and Ozawa S.: The digital watermarking method by using both patchwork and DCT: IEEE Multimedia Computing and System, (1999) 895-899.
13. Mohamed A. Suhail, Member, IEEE and Mohammad S. Obaidat, Senior Member, IEEE: Digital Watermarking-Based DCT and JPEG Model: IEEE Transactions on Instrumentation and Measurement, Vol. 52, No. 5 (2003)
14. Netravali A. N. and Haskell B. G.: Digital Pictures: Plenum Press, (1995)
15. Srikant R. and Agrawal R.: Mining generalized association rules: Future Generation Computer Systems, Vol. 13, No. 2-3 (1997) 161-180.
16. Shieh C. S., Huang H. C., Wang F. H. and Pan J. S.: An embedding algorithm for multiple watermarks: Journal of Information Science and Engineering, Vol. 19 (2003) 381-395.
17. Savasere A., Omiecinski E. and Navathe S.: An efficient algorithm for mining association rules in large databases: In Proceedings of the VLDB, Zurich, Switzerland, (1995) 432-444.
18. Wang M. S., Chang C. C. and Hwang K. F.: A watermarking technique based on one-way hash functions: IEEE Transactions on Consumer Electronics, Vol. 45, No. 2 (1999) 286-294.
19. Zaki M. J.: Scalable algorithms for association mining: IEEE Transactions on Knowledge and Data Engineering, Vol. 12, No. 3 (2000) 372-390.

A Reversible Watermark Scheme Combined with Hash Function and Lossless Compression

YongJie Wang[1], Yao Zhao[1], Jeng-Shyang Pan[2], and ShaoWei Weng[1]

[1] Institute of Information Science, Beijing Jiaotong University, P.R. China
[2] National Kaohsiung University of Applied Sciences, Taiwan

Abstract. The paper presents a reversible watermark scheme based on hash function and lossless compression technology. The paper extracts the abstract information of the image as hash code and lossless compresses the LSB(Least Significant Bit) of the image to make some space where the image hash can be completely inserted. It needs to be done without introducing a large distortion to the image. The scheme can achieve the authentication purpose by hash code and restores the original image by retrieving the watermark sequence completely because of the lossless decompression of the LSB of the image.

1 Introduction

In the Internet age, audio, video media, image can be easily modified, copied and transmitted. In some application field, such as the medical image, military image and X-ray image, they are not allowed to be illegal modified and manipulated, even if several bits changed.

The reversible watermarking techniques satisfy those requirements. The reversible watermark is also named as lossless watermark, invertible watermark and erasable watermark, that is to say the reversible watermark can completely restore the original image. Owing to the reversibility, the reversible watermark is applied in the medicine fields, physical fields and so on. The retrieved watermark bits compare with the original watermark bits to detect whether the image is authentic.

In [1], Honsinger proposed a reversible watermarking as an application of image authentication in the patent. He realized the reversibility of the watermarking processing by modulo addition and a robust spatial additive watermark. It seems to be the first report on the field of reversible watermark. Another reversible watermark scheme is presented in [2] by Fridrich. The watermark is embedded in the saved space by lossless compression on the bit-plane.

This reversible watermark scheme is based on hash function and lossless compression. So we will review some relevant knowledge about the hash function and lossless in the next section, then propose our reversible image watermarking scheme and the simulation results in the Sect. 3 and Sect. 4, respectively.

2   Relevant Knowledge

This section will introduce two skills applied in the reversible watermark. One is hash function, and another is the lossless compression skill.

2.1   Hash Function

In information security fields, information authentication and digital signature technology develops rapidly. The hash function plays an important role in those applications. In the modern times, the message authentication and digital signatures technology had a great development in the information security. In [3], the hash function is introduced like that. As with the message authentication code, a hash function accepts a variable-size message M as input and produces a fixed-size hash code $H(M)$, sometimes called a message digest, as output. A hash value can be calculated by the formula as follows.

$$h = H(M) \tag{1}$$

The hash code is a function of all the bits of the message and provides an error-detection capability: A change to any bits in the message results in a change to the hash code. So the message receiver can recalculate the hash code to detect the integrality of the information. Usually, the hash function needs not the secret key, but we will use the secret key in this watermark scheme for the security of watermark.

To be useful for message authentication, a hash function H must have the following properties (adapted from a list in [NECH92]):

1) H can be applied to a block of data of any size.
2) H produces a fixed-length output.
3) $H(x)$ is relatively easy to compute for any given x, making both hardware and software implementations practical.
4) For any given code h, it is computationally infeasible to find x such that $H(x) = h$. This is sometimes referred to in the literature as the *one-way* property.
5) For any given block x, it is computationally infeasible to find $y \neq x$ with $H(y) = H(x)$. This is sometimes referred to as *weak collision resistance*.
6) It is computationally infeasible to find any pair (y, x) such that $H(y) = H(x)$. This is sometimes referred to as *strong collision resistance*.

2.2   Lossless Compression

Lossless Compression is also called as reversible compression, distortion-free coding, or noise-free coding and so on. This is a kind of developed technique. We are very familiar with Winzip software which is a good example of lossless compression. Natural images have a high degree of redundancy due to the correlation of pixels with their surrounding pixels. Information theory gives a basis for reducing redundancy, thereby achieving data compression. Accordingly, image data

compression consists of two functions: decorrelation and encoding. Sometimes we also call compression progress as removing redundancy progress. The techniques in which the original image can be completely reproduced are said to be information-preserving or reversible.

3   Proposed Reversible Watermarking Approach

3.1   Embedding and Extracting Process

First map the image to get hash code. Hash function has several selections such as SHA_1 and MD5. Hash code consists of digest information of the original image. In the following section, the paper will authenticate whether the image is modified or not. After getting the hash code, our problem is how to embed the hash code in the original image. Two problems are needed to notice:

1) The method cannot impact the visual quality of watermarked image. PSNR of watermarked image gets to a certain value.
2) The method can restore the original image from watermarked image, that is to say the method can remove the watermark bits and restore the original image.

To satisfy the first point, the paper selects the lower bit-planes to make space for hash code. It is very difficult for the LSB of image to resist some attacks. For the second point, the paper compresses the lowest bit-plane to make space for hash code. The hash functions are typically used for digital signatures to authenticate the message being sent so that the recipient can verify that the message is authentic and that it came from the right person.

We show the flow chart of the embedding process and the detecting process in Fig. 1 and Fig. 2 as follows respectively.
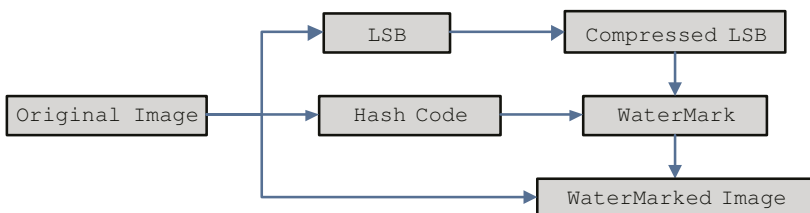
Fig. 1. Embedding Flow Chart.

There are a lot of methods for lossless compress such as arithmetic coding, Huffman coding, Run Length encoding and LZ78 compression algorithm based on the codebooks. By lossless compressing the image (or its feature), we can make some space where the image hash could be inserted. We propose to use the lowest bit-plane that, after lossless compression, provides enough space for the image hash, but bit-planes only include the symbols '0' and '1'. If applied
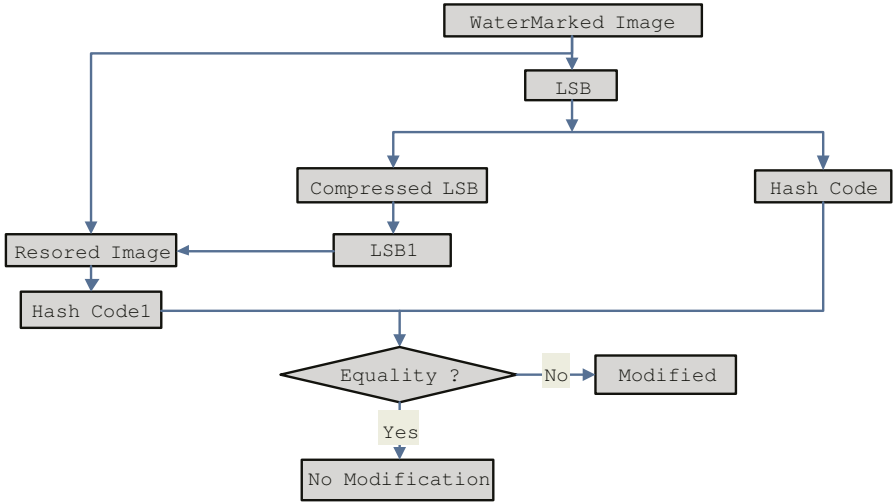
Fig. 2. Detecting Flow Chart.

to compress bit-planes, Huffman coding is of no good effect. The paper applies the Huffman coding and run-length coding to compress bit-planes. Run-length coding (which also can be found as part of the JPEG standard) statistically encodes the symbols '0' and '1' of the bit-planes to get a bit-string, and then Huffman coding encodes the bit-string to compress the bit-planes of the image. In compression process, we make use of the binary property of the bit-planes.

The process is described as follows:

The paper assumes the first symbol of the scan line is '0'. Start the coding by the run-length of the all the symbol '0'. The coder then codes the next new symbol '1' and its run-length repeatedly until the end of the scan is reached. If the first symbol of the scan line is '1', the encoder inserts a codeword for the run-length of zeros. For example, a binary bit-string "0000111000111111" coded as "5 3 3 6", "11110000011111" is coded as "4 5 5". After that we perform the Huffman algorithm on the result of the runlength processing. Then the Huffman code will be the part of the watermark. The encode and decode flow chart are proposed in Fig. 3 and Fig. 4.

4   Experimental Results

The test image is 256-grayscale image "Lena" with size 256×256 shown in Fig. 5. Hash function applies bit-by-bit exclusive-OR(XOR) the algorithm. The length of hash code is 8bit. The probability which the modified image couldn't be detected is 2^{-8}. The improved method is to select the SHA_1 or MD5 algorithm or lengthen the length of hash code. The compression algorithm is adopted as the above Huffman coding and run-length coding. Embed the hash code in the LSB of the image to get good visual quality. Fig. 6 is the watermarked figure.
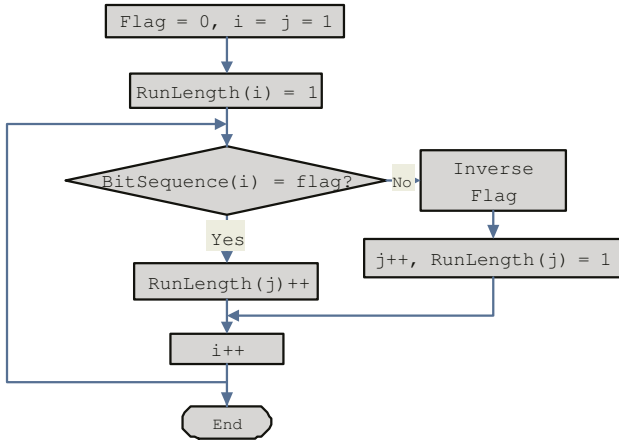
Fig. 3. Encode of Runlenth.
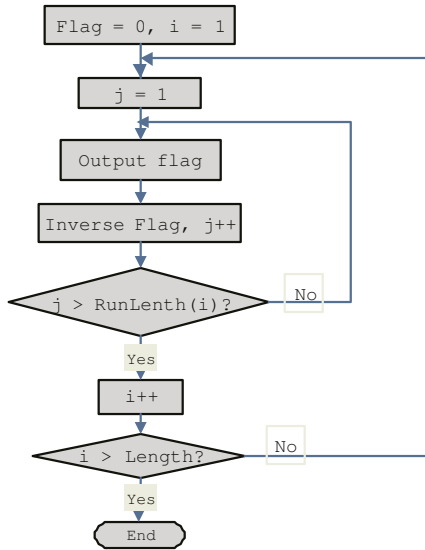
Fig. 4. Decode of Runlenth.

Fig. 5. Original.

Fig. 6. Watermarked.
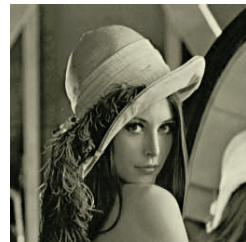
Fig. 7. Restored.

First without any attacks, the paper retrieves the watermark bits and get the Hash code and compressed LSB. Decompress the LSB and get the restored image(See Fig. 7). Hash functions are typically used to authenticate whether the watermarked image is modified or not. If hash code is not changed, the paper compares the original image with the restored image bit by bit, and finds whether they are totally the same which proves the feasibility of the hash code.

Then we perform a series of attack on the embedded image, seeing the figure Fig. 8 to Fig. 12. These distortions are so trivial that we can not find the modification place by our eyes, and the watermark scheme can detect the modification very sensitivity.

Fig. 8. Adds Gaussian white noise with zero mean and 2^{-4} variance.

Fig. 9. Gaussian low-pass filter of size [3 3] with standard deviation 0.5(positive).

Fig. 10. Modified random pixels with random value.

Fig. 11. Rotation by small angel (0.1°).

Fig. 12. Scaling by 1.01 times.

5   Conclusion

In a word, this arithmetic can satisfy two characteristics of reversible watermark.

1) The original image can be recovered;
2) It is sensitive for the modification of image.

At the same time, it does not cause obvious distortions of the image after embedded the watermark. It can obtain the imperceptibility of vision just like other watermarking scheme. There are two advantages of this method:

1) Arithmetic is simple and easy to implement by software or hardware;
2) The used arithmetic is mature, security and robustness;

Thus this arithmetic has the good feasibility in practicality. For example, this reversible watermark technology can be applied to the digital camera of prison. The prison can take photo for the prisoner and keep it in the archives. It does image authentication when the prison need to use these images so that she or he can find whether these images have been modified lawlessly. It is good for the management of the prison. Another application of this reversible watermark scheme is the protection of the CT image in hospital. So the foreground of this reversible watermark theme is optimistic very much. However, this arithmetic has its own limitations. First, it is hard to confirm where the modification happens. Second, it does not refer to the correlation of the whole image, but only uses the low effective bit-plane. We need to improve this watermark theme in order to overcome these limitations. At the same time the self-recoverability of the reversible watermark will be considered. It means the watermarked imaged can recover itself against some little distortions. Then the feasibility of application will be upgraded after these improvements what will be deal with in the next research.

References

1. C. W. Honsinger, P. Jones, M. Rabbani, and J. C. Stoffel :Lossless Recovery of an Original Image Containing Embedded Data. US patent, No 77102/E-D (1999)
2. J.Fridrich,J.Goljan,and R.Du: Invertible Authentication. In Proceedings of SPIE, Security and Watermarking of multimedia Content, San Jose (2001)
3. William Stallings. (ed.): Cryptography and Network Security: Principles and Practice. Prentice-Hall (2002)

Improved Method of Spread Spectrum Watermarking Techniques Using Pixel-Correlations

Zheng Liu

C4 Technology, Inc., Meguro Tokyu Bldg., 2-13-17 Kamiosaki Shinagawa-ku,
Tokyo 141-0021, Japan
zliu@c4t.jp
http://www.c4t.jp/

Abstract. In recent years, the method of spread spectrum watermarking (SSW) has become the one of the most promising technique used for digital watermarking technology. However, the major problems of SSW technology is that the effectiveness of SSW detection will drop fairly without subtracting the host image from watermarked image because of the effect of correlation remains calculated between the host image and original pseudo-random sequences. In our previous work, we have proposed a method to improve the detection of SSW techniques by reduction of correlation remains using pixel-correlations in an image. In this paper, we have investigated the effectiveness of the proposed method with image experiments in detail. The experimental results illustrate that the correlation remains can be reduced greatly by using pixel-correlations.

1 Introduction

With the rapid growth of network distributions of still image, video, and audio, there is an urgent need for copyright protection against piracy. As effective measures, in recent years, many digital watermarking schemes have been proposed for intellectual property right protection of digital media data, in which one of the most promising technique is what called as spread spectrum watermarking (SSW) technology [1], [2], [3], [4], [5], [6], [7], [8], [9].

In our previous work [10], we have proposed a method to improve the effectiveness of SSW techniques by reduction of correlation remains using pixel-correlations information. The principle is that in the general images, there is high correlated information among the adjacent pixels. The fact implies that if we calculate the correlations between a pseudo-random sequence and the block-images with the same size, the correlations calculated in two block-images will be about the same if the center positions of the two block-images are distanced in 1 pixel only. Therefore, we can use this kind of correlation information to estimate the correlation remains calculated in a block-image so as to subtract them form the calculated correlations.

In this paper, we have investigated the effectiveness of proposed method with image experiments in detail. The experimental results illustrate that the effect of correlation remains can be improved greatly by the proposed method and a better effectiveness can be acquired comparing with the traditional method.

2   Review of Proposed Method

At first, let's consider a typical equation of detection for SSW techniques as follows.

$$
\begin{aligned}
P_c &= G_c' \cdot W \\
&= (G_c + \alpha \cdot W) \cdot W \\
&= \frac{1}{N} \sum_{i=0}^{N-1} g_c(i) \cdot w(i) + \frac{1}{N} \sum_{i=0}^{N-1} w(i) \cdot w(i) \\
&= PR_c + PW_c,
\end{aligned}
\tag{1}
$$

where G_c and G_c' represents a block-image before and after the pseudo-random sequence, W, is embedded respectively; PR_c and PW_c represents the correlation remains and the autocorrelation, calculated between G and W, W and W, respectively. As indicated in [10], if the correlation remains is too large, the autocorrelation will disappear in PR_c. Therefore, the correlation remains greatly reduces the effectiveness of SSW technology.

In our previous work [10], we have proposed a method to improve the effectiveness of SSW techniques by reduction of correlation remains using the information of pixel-correlation. The principle of proposed method can be briefly described as follows.

$$
\begin{aligned}
P_s &= G_s' \cdot W \\
&= (G_s + \alpha \cdot W_s) \cdot W \\
&= \frac{1}{N} \sum_{i=0}^{N-1} g_s(i) \cdot w(i) + \frac{1}{N} \sum_{i=0}^{N-1} w_s(i) \cdot w(i) \\
&= PR_s + PW_s,
\end{aligned}
\tag{2}
$$

where G_s is a block-image with the same size of G_c and with a center distanced 1 pixel form that of G_c; G_s' is G_s embedded with the pseudo-random sequence W; PR_s and PW_s represent the correlation calculated by G_s and W and by W_s and W, respectively. Because the W_s is a block one with the same size of W and with a center distanced 1 pixel from that of W, there is $PW_s \approx 0$. Meanwhile because the high pixel-correlations between G_c and G_s, there is $PW_c \approx PW_s$ statistically. Therefore, we can compute the correlations in the center block-image G_c with reduction of correlation remains using pixel-correlations as follows.

$$
P = P_c - (P_l + P_r + P_t + P_b)/4 \approx PW_c.
\tag{3}
$$

where the P_l, P_r, P_t and P_b is the correlations calculated in Eq. 2 for the block-images positioned in left, right, top, and bottom hand of G_c with the center distanced 1 pixel form the center of G_c. In other words, we can reduced the correlation remains by using the pixel-correlations calculated from the adjacent block-images distanced 1 pixel from the center one in left, right, top, and bottom-hand.

3   Image Experiments

In this work, we have completed the 2 sets of image experiment in order to investigate the effectiveness of proposed method in detail and the results of the experiments are described as follows.

Fig. 1 shows the two sets of images in size 640×480 pixels, which consists of 8 pictures of portrait and town and represents the images containing low and high frequencies respectively. The major purpose is to compare the effectiveness of proposed method in the images consisting more high and low frequency components, respectively.

(1)   (2)

Fig. 1. Images of set 1 and 2 consisting of 8 pictures in size 640×480 pixels, indexed from left to right 1-4 in up ones and 5-8 in low ones, respectively.

In order to investigate in pixel-correlations in images of set 1 and 2, we calculated the histogram of differences among adjacent pixels in images of set 1 and 2 in such a way that for each pixel, the average of 4 differences between that pixel and its 4 adjacent pixels (left, right, up, and down-hand with 1 pixel from the center one) are calculated and then split into equal-sized bines. And then, the half count width (HCW) is calculated as a width of difference centering in the origin, within which the pixel counts is half of the total one.
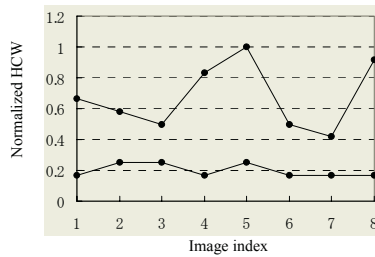
Fig. 2. HCW (half count width) of the images 1-8 in set 1 (low) and set 2 (up), normalized by the maximum in image 5 of set 2.

Fig. 2 shows the HCW of the images 1-8 in set 1 and 2, which are normalized by the maximum in image 5 of set 2. The normalized HCW in Fig. 2 indicates that the images of set 2 contain more high frequency components compared with images of set 1.

The watermark embedding was completed in the way as follows. First, we generated the pseudo-random sequences of period 4095 using an $n=12$ stage linear feedback shift register with an initiator 2048. The length of pseudo-random sequences is

4096, which generated in a length of 2^n with a copy of the first element to the last one so as to make the length of pseudo-random sequence as large as the size of a block-image in 64×64 pixels positioned in the image center. Second, we divided the center area of the images into 9 block-images with the same size of 64×64 pixels, and embedded the pseudo-random sequence into 9 block-images respectively. The watermark detection was completed with the proposed method using pixel-correlations information as described in previous section.

Fig. 3 (a) and (b) shows the correlations calculated between the center block of the images of set 1, 2 and the pseudo-random sequences with the initiator 1 to 4095, without using pixel-correlations, respectively. Fig. 4 (a) and (b) show the correlations calculated with the reduction of correlation remains using pixel-correlations, respectively.
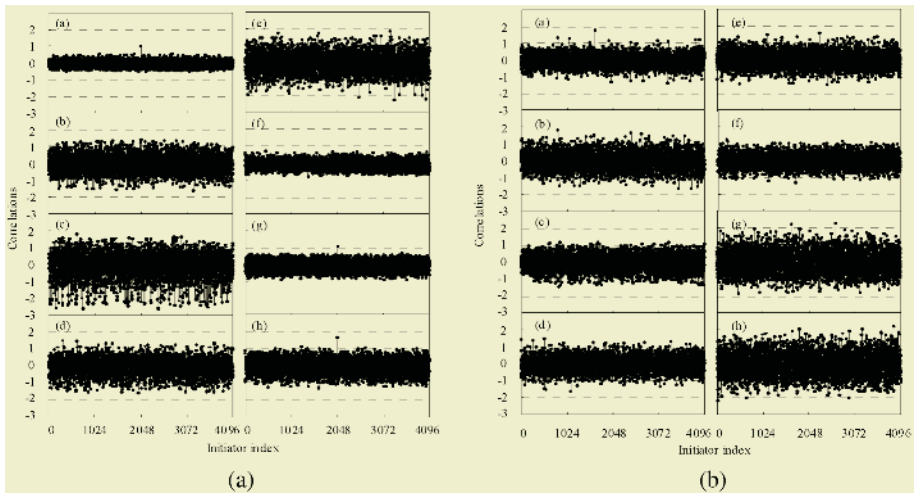
Fig. 3. Correlations calculated between the center block of the images and pseudo-random sequences with the initiator 1 to 4095, without using pixel-correlations. The graphics (a)-(h) in (a) and (b) correspond to the images 1-8 in set 1 and set 2, respectively.

Comparing the results showed in Fig. 3 and 4, except the case of image 8 in set 2, remarkable effectiveness of reducing correlation remains using pixel-correlations can be observed both in the cases of set 1 and 2. In order to examine the effectiveness in detail, we analyzed the experimental results as follows.

Table 2 and 3 show the statistical values calculated from the image of set 1 and 2, in which HCW/2 represents the half of HCW, and AVE1-2, VAR1-2, MAX1-2 represent the averages, variances, and maximums of the cross-correlations without and with the reduction in correlation remains, respectively. The AVE in the last row represents the average of HCW/2, AVE1-2, VAR1-2, and MAX1-2, respectively.

The graphics of AVE1-2, VAR1-2 in Table 1 and 2 are showed in Fig. 5 (a) and (b) respectively, which represents the variances and averages of cross-correlations without and with the reduction of correlation remains for set 1 and 2, respectively. Examining Fig. 5, we see that in both cases with and without using pixel-correlations, the variations of the graphics 1, 2 and the graphics 3, 4 are changed in the image with

about same tendency. This implies that we can use either of them as a representative for the investigations in the following examinations.
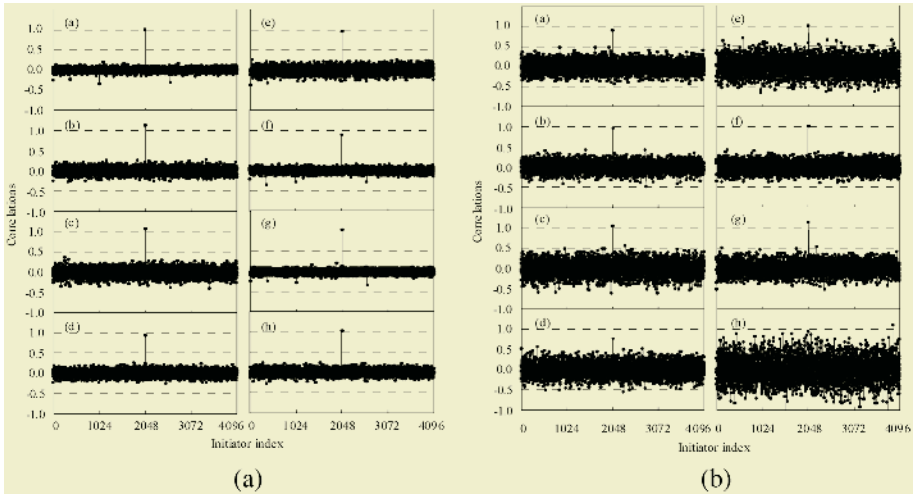
Fig. 4. Correlations calculated between the center block of the images and pseudo-random sequences with the initiator 1 to 4095, with the reduction of correlation remains using pixel-correlations from the 4 adjacent block-images. The graphics (a)-(h) in (a) and (b) corresponds to the images 1-8 in set 1 and 2, respectively.

Table 1. Statistical values calculated from the image set 1

NO	HCW/2	AVE1	AVE2	VAR1	VAR2	MAX1	MAX2
01	2	0.149	0.036	0.222	0.047	0.550	0.190
02	3	0.400	0.062	0.544	0.081	1.339	0.283
03	3	0.540	0.077	0.702	0.100	1.806	0.375
04	2	0.377	0.061	0.509	0.080	1.504	0.257
05	3	0.474	0.064	0.640	0.084	1.834	0.271
06	2	0.203	0.041	0.289	0.053	0.693	0.174
07	2	0.249	0.037	0.343	0.048	0.674	0.207
08	2	0.352	0.052	0.473	0.069	1.099	0.243
AVE	2.38	0.343	0.054	0.465	0.070	1.188	0.250

Table 2. Statistical values calculated from the image set 2

NO	HCW/2	AVE1	AVE2	VAR1	VAR2	MAX1	MAX2
01	8	0.284	0.104	0.383	0.137	1.759	0.406
02	7	0.403	0.094	0.541	0.123	1.822	0.433
03	6	0.349	0.132	0.455	0.171	1.183	0.565
04	10	0.345	0.129	0.459	0.168	1.505	0.554
05	12	0.350	0.158	0.472	0.206	1.604	0.701
06	6	0.326	0.096	0.432	0.125	1.260	0.439
07	5	0.521	0.112	0.696	0.147	2.269	0.529
08	11	0.528	0.223	0.710	0.292	2.189	1.025
AVE	8.13	0.388	0.131	0.519	0.171	1.699	0.582

Comparing the graphics 1 with 3 and the graphics 2 with 4 as showed in Fig. 5 (a) and (b), we see that with reducing correlation remains using pixel-correlations, both variances and the averages of the cross-correlations in cases of (a) and (b) are reduced remarkably.

Defining the ratio of an average of cross-correlations without and with the reduction of correlation remains as an improved rate of cross-correlation, the improved rate of average cross-correlations for the images in set 1 and 2 are 6.38 and 2.96, respectively. This indicates that with the increase of HCW, the effectiveness of using pixel-correlations will be decreased. The reason for this is that the pixel-correlations among adjacent pixels will decrease with the increase of the high frequency components in an image.

(a) The graphics of the images in set 1   (b) The graphics of the images in set 2

Image index

— △ — 1. Variance of cross-correlations without the reduction in correlation remains
— ■ — 2. Average of cross-correlations without the reduction in correlation remains
— □ — 3. Variance of cross-correlations with the reduction in correlation remains
— ▲ — 4. Average of cross-correlations with the reduction in correlation remains

Fig. 5. Variances and averages of cross-correlations calculated without and with the reduction of correlation remains; (a) and (b) corresponds to the results of set 1 and 2, respectively.

Moreover, comparing the graphics 1 with 3, and 2 with 4 in Fig. 5 (a) and (b) too, we see that with using pixel-correlations, the variations of variance and average of cross-correlations changed in image index become much gentle. This indicates that the higher the average of cross-correlation, the more effectiveness of reducing correlation remains using pixel-correlations can be acquired.

Fig. 6 (a) and (b) show the averages of cross-correlations without the reduction of correlation remains (graphics 1) and the HCW (graphics 2) in the images of set 1 and 2, respectively. Comparing Fig. 6 (a) with (b), an interesting fact can be observed that, although the averages of HCW of the images in set 1 are much lower than that of set 2, the average cross-correlations without the reduction of correlation remains are about same. The reason for this is that the difference between the total elements of 1 and 0 in a pseudo-random sequence is always 1, and the graphics 1 in Fig. 6 is an averages of the correlations calculated by using pseudo-random sequences with the initiator 1 to 4095.
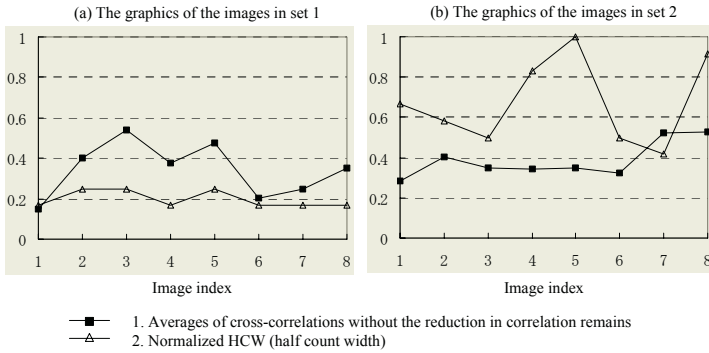
Fig. 6. Averages of cross-correlations calculated without the reduction of correlation remains (graphics 1) and HCW (graphics 2); (a) and (b) corresponds to the results of set 1 and 2, respectively.

4   Conclusions

In this paper, we investigated the effectiveness of the improved method for SSW techniques using pixel-correlations with image experiments in detail. The experimental results show that the correlation remains can be reduced remarkably using pixel-correlations.

Compared with the traditional method that calculates a precise threshold to determine the existence of watermark, the experimental results show that the proposed method can solve the problem of correlation remains substantially by reducing correlation remains using pixel-correlations among the adjacent pixels. The experimental results also show that the proposed method is effective to any kinds of images only if there are pixel-correlations among the adjacent pixels, and more effectiveness can be acquired if the images contain more low frequencies. The proposed method also can be used for SSW detection in other kind media, such as digital video and audio, only if there are pixel-correlations among adjacent pixels in frame images and the signals in the left and right channels [11].

References

1. Van Schyndel, R.G., Tirkel, A.Z, and Osborne, C.F.: "A digital watermark," Proceedings of IEEE ICIP 94, Vol. 2, (1994) 86-90
2. Cox, I.J., Joe Kilian, F. Thomson Leighton, and Talal Shamoon: "Secure Spread Spectrum Watermarking for Multimedia," IEEE Transactions on Image Processing, Vol. 6, No. 12, (1997) 1673-1687
3. Piva, A, Barni, M., Bartolini, F., and Cappellini, V.: "DCT-based Watermark Recovering without Resorting to the Uncorrupted Original Image," Proceedings of IEEE ICIP 97, Vol. 1, (1997) 520-523
4. Fridrich, J.: "Image Watermarking for Tamper Detection," Proceedings of IEEE ICIP 98, Vol. 2, (1998) 404-408
5. Zhu, W.-W., Xiong, Z.-X, and Zhang, Y.-Q: "Multi-resolution Watermarking for Images and Cideo: A Unified Approach," Proceedings of IEEE ICIP 98, Vol. 1, (1998) 465-468

6. Dugad, R., Ratakonda, K., and Ahuja, N.: "A New Qavelet-based Acheme for Qatermarking Images," Proceedings of IEEE ICIP 98, Vol. 2, (1998) 419-423
7. Marvel, L.M., Charles T. Retter, C.T., Boncelet, C.G.: "Hiding Information in Images," Proceedings of IEEE ICIP 98, Vol. 2, (1998) 396-398
8. Hartung, F., Su, J. K., Girod, B.: "Spread Spectrum Watermarking: Malicious Attacks and Counterattacks," Proceedings of SPIE, Vol. 3657, (1999) 147-158
9. Kutter, M., Winkler, S.: "A Vision-based Masking Model for Spread-spectrum Image Watermarking. " IEEE Transactions. Image Processing, Vol. 11, No. 1, (2002) 16-25
10. Liu, Z.: "Two Improved Methods for Spread Spectrum Watermark Technology." Proceedings of International Workshop on Digital Watermarking 2002, Nov. 21-22, Seoul, Korea (2002) 1-16
11. Liu, Z., Inoue, A.: "Audio Watermarking Techniques Using sinusoidal Patterns Based on Pseudo-random Sequences." IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 8, (2003) 801-812

Apply Semi-fragile Watermarking
to Authentication of Compressed Video Data

Tsong-Yi Chen[1], Da-Jinn Wang[2], Chih-Cheng Chiu[1], and Chien-Hua Huang[1]

[1] Department of Electronic Engineering, National Kaohsiung University of Applied Sciences,
Kaohsiung, Taiwan, R.O.C.
`chentso@cc.kuas.edu.tw`, `ss565575@ms19.hinet.net`
[2] Department of Information Management, National Kaohsiung Marine University,
Kaohsiung, Taiwan, R.O.C
`Wangdaj@mail.nkmu.edu.tw`

Abstract. This paper proposes an effective technique which can detect malicious manipulations under video lossy compressing data (e.g. H.263) and still-image lossy compressing data (e.g. JPEG). A block-classification strategy is used to divide DCT-blocks into the flat-blocks and the normal-blocks. Simple features of the both blocks are embedded invisibly. For later authentication, the watermarked frame is then put back into the compressed bit-streams.

The proposed method can detect and locate alterations of the tampered frame whether it is stored again into original bit-streams or not. This goal is for detecting tamper area and surviving most video lossy compression. Experimental results show that the proposed technique can detect various tampered areas and hence provides an effective image authentication for lossy compressed image and video in DVR (Digital Video Recorder) System.

Keywords: watermarking, H.263, JPEG, image authentication, compressed video.

1 Introduction

In the past, several techniques and concepts based on data hiding or steganography have been designed as a means for tamper detection in digital images and for image authentication − fragile watermarks, semi-fragile watermarks, and self-embedding. One class of authentication watermarks is formed by semi-fragile watermarks. Such watermarks are marginally robust and are less sensitive to pixel modifications. Thus, it is possible to use them for quantifying the degree of tamper and distinguish simple LSB shuffling from malicious changes, such as feature adding and removal.

More and more applications for tamper detection include authentication of digital data for courtroom evidence and copyright protection in DVR System. The reason is that the captured frame from DVR System will likely be maliciously tampered when they spread over internet or database. In this paper, we propose an effective technique which can detect malicious manipulations under video lossy compressing data (e.g. H.263) and still-image lossy compressing data (e.g. JPEG).

2 Proposed Data Embedding Scheme

The proposed watermarking process is based on 8x8-block DCT transform. The watermarking is performed on the Y component (luminance) of YCbCr color model. The basic process can be briefly described in the following.

At first, we hide the watermark into the chosen non-zero quantized AC coefficient (we called **NQAC**), denoted as $NQAC_i$, by backward zigzag-scan in each block. We set a threshold τ to divide the input frame into normal-blocks and flat-blocks. If the number of *NQAC* of the block is more than τ, the block is named a normal-block, the others is called a flat-block. The following subsections will illustrate how the watermark is embedded into these two blocks.

2.1 Normal-Block Embedding

In this paper, we also call the normal-blocks embedding as *Self-Embedding*. To consider about the speed of tamper detection, we can embed feature bit "1" into extra space as a key. We call the feature bit as *block classification bit*.

We extract other feature bits from quantized DC coefficient to embed into the chosen quantized AC coefficients. To emphasize the security of embedded feature bits, a fast one-dimensional pseudorandom number generating approach is used to shuffle the feature bits to disperse their energy relationship as feature bits f_i via exclusive-or operation where i denote *authentication strength*.

A degree of authentication according to *authentication strength*, but more authentication bits will make more degrade of image quality, vice versa. This is a trade-off between the quality of video and robustness of tamper detection. We embed the feature bits into LSB of the chosen nonzero quantized AC coefficients a_i by zigzag-scan.

The reason is that when the quantized AC coefficients are chosen to embed data, the artifacts caused by the replaced LSB's in the quantized coefficients will appear mostly in the sharp features in the image frames, like edges, lines, noise, etc., and will mostly be imperceptible. The quantized AC coefficients will be altered by data embedding as follows:

$$a_i' = \begin{cases} sign(a_i) * a_i, & if \ Bit_0(|a_i|) = f_i \\ sign(a_i) * AF(a_i), & if \ Bit_0(|a_i|) \neq f_i \end{cases} \quad (1)$$

Where "+1" or "-1" will be selected according to the sign of a_i. An *adjustment function*, AF, has two major features. In the first feature, the quantized AC coefficient "1" will be altered into "0" while the embedding feature bit f_i is "0". This will generate a data extracting fault due to the absence of quantized coefficients. The second feature is to transform the quantized AC coefficient 2 or -2 into 1 or -1 while the embedding feature bit is "1". This is an additional adjustment due to the result of the first feature. The *adjustment function* as follows:

$$AF(a_i) \Rightarrow \begin{cases} Bit_1(|a_i|) = f_i \oplus 1 & ,if \ |a_i| = 1 \\ Bit_1(|a_i|) = f_i \oplus 1 & ,if \ |a_i| = 2 \\ Bit_0(|a_i|) = f_i \end{cases} \quad (2)$$

For example, according to the result of *adjustment function*, the quantized AC coefficient "1" is "1", "-2" is "-1", "3" is "3", "-4" is "-5" while embedding feature bit is "1". Other quantized AC coefficient "1" is "2", "-2" is "-2", "3" is "2", "4" is "4" while embedding feature bit is "0".

121	130	136	129	112	98	94	95
127	133	139	136	125	111	101	97
132	134	138	142	139	127	109	95
130	129	131	138	142	133	110	92
125	122	122	129	135	128	109	93
124	122	121	123	124	121	113	106
130	130	129	123	118	118	123	128
137	139	137	127	117	118	132	146

(a)

132	136	135	126	111	97	91	90
136	137	137	131	119	107	97	92
138	137	137	135	129	118	104	95
136	134	133	134	133	124	109	96
131	129	128	130	131	125	112	101
130	128	127	127	126	123	117	113
133	133	131	128	124	124	126	129
136	138	136	130	124	125	133	141

(b)

Fig. 1. (a) 8x8 pixel-block, number of non-zero AC coeff. = 5 (by H.263, Qp = 7), (b) watermarking (a), authentication strength is 5 (worst-case).

2.2 Reduction of Clipping Errors

The quantized AC coefficient will likely be corrupted while enforcing normalization in spatial domain. The clipping error may be occurred due to that all pixels must be normalized to [0,255] in spatial domain. This is a *false alarm* for tamper detection. Find out maximum clipping error, and calculate its basis image from Eq. (3). The basis image with a scaling factor as a weight-mask (see Eq. (4)). Using the sign of the weight-mask to map the chosen quantized AC coefficient. Then, the maximum clipping error will not be raised any more. For example, as the Figure 2 (a)(b) shown, the pixel of coordinates (1,7) is 253 that into the maximum clipping error (269) will occur, and it will affect the chosen quantized AC coefficient. The basis image $B(x,y;u,v)$ is as

$$B(x,y;u,v) = \cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N} \tag{3}$$

from IDCT (inverse discrete cosine transform):

$$f(x,y) = \sum_{x=0}^{N}\sum_{y=0}^{N}\alpha(u)\alpha(v)C(u,v)\cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N}$$

The basis image with a scaling factor $\alpha(u)\alpha(v)$:

$$\alpha(u)\alpha(v)B(x,y;u,v) \tag{4}$$

184	161	161	173	170	183	226	252
126	122	133	193	235	244	253	253
45	80	116	194	244	244	253	250
55	56	103	200	244	232	237	247
72	49	111	193	225	239	248	252
85	75	140	177	192	241	250	134
88	69	104	123	131	164	163	136
66	75	94	66	68	80	90	74

(a)

182	160	160	171	168	182	225	254
126	121	132	193	235	243	264	269
44	80	116	193	243	244	256	249
54	54	101	199	243	230	235	246
71	47	109	191	224	237	246	257
84	75	140	176	191	241	250	233
87	68	103	123	131	163	162	136
64	74	93	84	66	79	89	72

(b)

Fig. 2. (a) 8x8 pixel-block; (b) (a) after non-linear quantizer (by H.263).

0.13	-0.17	0.16	-0.15	0.12	-0.1	0.07	-0.03
0.15	-0.2	0.19	-0.17	0.15	-0.12	0.08	-0.04
0.07	-0.09	0.09	-0.08	0.07	-0.05	0.04	-0.02
-0.03	0.05	-0..05	0.04	-0.03	0.03	-0.02	0.01
-0.12	0.17	-0.16	0.15	-0.12	0.1	-0.07	0.03
-0.17	0.24	-0.23	0.2	-0.17	0.14	-0.09	0.05
-0.16	0.23	-0.21	0.19	0.16	0.13	-0.09	0.05
-0.1	0.14	-0.13	0.12	-0.1	0.08	-0.05	0.03

(a)

160	-28	-5	1	2	2	0	0
17	-5	1	0	2	-1	-1	0
-8	13	6	-1	-2	0	0	0
7	0	3	0	0	1	1	0
-4	3	1	-1	1	0	0	0
-1	1	0	0	0	0	0	0
0	0	0	-1	-1	0	0	0
-1	0	0	0	0	0	0	0

(b)
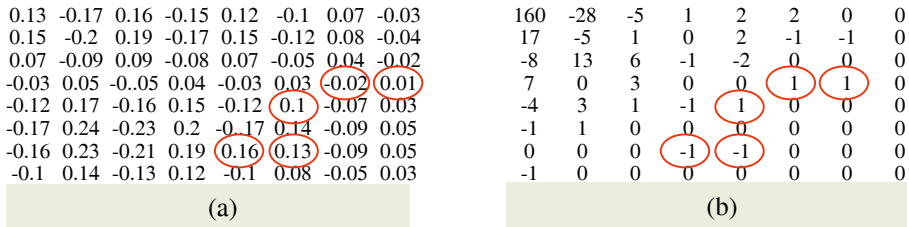
Fig. 3. (a) weight-mask for pixel (1,7); (b) authentication strength is 5.

Eq. (4) is a weight-mask which can map the 8x8 coefficient from DCT domain to spatial domain. As the Figure 3 (a)(b) shown, we must alter the sign of the chosen embedded 5 coefficient by weight-mask so that the pixel (269) will not be raised any more. It is useful for revertible capability of the chosen quantized coefficient.

Consequently, we can preprocess above two steps so that the chosen quantized AC coefficient can map back for the normal-block embedding

2.3 Flat-Block Embedding

For brevity, the flat-block embedding is described in the following points.

(1) Get a block classification bit: A feature bit "0" is embedded as an authentication key.
(2) Obtain and embed the watermark: Based on the characteristic of the few embedding capability in smooth region, we pick out the Bit5, Bit6, and Bit7 as the watermark bit from quantized DC coefficient, and embed they into the authentication-key. For the better trade-off between the robust of authentication and the capacity of authentication-key, we can replace previous pseudo-random number to with Bit0 and Bit1 so that we have 5 watermark bits to authenticate tampering blocks. It's very useful for decreasing strength of the authentication key and maintaining quality of the frame.

3   Proposed Tamper Detection Scheme

When we suspect that the protected frame is tampered by other users, we must confirm whether the frame is tampered or not. We proposed tamper detection approach as follow. First of all, extracting the eigenvalues of the normal-blocks from quantized AC coefficients, and another ones of the flat-blocks from key and Bit0, Bit1 of the quantized DC coefficients. Second, recovering the feature bits from the eigenvalues via previously pseudo-random number permuting. Third, two-step can detect tamper area of detected block. In the first step, we can use block classification bit from key to detect quickly each blocks which are tampered or not. In the second step, to calculate the Hamming Distance between the feature bits and compared bits which are extracted from the blocks. If the result of comparison has no difference, this block of the frame is not tampered. Otherwise, this block of frame is tampered, and the difference will be the positions of the tampered area of the original frame. We can see the proposed tamper detection scheme as Figure 4 shown.
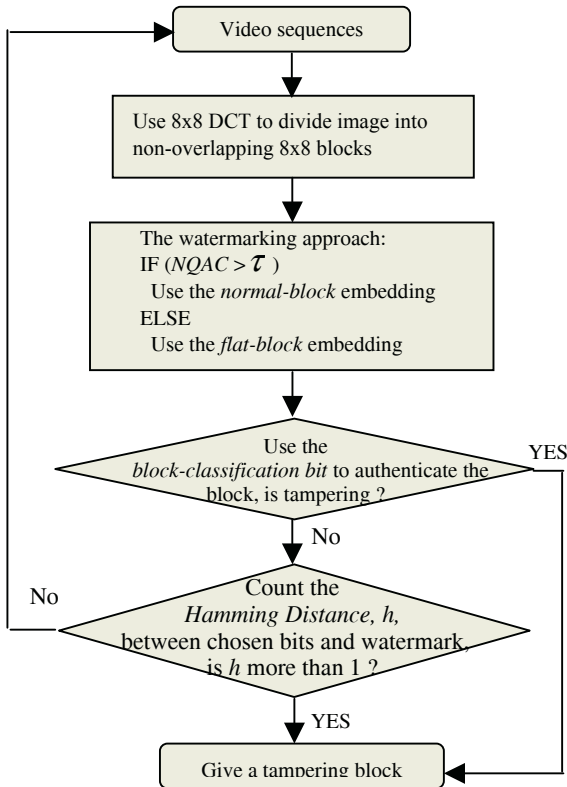
Fig. 4. The proposed block-based video tamper detection algorithm.

4   Experimental Results

The results are experiment by an IBM-compatible PC with Pentium Ⅳ 2.4 GHz CPU, 256 MB RAM, and the platform Windows XP. The image processing software is Ulead PhotoImpact 7.0. The digital video format of DVR System is Square pixel SIF (NTSC) (320 * 240).

The experimental results are discussed in two aspects. One is the quality of a frame after data embedding and extra space spending. Another is the authentication frame tamper detection. First we show the experimental results of a frame quality after data embedding, PSNR = 42.5 db, and the extra space spending is composed of the difference of original bit-streams and water- marked bit-streams, and the key. As Figure 6 shown. Number of the normal-blocks is 544, and the extra bit spending of the normal -blocks needs one bit. Number of flat-blocks is 656, and the extra bit spending of the flat-blocks needs three bits. the extra space spending is (12378 − 12169) + (3 * 656 + 544) / 8 = 523 bytes.

The key of extra space spending has two cases: the best-case is that all quantized DCT-blocks of the captured frame are normal-blocks (320 * 240 / (8 * 8) = 150 bytes), and the worst-case is that all quantized DCT-blocks of the captured frame are

flat-blocks (320 * 240 / (8 * 8) * 3 = 450 bytes). The region of extra space spending is (diff + 150 bytes) ~ (diff + 450 bytes). Diff denote the difference between original bit-streams and watermarked bit-streams.

5 Conclusion

In this paper, we propose an effective technique which can detect malicious manipulations on lossy compressed video data (e.g. H.263) and reduce the clipping-error problem. Although the quantized coefficients are more suitable than original coefficients for data embedding. It is noteworthy that the damaging problem of clipping errors can be caused by normalization in spatial domain. In this paper, we can reduce the problem of clipping errors.

Besides, we cannot embed too much data into the flat (smooth) block due to the value and number of the non-zero quantized AC coefficient is small and less. So, we use a specific approach to embed the watermark into the flat-block. Further, we can try to recover the tampered frame according to the few authentication-key.

Fig. 5. Original authentication frame (by H.263, Qp=7, bit-streams = 12169 Bytes).

Fig. 6. Watermarked frame, (PSNR is 42.5 db, bit-streams= 12378 Bytes).

Fig. 7. Illegal tampered frame.

Fig. 8. Tampered area by proposed approach.

References

1. Kwang-Fu Li, Tung-Shou Chen, and Seng-Cheng Wu, "Image tamper detection and recovery system based on discrete wavelet transformation," Communications, Computers and signal Processing, 2001 IEEE Pacific Rim Conference on, volume: 1, Aug. Page(s): 164 - 167 vol. 1.
2. M.Wu, and B. Liu, "Watermarking for Image Authentication," Proceedings of IEEE International Conference on Image Processing, Oct. 4-7, 1998, Chicago, Illinois, USA, vol. 2, pp. 437-441.

3. Tae-Yun Chung, Min-Suk Hong, Young-Nam Oh, Dong-Ho Shin, and Sang-Hui Park, "Digital Watermarking for Copyright Protections of MPEG2 Compressed Video," IEEE Trans. on Consumer Electronics, vol. 44, no. 3, Aug. 1998, pp. 895-901.

4. C.-Y. Lin, and S.-F. Chang, "Semi-Fragile Watermarking for Authenticating JPEG Visual Content," SPIE International Conf. on Security and Watermarking of Multimedia Contents II, vol. 3971, No. 13, EI '00, San Jose, USA, Jan 2000.

5. C.-Y. Lin and S.-F. Chang, "Issues and Solutions for Authenticating MPEG Video," SPIE International Conf. on Security and Watermarking of Multimedia Contents, vol. 3657, No. 06, EI '99, San Jose, USA, Jan 1999.

6. C.-Y. Lin, and S.-F. Chang, "A Robust Image Authentication Method Surviving JPEG Lossy Compression," Proc. IS&T/SPIE, San Jose, January 1998.

7. C.-Y. Lin, and S.-F Chang, "A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation," CU/CTR Technical Report 486-97-19, Dec. 1997.

8. L. M. Marvel, G. Hartwig, and C. G. Boncelet Jr., "Compression Compatible Fragile and Semi-Fragile Tamper Detection," SPIE International Conf. on Security and Watermarking of Multimedia Contents II, vol. 3971, No. 12, EI '00, San Jose, USA, Jan 2000.

9. E. J. Delp , C. I. Podilchuk, and E. T. Lin, "Detection of Image Alterations using Semi-Fragile Watermarks," SPIE International Conf. on Security and Watermarking of Multimedia Contents II, vol. 3971, No. 14, EI '00, San Jose, USA, Jan 2000.

10. N. D. Memon, P. Vora, and M. Yeung, "Distortion Bound Authentication Techniques," SPIE International Conf. on Security and Watermarking of Multimedia Contents II, vol. 3971, No. 15, EI '00, San Jose, USA, Jan 2000.

11. K. Toyokawa, N. Morimoto, S. Tonegawa, K. Kamijo, and A. Koide, "Secure Digital Photograph Handling with Watermarking Technique in Insurance Claim Process," SPIE International Conf. on Security and Watermarking of Multimedia Contents II, vol. 3971, No. 42, EI '00, San Jose, USA, Jan 2000.

12. Jiri Fridrich, and Miroslav Goljan, "Images with self-correcting capabilities," Proceedings of 1999 IEEE International Conference on Image Processing, volume: 3,PP. 24 -28 Oct. 1999, Page(s): 792-796 vol. 3.

13. I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," in Proc. IEEE Int. Conf. on Image Processing, Sep. 1996, vol. 3, pp. 243-246. atabases, San Jose, Jan. 1998.

An Approach
of Multi-level Semantics Abstraction

Hongli Xu and De Sun Zhijie Xu

Dept. of Computer Science & Technology, Beijing Jiaotong Univ.
Beijing, China 100044
`hlxu@center.njtu.edu.cn`

Abstract Semantics-based image retrieval is a challenging problem. In this paper, we propose an approach for image semantics abstraction, which constructs a multi-level semantics tree based on human subject and train hierarchical semantic classifier. According to our method, image features are selected by using priori knowledge. Then, those images are classified in every level by the classifier based on support vector machines (SVM). The SVM classifiers learn the semantics of specified classes from a training database of image. Experiments show that we can abstract multi-level semantics from image database by using less low-level features.

1 Introduction

Content-based image and video retrieval has always been the most popular way for information retrieval. Similarity match between low-level features is used to determine the retrieval interest. However, the retrieval result based on low-level feature similarity may not match with the user's perception of similarity. Nowadays, the mainstream of the research converges to the retrieval based on semantic meaning, whose goal is to extract cognitive concept of human by combining the low-level features in certain way. Image semantics abstraction has been the key to bridge the gap between low-level features and high-level semantics. Some researchers try to obtain semantic meaning from the multi-feature space by direct extraction of the objects [1], semantic visual template [2], supervised[3] or unsupervised training[4,5,6,8,9,10]and learning by relevance feedback[11,12]. These work are usually divided into two ways: one is from bottom to up, which maps low-level features to correspond to high-level abstraction; The other is from top to bottom, which use machine learning to obtain concept with low-level features.

In Visengine [2], it creates semantic visual template for main regions with plenty semantic information on the basis of weighted feature combination. Ana B. Benitez[5] proposes novel methods for the knowledge discovery integrating both the processing of images and annotations. Aleksandra Mojsilovic [7] models the semantic categories based on human perception and proposes a set of low-level image features to describe each category. D. Brahmi [11] proposes a model which makes use of the interaction between the user and the metadata, and which can find the semantic users' preference by analyzing their answers through the Relevance Feedback process. Philippe H. Gosselin [12] presents an approach to exploit the knowledge provided by user

interaction. The semantic annotations can be integrated into the similarity matrix of the database images.

Aditya Vailaya[6,8] shows that a specific high-level classification problem can be solved from relatively simple low-level features using Bayesian[6] and weighted k-NN Classifier[8] respectively. C.Town [9] catalogs images through construct neural network classifier for each semantic meaning. Dezhong Hong[10] uses fuzzy K-mean classifier to perform supervisory classification based on low-level features, deal with classification overlap problem.

In these methods, whether from bottom to up or from top to bottom, users have to participate to name the image mean and to select the discriminative features during defining image semantic meaning. However, the relations between images semantic meanings are rarely considered. In this paper, we will give a simple and effective semantics abstraction algorithm, which constructs a multi-level semantics tree based on user regarding the semantic relation, and chooses features according to priori knowledge, and then classifies the images using the SVM classifier. The algorithm can obtain the class semantics and its subclass semantics at the same time. Our experiments show that the semantic abstraction algorithm is effective.

The paper is organized as follow. In Section 2, we propose a multi-level semantics tree by human subject, and discuss the preprocessing steps including feature selection. In Section 3, we give the multi-level semantics classifier based on SVM. In Section 4, we present the experimental results of multi-level semantics abstraction for a set of five semantic concepts. In Section 5, we conclude the application of the multi-level semantics abstraction and present our further work.

2   Multi-level Semantics Tree by Human Subject (MLST)

Human usually understand image semantics that are hierarchical, just as Fig.1. If we analyze the hierarchical semantic structure, we can discover the relation between levels: lower level semantic meaning inherits higher level semantics, and only the same level semantics need be classified. For example, inflorescence and flower possess the first level flowerage' meaning in the secondary level semantics. Therefore, when we extract lower level semantics, we may discard the redundant features, and select the best discriminating features for classifier. User can build MLST by himself, so that the classifier based on MLST is suitable for semantic varieties and user' requirement.
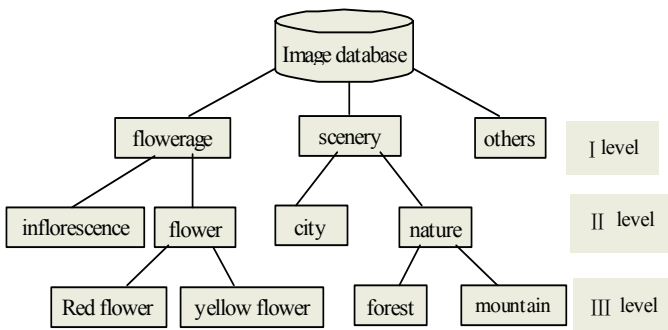
Fig. 1. Multi-level Semantics Tree By Human Subject

In our experiment, we construct the multi-level semantics classifier according to the multi-level flowerage semantics tree. Our system allow users to create their multi-level semantics tree, and select each level features to construct classifier by themselves. In addition, we may automatically select most discriminative features by PCA and entropy methods. For example, inflorescence classifier only uses texture feature, but the classifying precision is enough high.

3   Image Semantics Classifier

3.1   Image Semantic Description

In image classification based on SVM, every semantic class is considered as independent two-class problem to improve classifier precision. Given m classes in the image database $L = \{A_1, A_2, \ldots A_m\}$, each semantic class A_i has N_i pictures. We translate the classification problem of m classes into that of two classes: For any A_i, positive case is from images of the class; negative case is from the other. Assuming A_i has N_i positive cases, the number of negative case is $\sum_{j=1, j \neq i}^{m} N_j$.

For $\forall A_i \in L$, the training collection T is

$$T = \{(x_1, y_1), (x_2, y_2), \ldots (x_l, y_l)\}$$

$\{x_i, y_i\}$ is defined as a sample with semantic meaning.

An image feature vector is denoted as $x_i \in R^n$, such as color, texture, shape etc.

$y_i \in \{+1, -1\}$, $y_i = +1$ represents $x_i \in A_i$, the image with x_i attribute semantic A_i class; $y_i = -1$ represents $x_i \notin A_i$.

The feature vectors of the trained images are the input of SVM classifier. Semantic A_i classifier is confirmed. SVM-based classification integrate human perception for image. Through training, classifier is going to select the fittest low-level feature collection to simulate human visual understanding. According to semantic inheritance relation in multi-level semantics tree, we train SVM-based classifier. After performing the test for testing image collection, we can capture the approximate relation of a sample corresponding semantic meaning, and obtain multi-level semantics abstraction.

Given SVM_Model_i is a trained SVM classifier with semantic A_i, semantic A_i is described as: { A_i, SVM_Model_i, Fea_i }

A_i is semantic word, such as flower, forest, city etc. SVM_Model_i is a trained SVM classifier with corresponding semantics, and Fea_i is low-level features of A_i. In our experiment, Fea_i is represented as {fi,fj,…,Num}, fi∈(color, texture, shape,…), and num is the sum of feature dimension.

Our method is a mapping from high-level semantics to low-level feature by using SVM classifier to model SVM_Model_i.

3.2   Multi-level Semantics Classification

Multi-classification problem is realized by decision-making tree based SVM method. In decision-making tree, each node represents a classifier, and each divaricating is an output, and leaves represent classes.
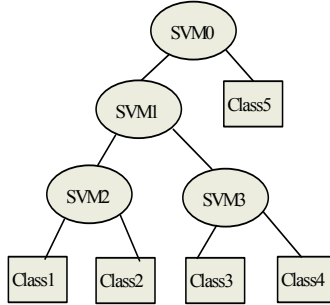
Fig. 2. SVM decision-making tree

For Multi-classification, if there is enough prior knowledge, the tree structure is unbalancing binary tree, shown in Fig.2. In the paper, we use prior knowledge to guide leaf nodes partition and construct a best SVM decision-making tree. Assuming a picture is not classified, we can deduce the SVM_Model_i by the decision-making tree, and extract the Fea_i as the input of SVM_Model_i. Then, pictures is classified level by level and in the end, the picture will be classified to A_i and return A_i semantic word at the same time.

Our semantics extracting procedures are described as follows:

Step 1 train A_i classifier.

Step 2 extract low level features of test pictures according to Multi-level Semantics Tree.

Step 3 input low level feature vectors to SVM_Model_i.

Step 4 conclude if the picture includes A_i semantics.

Step 5 if step 4 is 'Yes', return A_i description; if not, return other.

4   Experiment Results

4.1   Experiment Environment

In our experiments, the image database consists of 1500 images, including flower, forest, mountain, and sky. Multi-level flower semantics sample images are shown in the Fig.3. Experimental environment includes Intel (R) 4 PC, 256M memory, Windows 2000, Microsoft Access 2000, VC++6.0. Table 1 gives the training and testing number of every class image. Table 2 is image features in our experiment.

There are three kernel functions: Polynomial kernel function, Gaussian kernel function and Sigmoid kernel function. We select the Gaussian kernel ($\sigma=0.5$).

inflorescence   ikebana   flower

Fig. 3. Multi-level flower semantics of flowerage

Table 1. The training and testing number of every class image

Class	Ikebana	Flower	Inflorescence	Sky	Forest	Mountain
Training	40	50	50	20	30	30
Testing	40	100	80	20	45	30

Table 2. Abbreviations of image features

Feature	abr.	Feature	abr.
HSV Cumulative Histogram Training	HCH	RGB Moment	RM
GRAYCumulative Histogram	GCH	Co-occurrence	COO
COLOR Histogram	CH	COLOR Cumulative Histogram	CCH

4.2 Multi-level Semantic Classifier

For evaluating experiment result, we define precision rate and recall rate:
precision rate=the number of classifying correct/ the total number of the class
recall rate= return correct number / the return number

Table 3. Experiment results of flowerage classifier

precision recall	HCH/COO	RM/COO	GCH	GCH/COO	HCH/RM/COO	GCH/HM/COO
flowerage	**0.83**	0.85	0.57	0.73	0.86	0.73
	0.94	0.84	0.61	0.58	0.84	0.64

Table 4. Inflorescence and flower classifier results in flowerage

Precision recall	COO	RM/COO	CC/COO	GC/COO	HC/RM/COO	GC/RM/COO
inflorescence	**0.73**	0.69	0.75	0.58	0.65	0.70
	0.90	0.87	0.82	0.62	0.72	0.70
flower	0.69	**0.85**	0.67	0.71	0.71	0.72
	0.67	**0.82**	0.87	0.87	0.85	0.90

According to Table 3 and Table 4, we can conclude:

1) The different semantic class should be represented by different feature collection. Our results show that HSV cumulate histogram and gray co-occurrence describe flowerage very well, and the classifier discriminate flowerage from others (sky, mountain, forest).

2) We cannot use the same feature to discriminate the different sub-class in the same class, such as inflorescence and flower. And the fact is not true that more features there are, better the performance is. For example, the precision of inflorescence described with COO is 0.73, and the recall is 0.9, but after adding GC, the precision is low to 0.58, and the recall is 0.625.

3) In flowerage class, because inflorescence pictures include various little flowers, and main objects are gathered together in flower and ikebana pictures, we select the texture feature to train the inflorescence classifier. For discriminating flower and ikebana, RGBMOMENT is more effective.

4) Every semantic class is represented as:

{Flowerage, *SVM_Model_flowerage*, HCH, COO }
{Inflorescence, *SVM_Model_inflorescence*, COO }
{Flower, *SVM_Model_flower*, RM }

When training independently inflorescence, ikebana and red flower classifiers to distinguish from others (sky, mountain, forest, etc.), Table 5 shows the experiment results comparing with MLST. Obviously, Multi-level semantic classifier is more effective in classifying inflorescence and flower. It is not better than the independent classifier for red flower. Because the SVM based classifier is not good for the less difference between positive and negative case. When we use PCA method to classify them, the precision is 0.65 and the recall is 0.825

Table 5. MLSC comparing with independent classifier

Classifier precision	flowerage	inflorescence	flower	red flower
Independent classifier	0.83 (HCH/COO)	0.44 (COO)	0.54 (RM/COO)	0.24 (HCH/COO)
MLSC	0.83 (HCH/COO)	0.78 (COO)	0.85 (RM/COO)	0.35 (HCH/COO)

5   Conclusions and Further Work

In this paper, we propose an approach of multi-level semantics abstraction. According to multi-level semantics tree, we construct multi-level semantics classifier, and use prior knowledge to select low-level features. Multi-classification problem is solved by decision-making tree based SVM method. Experiment Results show this method is effective and semantic retrieval can be implemented by the method. Our further work include three aspects: a) to implement the MLST for more classes of image, b) to improve the method of the low-level feature selected and consider what MPEG-7 has defined, and c) to select the different classifier model by the feature difference.

Acknowledgements

References

1. Chad Carson, et.al. "Blobword: Image Segmentation Using Expectation-maximization and Its Application to Image Querying". IEEE Trans. PAMI. Vol.24, No.8, pp.1026-1038, Aug. 2002

2. JIAN-YONG SUN, ZhENG-XING SUN et.al "A SEMANTIC-BASED IMAGE RETRIEVAL SYSTEM:VISENGINE" Proceedings of the First International Conference on Machine Learning and Cybernetics, Beijing, 4-5 November 2002

3. J.Z.Wang,J.Li,and G.Wiederhold,"SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries"IEEE Trans.Pattern Anal.Machine Intell , vol .23, No9 . 2001

4. G.Sheikholeslami,W.Chang,and A.Zhang,"SemQuery Semantic Clustering and Querying on Hetergeneous Features for Visual Data" IEEE Trans.Knowledge and Data Engineering,vol.14,no 5,2002

5. Ana B. Benitez, Shih-Fu Chang. Automatic Multimedia Knowledge Discovery, Summarization and Evaluation. IEEE Transactions on Multimedia, 2003
   (http://www.ee.columbia.edu/dvmm/publicationPage//Author//AnaB.Benitez.html)

6. A Vailaya , A Jain M Figueiredo, H J Zhang. Content-Based Hierarchial Classification of Vacation Images .IEEE Int'l Conf .On Multimedia Computing and Systems ,Jun. 1999

7. eksandra Mojsilovic and Bernice Rogowitz,. Capture image semantics with low-level descriptors.Image Processing, 2001. Proceedings. 2001 International Conference on Volume: 1 , 7-10 Oct. 2001 .Pages:18 - 21 vol.1

8. A . Vailaya , A. Jain and H .J .Zhang , On Image Classification :City Images vs Landscapes, Pattern Recognition , Page : 1921-1935 , 1998

9. C.Town,et al."Content Based Image Retrieval Using Semantic Visual Categories " TR-2000,14,AT&T Lab.

10. DeZhong Hong et.al.Refining Image Retrieval Based on Context-Driven Methods.SPIE v.3656.PP581-592,1999

11. D. Brahmi, D. Ziou, Improving CBIR Systems by Integrating Semantic Features, 1st Canadian Conference on Computer and Robot Vision (CRV'04) May 17 - 19, 2004 pp. 233-240

12 . Philippe H. Gosselin, Matthieu Cord, Semantic Kernel Updating for Content-Based Image Retrieval IEEE Sixth International Symposium on Multimedia Software Engineering (ISMSE'04) Miami, Florida December 13 - 15, 2004 pp. 537-544

Effective Video Scene Detection Approach Based on Cinematic Rules

Yuliang Geng, De Xu, and Aimin Wu

Institute of Computer Science and Technology,
Beijing Jiaotong University, Beijing, 100044, China
gengyuliang@hotmail.com, xd@computer.njtu.edu.cn,
wuaimin@dongying.gov.cn

Abstract. Scene detection is an essential step to organize the video data properly for content-based video analysis and its application. In this paper, an effective scene detection approach is proposed, which exploits the cinematic rules used by filmmakers as guideline to compute shot similarities and identify the video scenes in narrative film. First, a clustering method with time constraint is used to group the shots into scene slices. Then dialog scene's alternative structure and audio correlation are used to identify dialog scene. Finally, motion and audio correlation guided by cinematic rules are used to further detect action scene. Experimental results show that the proposed method works well and can deal with complex scene.

1 Introduction

Extracting scene structure information of film will help create film summaries, and enable a meaningful nonlinear navigation of film. There are three main scene categories: dialogs without action, actions without dialog and dialogs with action. These are of course simplified categories, but most of real scenes fall within them [1], so this classification is essential to the study of video structure and extraction of high semantic information. There are a number of algorithms proposed to detect the scene boundary in recent literatures [2-6]. Most of prior work is focused on visual information. In [2], both motion and color features were utilized to compute shot similarities and a graph partition algorithm based on undirected graph was employed to compute video scene. In [3,4], color feature was utilized to identify scene boundaries.

As the film data is a rich mixture that involves audio, video and text information, many researchers used both audio and video information to detect scene boundary [5,6]. In [5], audio features based on shot were extracted so as to avoid the conflicts between the audio and video data. But it was not a proper way to compute audio statistical features in a shot. In [6], the audio stream was first classified into four semantic classes, namely, silence, music, speech and environmental sound segments, and then the clustering based on audio features and visual features are used to identify the scene boundaries respectively. However, the conflict between the audio and video segments is a difficult problem. The cinematic rules are also important clues to detect the scene boundaries. Some researchers tried to exploit the cinematic rules to guide the video scene parsing or refine the detection results [3, 4]. But how to quantify cinematic rules effectively is a still a difficult problem.

As mentioned above, exploiting multiple features involved in video is an effective method to detect the scene boundaries. In this paper, we employ the cinematic rules as guideline to extract features, compute shot similarities and identify the scene structures of narrative film. First, video sequence is divided into shots through shot boundary detection; Color and motion features based on shot are extracted; Audio features based on semantic segments are also extracted. Then, a clustering method based on color feature with time constraint is used to group the shots into scene slices. Next then, dialog scene's alternative structure and audio correlation are used to identify dialog scene. Finally, motion and audio correlation guided by cinematic rules is used to further detect the action scenes. Experimental results show that the proposed method works well and can deal with complex scenes.

The organization of the paper is as follows. Section 2 gives an overview of the cinematic rules. Feature extraction and the proposed algorithm are presented in Section 3 and 4. Section 5 and 6 give the experimental results and draw the conclusion.

2   Cinematic Rules for Scene Analysis

Cinematic rules is a group of rules to define how the cut connection is made simply [1]. Understanding cinematic rules, which are used by filmmakers to construct meaning, will help to understand the scene structure of film [1,7].

Rule 1. The 180° Rule: The 180° rule dictates that the camera should stay in one of the areas on either side of the axis of action.

Rule 2. Action-matching Rule: Action-matching indicates motion direction should be the same on in two consecutive shots that record the continuous motion of actor.

Rule 3. Film Rhythm Rule: Film rhythm indicates the perceived rate and regularity of sounds, series of shots, and motion within the shots. Rhythmic factors include beat, accent, and tempo. In same scene, the shots have the similar rhythm.

Rule 4. Establishing Shot: Establishing shot shows the spatial relations among the important figures, objects, and setting in a scene. Usually, the first and last few shots in a dialog scene are establishing shots.

3   Features Extraction

3.1   Audio Features

As film rhythm rule said, the shots in the same scene have similar rhythm, and audio is an important rhythmic factor. That is to say, the shots may have same or similar environment sounds and topic sounds in the same scene. To express this nature of scene effectively, we extract the audio features based on the semantic segmentation of shot. First, the audio stream is classified into four types of semantic segments: silence, speech, music, and environment sound in each shot [8]. Then the ratio of the number of frames in semantic segment i to the total number of frames in shot a is computed, and is denoted as $RSS_i(a)$. Except for the silence segment, we respectively extract the audio features of each semantic segment: Short Time Average Energy

(*STE*), Frequency Centroid (*FC*), and Band Width (*BW*) [9]. The mean of each feature in corresponding semantic segment is computed and denoted as $STE_i(a)$, $FC_i(a)$ and $BW_i(a)$. Then we normalize them using Gauss normalization formula and represent the audio features of semantic segment i in shot a as a triplet, and denotes it as $\mathbf{NAF_i(a)}$.

3.2 Visual Features

First, we extract color feature in this section. The video is first divided into shots [10] and each shot is represented by a set of key frames [11]. The key frame is represented with the HSV color histogram. We employ the histograms intersection to represent color similarity of frames f_i and f_j, denotes it as $D(f_i, f_j)$.

In film editing, motion is an important rhythmic factor, and the filmmaker often uses a series of short shots or shots with high motion to create tense or strong atmosphere. And two scenes with high motion rhythm may not be juxtaposed together. Here we extract shot change rate (*SCR*) to describe this attribute. Moreover, filmmaker obeys the action-matching rule to sustain the motion continuity. So we extract the dominant motion direction histogram of shot (*DMD*) to materialize this rule.

In [12], Xue proposed a method for computing dominant motion of single frame; we employ it to further compute the dominant motion of shot. Let S_{vm} denote the set of valid motion vectors and N_{vm} denote the number of valid motion vectors in frame. The motion vector direction histogram, $H_{md}(i, j)$, represents the percentage of vectors at the jth direction in the valid motion vectors set of frame i.

As the motion continuity, we utilize a sliding window to estimate the dominant motion histogram of shot a, $DMD_a(k)$. First, the average histogram in the sliding window is calculated, if the dominant motion direction of current frame consists with the dominant motion direction, m, of the average histogram in the sliding window, then $DMD_a(m)$ is added by one. Next then move the current frame over frame sequence at one-frame increment, and the sliding window moves consequently. Repeat it until the all frames have been computed. Thus, we obtain the dominant motion direction histogram of shot a and normalize it as $NDMD_a(k)$.

Since the magnitudes of the motion vectors reflect the activity/energy of shot, we compute average motion energy of shot a, $ME(a)$, to depict shot activity. The shot cut frequency, $SF(a)$, is defined as the inverse of shot length. Then, The shot change rate (*SCR*) is computed as the weighted sum of the normalized shot average motion energy $NME(a)$ and the normalized shot cut frequency $NSF(a)$ as

$$SCR(a) = \alpha \times NME(a) + \beta \times NSF(a) \tag{1}$$

where the α and β is the weight value, and $\alpha + \beta = 1$. Here we set $\alpha = \beta = 1/2$.

4   Scene Boundary Detection

4.1   Scene Slice Clustering

As the 180° rule said, the shots in same scene may have similar background and lighting, so they should have similar visual feature. We utilize color feature extracted in

section 3.2 to cluster the similar shots into one scene slice. First, the inter-shot similarities are computed in sliding window, and the shots whose similarities are greater than a given threshold Th_{vis}, are clustered into the same class. The visual similarity of shot a and shot b is defined as

$$VSS(a,b) = \max_{i \in KFS_a, j \in KFS_b} (D(f_i, f_j)) \tag{2}$$

where KFS_x is the key frame set of shot x.

Then move the sliding window over shot sequence at one-shot increment. Compute the similarities between the new shot cur and other shots in the sliding window, and label shot cur as the class of the shot that has the maximum similarity with shot cur. If the maximum similarity is less than the threshold Th_{vis}, a new class is created. Repeat it until all the shots have been computed. Thus, each shot is clustered into one scene slice (SS) and has a class label of its scene slice.

4.2   Dialog Scene Detection

As cinematic rule 4 said, a typical dialog scene consists of two type shots. Establishing shots give us the overall spatial relationship, and alternative shot/reverse shots describe the interaction between actors in detail. Many literatures have exploited the alternative visual structure to detect dialog scene. Here we employ our prior work [13] to cluster shots into shot slices. As parallel scene has the same alternative structure as dialog scene does [1], it is often mistaken for a dialog scene. We compute the audio correlation between shots in the different scene slices to distinguish them because the two scene slices in parallel scene have different audio effects that represent different scene atmosphere. The audio correlation is defined as

$$AudioCor(a,b) = \sum_{i=1}^{4} \alpha_i \times f(\mathbf{NAF_i(a)} - \mathbf{NAF_i(b)}) \tag{3}$$

$$f(\mathbf{NAF_i(a)} - \mathbf{NAF_i(b)}) = \begin{cases} 1 & \text{if } i \text{ is silence} \\ 1 - \sum_j \beta_j \mid NAF_i^j(a) - NAF_i^j(b) \mid & \text{else} \end{cases} \tag{4}$$

Where $\alpha_i = \min(RSS_i(a), RSS_i(b))$, and β_j is the weight value and $\sum_j \beta_j = 1$, here is

set as 1/3. $NAF_i^j(x)$ is the jth component of $\mathbf{NAF_i(x)}$.

If the average of audio correlation between the shots in the interleaving scene slices is greater than a given threshold Th_{aud}, the shots are merged; otherwise, the interleaving shots belong to different scene respectively.

Because establishing shots are often the first or last shots in dialog situation as rule 4 said, we identify establishing shots as follows.

For a dialog scene $\{S(i), S(i+1),\ldots, S(i+m)\}$, if the shot $S(i-1)$ or $S(i+m+1)$ does not belong to any scenes, the visual similarity $VSS(S(i-1), S(i+m+1))$, is computed. If the similarity is greater than the given threshold Th_{vis}, then this two shots are merged into this dialog scene as establishing shots.

4.3   Action Scene Detection

Action scene consists of a series of shots with significant change in locations, such as the tracking scene. So the shots in action scene have low visual similarities and weak structure, and after the first two steps, the action scene is divided into several scene slices or single shots. How to construct the action scene is still a difficult problem. Fortunately, action-matching and film rhythm rules give us a guideline to further detect action scene.

As the action-matching rule said, two consecutive shots that record the continuous motion of a performer should have consistent motion direction. At the same time, the consecutive shots that have similar shot change rate may belong to one scene. Another important clue is audio features discussed in section 3.1. So we exploit shot's dominant motion direction, shot change rate and audio features to detect action scene.

First, we define the motion correlation between shot a and b as follows

$$MotionCor(a,b) = \max(SimSCR(a,b), SimDMD(a,b)) \tag{5}$$

where $SimSCR(a,b)$ is the similarity between a and b on shot change rate; $SimDMD(a,b)$ is the similarity between a and b on dominant motion direction.

$$SimSCR(a,b) = 1 - |SCR(a) - SCR(b)| \tag{6}$$

$$SimDMD(a,b) = \sum_{i=1}^{4} \min(NDMD_a(i), NDMD_b(i)) \tag{7}$$

Then, the action scene detection algorithm is implemented as follows:

If two adjacent shots a, b do not belong to any scene slices, calculate the motion correlation and audio correlation respectively. One of the following conditions holds, the two shots are merged.

$$MotionCor(a,b) > Th_m; AudioCor(a,b) > Th_{aud} \tag{8}$$

where Th_m and Th_{aud} are given thresholds.

If a single shot a and a scene slice SS are adjacent, the correlation between shot a and scene slice SS is computed as Eq. (9) shown, that is, the correlation between shot a and the shots in window M that is neighboring to shot a in scene slice SS. If the correlation is greater than a given threshold (Th_m for motion feature, Th_{aud} for audio feature), the shot a is merged into scene slice SS.

$$ShotSceneCor(a, SS) = \max(\min_{i \in M}(MotionCor(a,i)), \min_{i \in M}(AudioCor(a,i))) \tag{9}$$

If two scene slices are adjacent, calculate the correlation between them as follows.

$$SceneCor(SS_1, SS_2) = \min_{i \in M_1}(ShotSceneCor(i, SS_2)) \tag{10}$$

M_1 is a window that is neighboring to SS_2 in scene slice SS_1. If two scene slices, SS_1 and SS_2 have qualified audio correlation and motion correlation, merge them.

5 Experimental Results

To evaluate the performance of proposed scene detection algorithm, we collect the test data set from five narrative movies, namely, *Rain man*, *Leon*, *Shrek*, *The Shaolin Temple* and *Ghost*. For the subjectivity of the scene boundary, we employ a group of human observers to identify the scene boundaries as ground truth. Table 1 summarizes the test videos and the experimental results. The experimental results at different steps are also shown as I, II and III in the fifth column.

As the experimental results shown, the scene slice clustering is very effective method to detect the scene without action, but it is not enough to detect dialog and action scenes. In the second step, the alternative structure is used to identify dialog scene and achieve good performance. Most establishing shots are identified. And one parallel scene happened in *Ghost* is correctly distinguished from dialog scenes by using audio correlation. At the third step, the audio and motion correlation functions effectively eliminate the positive false scenes caused by motion. Of course, there exist few missed scenes and positive false scenes. The missed scenes mainly occurred at the first step because these scenes have very color similarity with adjacent scene. The positive false scenes are due that the action scene that involves irregular motion is split into several scene slices.

Table 1. Experimental result of video scene detection approach

Video	Frame #	Shot #	Ground Truth Scene #	Detected Scene #			Miss #	Positive False #
				I	II	III		
Rain Man	35040	189	15	70	46	17	1	3
Leon	12118	113	9	51	28	11	0	2
Shrek	20280	201	8	57	38	8	0	0
The Shaolin Temple	16719	168	9	78	42	11	1	3
Ghost	28698	207	12	70	28	12	0	0

Table 2 shows the experimental results compared with the approach [2], denoted by GTA. The experimental results show that our method, denoted by AbCR, can obtain better performance, and can deal with the more complex scenes. For example, most action scenes can be detected correctly in AbCR. For dialog scene, AbCR has more precise result, such as establishing shot detection, parallel scene distinguish.

Table 2. Experimental results compared with GTA. Video data is from *Rain man* and *Leon*

Algorithm	Video	Frame #	Shot #	Ground Truth Scene #	Detected Scene #	Miss #	Positive False #
AbCR	*Rain Man*	35040	189	15	17	1	3
GTA					21	1	7
AbCR	*Leon*	12118	113	9	11	0	2
GTA					13	0	4

6   Conclusions

In this paper, we presented an approach for video scene detection in narrative film. Since the video data is a rich mixture, we utilize color, motion and audio features to detect video scene. Meanwhile, the cinematic rules used by filmmakers are employed as guideline to extract feature and identify scene structure. Experimental results show that our method works very well and can deal with complex scene. Of course, there exist failed examples. Thus, in video scene detection, how to exploit the cinematic rules and video features involved in video effectively is our further work. Extracting highlight semantic scene is another important work.

References

1. Arijon, D.: Grammar of the film language. Silman-James Press, Los Angeles CA (1991)
2. Lu, H., Tan, Y.P.: An Efficient Graph Theoretic Approach to Video Scene Clustering. In: Proceedings of ICICS-PCM, Vol. 3. (2003) 1782–1786
3. Truong,B.T., Venkatesh S.,Dorai C.: Scene Extraction in Motion Pictures. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13. Issue.1. (2003) 5–15
4. Tavanapong W., Zhou J.: Shot Clustering Techniques for Story Browsing. IEEE Transactions on Multimedia, Vol. 6, Issue. 4 (2004) 517–527
5. Chen, S.C., Shyu, M.L., Liao, W., Zhang, C: Scene Change Detection by Audio and Video Clues. In: Proceedings of IEEE ICME, Vol. 2 (2002) 365–368
6. Zhu, Y.Y, Zhou, D.R.: Scene Change Detection Based on Multimodal Integration. In: Proceedings of SPIE, Vol. 5286 (2003) 968–973
7. http://classes.yale.edu/film-analysis/index.htm
8. Lu, L., Jiang, H., Zhang, H.J.: A Robust Audio Classification and Segmentation Method. In: Proceedings of the ninth ACM international conference on Multimedia (2001)
9. Wang, Y., Liu, Z., Huang, J.C.: Multimedia Content Analysis-Using Both Audio and Visual Clues. IEEE Signal Processing Magazine, Vol. 17, Issue. 6 (2000) 12–6
10. Geng, Y.L., Xu, D.: A Unified Framework for Shot Boundary Detection. Journal of Image and Graphics (Chinese) Vol.10, No.5 (2005)
11. Geng, Y.L., Xu, D.: Efficient Key Frames Extraction Based on Hierarchical Clustering. In: Proceedings of IEEE TENCON, Vol. B (2004) 77-80
12. Xue, X.Y., Xiao, Y.N., Ding, C.M.: Qualitative Analysis of Dominant Motion for Compressed Video Stream. Chinese Journal Computers, Vol. 25 No.12 (2002) 1428–1433
13. Cheng, W.G., Xu, D., Lang, C.Y.: An Efficient Method for Video Scene Detection. Journal of Image and Graphics (Chinese), Vol. 9 No.8 (2004) 984–990

3-DWT Based Motion Suppression
for Video Shot Boundary Detection

Yang Xu, Xu De, Guan Tengfei, Wu Aimin, and Lang Congyan

Dept. of Computer Science & Technology, Beijing Jiaotong Univ., Beijing, China 100044
yangxubj@sohu.com

Abstract. This paper presents a motion suppression technique for video shot boundary detection based on 3D wavelet transform (3-DWT). Dramatic motion can be characterized in terms of energy and variance, and be differentiated from various video effects. Motion suppression value (MSV), which indicates intensity of motion energy, is extracted and used to suppress motion influence suffered by traditional video shot boundary detection algorithms. Unlike previous methods, our technique is robust to dramatic motion inherent in video sequence, experimental result validates our approach.

1 Introduction

Video shot boundary detection is a basic and important step for higher level video content analysis, such as video indexing and classification. To fulfill the task of video shot boundary detection, we need to detect and locate the transition of two adjacent shots in video stream. Generally there are two types of transitions, respectively abrupt transition (CUT) and gradual transition (GT). In this paper, we only refer to the most commonly used GT: fade in/out, and dissolve.

Early techniques for video shot boundary detection focus on CUT detection. A survey of existing method is presented in [1]. Color histogram-based method [2] is the most popular and robust one for CUT detection, however, it is usually combined with edge histogram-based method [3] to avoid the influence of illumination changes.

GT detection usually is more difficult because the change between consecutive frames is more gradual compared to a sharp CUT. Specifically, dramatic motion such as camera motion or large object motion in video sequence sometimes exhibits the same pattern with GT in metric curve, which makes GT detection a hard problem. Although motion-based methods were proposed, extraction of motion information suffers from several problems such as sensitive to video noise and illumination changes. Furthermore, single motion feature is too simple to represent content of video, it should be integrated with other methods to detect video shot boundary. Zhang et al. [4] proposed one of the most successful early methods, Twin-Comparison method, which attempts to handle both CUT and GT with two thresholds respectively for CUT and GT detection. However, in Twin-Comparison technique, it is important that very adaptive thresholds must be selected so that GT and motion are adequately discriminated. Suppression of motion influence is necessary for robust GT detection.

In this paper, dramatic motion is differentiated from various video effects including noise, illumination changes, GT, CUT. By characterizing dramatic motion under the

framework of 3-DWT, we extract motion intensity and define Motion Suppression Value (MSV) which is integrated into traditional histogram-based, edge-based methods for video shot boundary detection. Our algorithm has overcome the tough problem of dramatic motion in video shot boundary detection. The rest of this paper is organized as follows. In section 2, motion characterization approach is discussed. Section 3 is the extraction of MSV and the motion suppression algorithm for video shot boundary detection. We present some experimental results in section 4. Finally, section 5 concludes this paper.

2 Motion Characterization Approach

Video sequences can be perceived as 3D data volumes that have strong correlation both in space and in time. Efficient separable 3D wavelet transform is naturally an important tool for motion intensity representation, because wavelet representation has good localization properties in both frequency and spatial domains. However, in the real situation for video shot boundary detection, dramatic motion must be characterized and differentiated from other video effects including video noise, CUT, GT and illumination changes.

2.1 3D Wavelet Transform and Denoising

A segment of video sequence which contains certain consecutive frames, can be decomposed into eight subbands along two spatial axes which are labeled as x and y, and one temporal dimension labeled as z. Let L^λ and H^λ, respectively represent the high and low pass directional filter along directions of λ axis, where $\lambda = \{x, y, z\}$. We have one low pass subband, namely $L^x L^y L^z$, and seven wavelet subbands. We only choose $L^x H^y H^z$, $H^x L^y H^z$, $H^x H^y H^z$ subbands for motion characterization. Mathematically, adding all three matrices of subband and considering the result as a matrix of $I \times J \times K$ elements, namely $M[i, j, k]$. With this wavelet analysis, only higher frequency spatio-temporal edge information contributes to matrix M, which makes it robust to illumination changes.

However, motion characterization needs eliminating video noise. The elimination is achieved by thresholding $M[i, j, k]$ mentioned above. Similar to [5], our threshold selection is implemented as followling

$$T = med\{|M[i, j, k]|\}\sqrt{2\ln(N\log_2^N)} \tag{1}$$

where $N = I \times J \times K$, all the absolute coefficients of $M[i, j, k]$ are contained in the set $\{|M[i, j, k]|\}$. $med\{\}$ picks up median value in the set $\{|M[i, j, k]|\}$.

Since GT occurs over multiple frames, and usually it will last for 20 or more frames. The change between consecutive frames of GT is more gradual compared to a

sharp CUT, which makes GT produce only lower wavelet coefficient and exhibits almost the same pattern as noise. Therefore, coefficients of GT can also be seriously attenuated by our noise thresholding technique. Thus, after eliminating noise and GT, $F[i,j,k]$ is suitable for defining features for motion characterization.

2.2 Features Definition

We define two features, namely segment energy and variance, for motion characterization. Segment energy is defined by

$$SE = \sum_i \sum_j \sum_k \left(F[i,j,k] \right)^2 \tag{2}$$

The variance which measures the temporal texture of video segment is calculated by

$$Variance = \sum_k \left(\sum_i \sum_j \left(F[i,j,k] \right)^2 \bigg/ SE - 1/K \right)^2 \tag{3}$$

Variance can measure energy distribution along time axis. When energy is almost distributed in one frames, which indicates that a CUT may happen in the video segment, variance can be very high.

2.3 Motion Characterization

Video effects include video noise, CUT, GT, dramatic motion. By characterizing various video effects in wavelet domain, with features of SE and $Variance$, dramatic motion and CUT can be differentiated from other video effects.

Noise in 3D spatio-temporal data usually can be modeled as additive Gaussian distribution, and almost equally full in 3D space. Therefore, after denoising, both SE and $Variance$ of noise data is small. As we mentioned above, GT editing effect can be seen as some kind of special noise and GT exhibits the same pattern of features with video noise. CUT contains a sharp change at certain frame in 3D spatio-temporal data, so almost all of energy of wavelet coefficients are distributed in that frame. This results in high SE and very high variance. For camera operations or fast movements of large object, they contribute big coefficients to either of three subbands coefficients mentioned above, at the position of 3D spatio-temporal edges. Their energy is high, however, since edges is distributed all over time axis, their variance is low.

3 Motion Suppression Algorithm

3.1 Objective of the Algorithm

The core of motion suppression technique is to properly assign a motion suppression value for certain video segment which is decomposed by one level 3-DWT. The value

is used to attenuate motion influence suffered by conventional color histogram-based [2] and edge histogram-based [3] method for video shot boundary detection. Specially, all the inter-frame comparison results obtained by traditional methods over that segment are divided by motion suppression value. On the one hand, the more dramatic motion intensity is, the bigger motion suppression value should be assigned; on the other hand, we should preserve CUT or GT while suppressing dramatic motion.

For each video segment decomposed by one level 3-DWT, we first determine whether the segment contains CUT or dramatic motion. If CUT, suppression value is set to be 1, which means there is no suppression in CUT detection. If dramatic motion, suppression value must be assigned according to the estimated motion intensity.

3.2 Motion Suppression Value Extraction

We assign motion suppression value (MSV) according to the $F[i, j, k]$ mentioned in section 2. MSV is defined by

$$MSV = Max\left(\frac{AME}{AMT}, 1\right)$$

(4)

where AME denotes average motion energy, defined by

$$AME = \sum_i \sum_j \sum_k \left(F[i, j, k]\right)^2 \Big/ I \times J \times K$$

(5)

AMT denotes average motion energy threshold which is a empirical value. AMT is easy to choose and stable for various video due to it is an average quantity. Since GT just like noise, has been eliminated by thresholding technique in section 2, GT produces very low AME. Thus we achieve the goal of preserving GT.

4   Experimental Result

In our experiment, each video segment decomposed by 3D wavelet transform contains 20 frames. Harr wavelet is adopted in denoising signals since it preserves certain high-frequency signals (such as edges). The threshold of $Variance$ for CUT discrimination is 0.30. And the parameter AMT which should be tuned is set to be 2.3e-005, it's stable to various video sequence due to it is an average quantity.

Figure. 1 shows the effectiveness of our motion suppression technique. The tested video is a MTV named "Britney Spears" which contains various video effects including dramatic motions, CUT, and GT. After integrating motion suppression technique into two traditional methods, dramatic motion influence can be well suppressed.

For detecting various types of video shot changes, the proposed motion suppression technique is tested on five digital videos with about 12000 frames (about 8 minutes). The contents of the experimental data are MTV, martial arts movies and sports. There are quite a lot dramatic motions in each video. By integrating color-histogram and edge-histogram as well as motion suppression technique, final results are ob-

tained and shown in Table 1. We can see our motion suppression technique is robust to dramatic motion influence. In term of precision and recall, GT detection is fairly stable although there are quite a lot dramatic motions.

(a) Color histogram-based method.   (b) Edge histogram-based method.

(c) Motion suppression technique for (a).   (d) Motion suppression technique for (b).
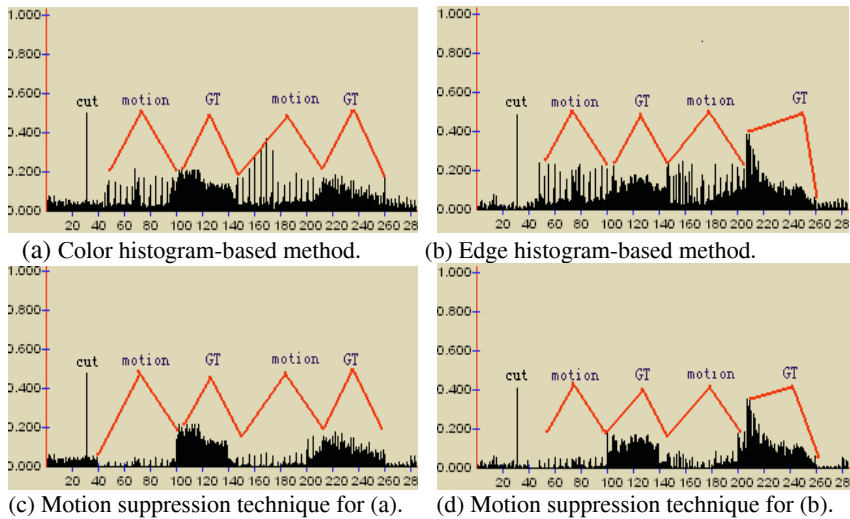
Fig. 1. Comparative results between traditional shot boundary detection algorithms and our motion suppression technique.

Table 1. Experimental result for five video sequences which includes dramatic motions

Video sequence	Description	Frames	Cut	GT	Motion	Precision		Recall	
						Cut	GT	Cut	GT
Britney Spears (MTV)	7 cut, 4 gradual	2000	7	5	4	98%	86%	100%	80%
Knight Spirit (MTV)	10 cuts, 1 gradual	1200	11	2	3				
The Matrix Reloaded	45 cuts, 4 gradual	5900	46	3	5				
JordanVSCarter (Sports)	7 cuts, 2 gradual	1500	7	1	4				
Basketball Block (Sports)	15 cuts, 4 gradual	1600	15	3	7				

5   Conclusion

In this paper, we characterize and differentiate dramatic motion from various video effects based on 3D wavelet transform. Motion suppression value (MSV) is robustly extracted and integrated into traditional histogram-based, edge-based methods for video shot boundary detection. Experimental results show that our motion suppression technique has suppressed the influence of dramatic motion in video shot boundary detection. Moreover, our technique can also be integrated into other shot boundary detection methods to improve their performance.

References

1. Hanjalic, A.: Shot-boundary detection: unraveled and resolved?. IEEE transaction on Circuits and Systems for Video Technology, Vol. 12, No. 2 (2002) 90-105
2. Nagasaka, A., Tanaka, Y.: Automatic video indexing and full-video search fro object appearances. In Visual Database Systems II, Netherlands: North-Holland (1992) 113-127
3. Zabih, R., Miller J., Mai, K.: Feature-based algorithm for detecting and classifying scene breaks. In Proc. ACM Multimedia, San Francisco (1995) 189-200
4. Zhang, H.J., Kankanhalli, A., Smoliar, S.W.: Automatic partitioning of full-motion video. Multimedia Systems Journal, Vol. 1, No. 1 (1993) 10-28
5. Chen, Z., Ning, R.: Breast volume denoising and noise characterization by 3D wavelet transform. Computerized Medical Imaging and Graphics, Vol. 28 (2004) 235-246

Neural Network Based Image Retrieval with Multiple Instance Leaning Techniques

S.C. Chuang, Y.Y. Xu, and Hsin-Chia Fu⋆

Department of Computer Science and Information Engineering,
National Chiao-Tung University
Hsinchu 300, Taiwan
{scchuang,yyxu,hcfu}@csie.nctu.edu.tw

Abstract. In this paper, we propose a *Generalized Probabilistic decision based Neural Network* (GPDNN) for content-based image retrieval (CBIR). Instead of receiving the numerical values of each data points as the input, the proposed GPDNN models the I/O relationship via the distribution of input data and their corresponding outputs. The GPDNN involves the Multiple-Instance learning techniques to learn a desired concept. A set of exemplar images are selected by a user, each of which is labeled as conceptual related (positive) or conceptual unrelated (negative) image. Then, by using the proposed learning algorithm, an image classification system can learn the user's preferred image class from the positive and negative examples. The experimental results show that for only a few times of relearning, a user can use the prototype system to retrieve favor images from the database.

1 Introduction

Content-based image retrieval (CBIR) system has been build over the last few decades,and is a interesting problem due to abundant contents and information hiding in images. How to extract proper features is still an open issue for the image retrieval problem.

Features used to represent the contents of the image can be roughly divided into two types: (1) global features and (2) local features. Global features are the features extracted from a whole image, and local features are the features extracted from local regions (subimages) of the image. Since the global features lack spatial information among the interested regions in the image, the retrieval results are not satisfactory. Hence, local features are adopted to represent the interested image more appropriate. Some of the systems, i.e., SaFe [1], uses the local features and spatial relationship among the interested subimages in the image as a new kind of features. The challenge of the local features extraction is how to automatically choose interested subimages from a sample image. Even for human, different interested subimages are selected from a single image due to their different point of views.

However, segmenting and deciding the interested subimages from an image automatically and precisely are very difficult tasks. In order to automatically extract preferred subimages of images without image segmentation, the image retrieval problem is reduced to the *multiple-instance learning problem*. The label is either a positive or negative, which is given to an image (bag) in multiple-instance based image retrieval problem. A positive bag contains a set of instances (subimages) if at least one instance in the bag belongs to the user's favored image class. A negative bag contains a set of instances if all instances in the bag do not belong to the user's favored image class. Since the labeling of an instance to be neither precisely nor completely, the conventional supervised learning framework is not suitable to learn these problems.

Recently, a few multiple-instance learning framework were proposed. Dietterich et al.,[2] proposed an algorithm for learning axis-parallel concepts in the drug discovery problems, but the proposed algorithm can only suitable to approximate the distributions that are as axis-parallel rectangles. O. Maron and A. Lakshmi Ratan [3] proposed the Quasi-Newton based framework to maximize various Diverse Densities to learn multiple-instance problems, where every instance in positive bags is used as an initial point and the learning speed of the framework is slow down.

In this paper, the image retrieval problem is reduced to the multiple-instance problem, and a Generalized Probability Decision Neural Network (GPDNN) is proposed to model the particular image concepts. GPDNN uses the data distributions instead data values as network input, which can approximate feature distributions precisely since image features in the form of distributions can be a better representation than in the numerical forms.

The paper is organized as follows. Section 2 presents the proposed GPDNN. Then, the learning methods are described in Section 3. The implementation and some experimental results are presented in Section 4. Concluding remarks are presented in Section 5.

2   Generalized Probabilistic Decision-Based Neural Network

The proposed GPDNN is a generalized model of its predecessor, SPDNN [4]. One major difference between the SPDNN and GPDNN is that GPDNN uses the data distributions instead data values as network input. it is beneficial that, for some pattern recognition problems, image features in the form of distributions can be a better representation than in the numerical forms. The schematic of a GPDNN is depicted in Fig. 1. Similar to its predecessor SPDNN, a GPDNN has a modular network structure. One subnet is designated to model one object class. A detailed description of GPDNN model is given in the following sections.

2.1   Discriminant Functions of GPDNN

The discriminate function of the multi-class GPDNN measures the difference between the input distribution and the distributions which are modelled or learned
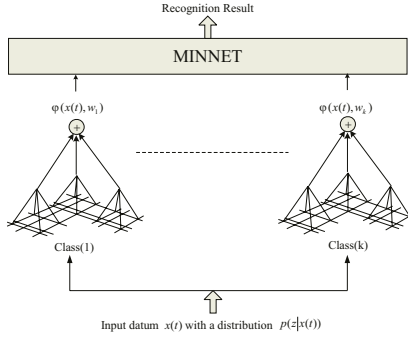
Fig. 1. The schematic diagram of a *k-class* GPDNN. Each subnet is designed to represent a class. (The detail of a subnet is shown in Fig. 2.) Input to a GPDNN is a data distribution instead of a numerical values of the data.

in its subnets. Let the likelihood function for a class ω_i be $p(\mathbf{z} \mid \omega_i)$. For a testing data $\mathbf{x}(t)$ with a distribution $p(\mathbf{z} \mid \omega(t))$, the difference of the distribution between $p(\mathbf{z} \mid \omega_i)$ and $p(\mathbf{z} \mid \mathbf{x}(t))$ can be defined as $\mathcal{P}_{i-t} \equiv p(\mathbf{z} \mid \omega_i) - p(\mathbf{z} \mid \mathbf{x}(t))$. In order to evaluate or measure the similarity between $p(\mathbf{z} \mid \omega_i)$ and $p(\mathbf{z} \mid \mathbf{x}(t))$, the discriminate function of a multi-class GPDNN is defined as

$$\varphi(\mathbf{x}(t), \mathbf{w}_i) = \int_{R^D} \mathcal{P}_{i-t}^2 d\mathbf{z}, \tag{1}$$

where $\mathbf{w}_i$ is the parameter set of the subnet i. Before exploring the expression of (1) in detail, we first introduce the product moment $\mathcal{F}$ of two distributions $\mathcal{P}_a$ and $\mathcal{P}_b$: $\mathcal{F}(\mathcal{P}_a, \mathcal{P}_b) = \int_{R^D} \mathcal{P}_a \mathcal{P}_b d\mathbf{z}$. Suppose $\mathcal{P}_a$ and $\mathcal{P}_b$ are two mixture Gaussian distributions: $\mathcal{P}_a = \sum_{r_a=1}^{R_a} P(\theta_{r_a}) p(\mathbf{z} \mid \theta_{r_a})$ and $\mathcal{P}_b = \sum_{r_b=1}^{R_b} P(\theta_{r_b}) p(\mathbf{z} \mid \theta_{r_b})$, where $p(\mathbf{z} \mid \theta_{r_a})$ and $p(\mathbf{z} \mid \theta_{r_b})$ are the Gaussian clusters in $\mathcal{P}_a$ and $\mathcal{P}_b$ with the parameter sets $\theta_{r_a} = \{\mu_{r_a}, \Sigma_{r_a}\}$ and $\theta_{r_b} = \{\mu_{r_b}, \Sigma_{r_b}\}$, respectively. $\mu_{r_a} = [\mu_{r_a(1)} \cdots \mu_{r_a(D)}]^T$ and $\mu_{r_b} = [\mu_{r_b(1)} \cdots \mu_{r_b(D)}]^T$ are the mean vectors, and $\Sigma_{r_a} = diag[\sigma_{r_a(1)}^2 \cdots \sigma_{r_a(D)}^2]^T$ and $\Sigma_{r_b} = diag[\sigma_{r_b(1)}^2 \cdots \sigma_{r_b(D)}^2]^T$ are the diagonal matrixes. Then, the product moment $\mathcal{F}$ can be expressed as:

$$\mathcal{F}(\mathcal{P}_a, \mathcal{P}_b) = \sum_{r_a=1}^{R_a} \sum_{r_b=1}^{R_b} P(\theta_{r_a}) P(\theta_{r_b}) \mathcal{G}(\theta_{r_a}, \theta_{r_b}),$$

where

$$\mathcal{G}(\theta_{r_a}, \theta_{r_b}) = \frac{\exp\left\{-\frac{1}{2} \sum_{d=1}^{D} \frac{(\mu_{r_b(d)} - \mu_{r_a(d)})^2}{\sigma_{r_b(d)}^2 + \sigma_{r_a(d)}^2}\right\}}{\sqrt{(2\pi)^D \prod_{d=1}^{D} (\sigma_{r_b(d)}^2 + \sigma_{r_a(d)}^2)}}, \tag{2}$$

and (1) can be expressed as:

$$\varphi(\mathbf{x}(t), \mathbf{w}_i) = \mathcal{F}(\mathcal{P}_i, \mathcal{P}_i) - 2\mathcal{F}(\mathcal{P}_i, \mathcal{P}_t) + \mathcal{F}(\mathcal{P}_t, \mathcal{P}_t). \tag{3}$$

The discriminate function φ can be enumerated by the expression (3). The following section describes how to implement a neural network for the calculation of (3).

2.2 Architecture

According to expression in (3), the discriminate function $\varphi(\mathbf{x}(t), \mathbf{w}_i)$ can be implemented by a two-layer pyramid network as shown in Fig. 2. The bottom layer contains three structurally identical pyramid subnetworks, each of which computes the $\mathcal{F}(\mathcal{P}_i, \mathcal{P}_i)$, $\mathcal{F}(\mathcal{P}_i, \mathcal{P}_t)$, and $\mathcal{F}(\mathcal{P}_t, \mathcal{P}_t)$, respectively. By summing the results of these three components, the difference from the testing data set $\mathbf{x}(t)$ to the class ω_i is output at the top layer of the pyramid subnetwork.
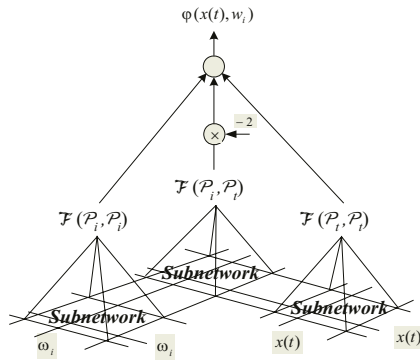
Fig. 2. The diagram of the subnet in *Generalized Probabilistic decision based Neural Network* (GPDNN).

Fig.3 depicts the internal architecture of the pyramid subnetwork corresponding to $\mathcal{F}(\mathcal{P}_i, \mathcal{P}_t)$. Suppose that the mixture Gaussian distributions $\mathcal{P}_i$ and $\mathcal{P}_j$ consist of R_i and R_t Gaussian clusters, then the subnetwork for $\mathcal{F}(\mathcal{P}_i, \mathcal{P}_t)$ contains a $R_i \times R_t$ input grids, each of which is marked as G_{r_i, r_t} node, R_i hidden nodes, and one output node.

While the subnetwork receives a testing data $\mathbf{x}(t)$ with a distribution $\mathcal{P}_t$, each G_{r_i, r_t} node performs the computation of (2) to measure the difference between the Gaussian clusters r_i in $\mathcal{P}_i$ and the Gaussian clusters r_t in $\mathcal{P}_t$. Then, the outputs of the G_{r_i, r_t} nodes are all weighted by $P(\theta_{r_t})$ and are summed to the hidden node h_{r_i}: $h_{r_i} = \sum_{r_t=1}^{R_t} P(\theta_{r_t}) G_{r_i, r_t}$. Finally, the output of each hidden node h_{r_i}, $r_i = 1, \ldots, R_i$, are all weighted by $P(\theta_{r_i})$ and are summed to the output node of the pyramid subnetwork: $\mathcal{F}(\mathcal{P}_i \mid \mathcal{P}_t) = \sum_{r_i=1}^{R_i} P(\theta_{r_i}) h_{r_i}$.

After the outputs of all the subnets are obtained, the MINNET in GPDNN is activated to select the minimum of the values from the lower subnet and its corresponding subnet I.D. That is, if the output value of subnet i is the minimum among the outputs of all subnets in a GPDNN, the testing data $\mathbf{x}(t)$ is classified as ω_i.
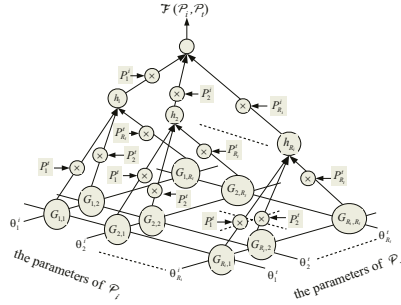
Fig. 3. The internal architecture of a model in Fig. 2 for computing of $\mathcal{F}(\mathcal{P}_i, \mathcal{P}_t)$.

3   Learning Rules for GPDNN

The GPDNN adopts the Multiple-Instance learning techniques to learn a desired concept. Every training bag is labelled as positive if at least one instance in the bag belongs to the user's favored image class, and is as negative if all instances in the bag do not belong to the user's favored image class. The label information and the following rules are used to modify (learn) the distribution parameters $\mathbf{w}_i$, such that the difference between training patterns and the discriminant function of a GPDNN becomes smaller in each iteration.

Reinforced Learning rule:
$$\mathbf{w}_i^{(m+1)} = \mathbf{w}_i^{(m)} + \eta \nabla \varphi(\mathbf{x}(t), \mathbf{w}_i) \tag{4}$$
Antireinforced Learning rule:
$$\mathbf{w}_j^{(m+1)} = \mathbf{w}_j^{(m)} - \eta \nabla \varphi(\mathbf{x}(t), \mathbf{w}_j) \tag{5}$$

The gradient vectors in (4) and (5) are computed as follows:

$$\frac{\partial \varphi(\mathbf{x}(t), \mathbf{w}_i)}{\partial \mu_{r_i(d)}} = 2P(\theta_{r_i}) \left[\sum_{r_n=1}^{R_i} \left(\frac{P(\theta_{r_n})\mathcal{G}(\theta_{r_n}, \theta_{r_i})}{\sigma_{r_n(d)}^2 + \sigma_{r_i(d)}^2} \right) \cdot (\mu_{r_n(d)} - \mu_{r_i(d)}) \right.$$
$$\left. - \sum_{r_t=1}^{R_t} \left(\frac{P(\theta_{r_t})\mathcal{G}(\theta_{r_t}, \theta_{r_i})}{\sigma_{r_t(d)}^2 + \sigma_{r_i(d)}^2} \right) \cdot (\mu_{r_t(d)} - \mu_{r_i(d)}) \right], \tag{6}$$

$$\frac{\partial \varphi(\mathbf{x}(t), \mathbf{w}_i)}{\partial \sigma_{r_i(d)}^2} = P(\theta_{r_i}) \left[\sum_{r_n=1}^{R_i} \left(\frac{P(\theta_{r_n})\mathcal{G}(\theta_{r_n}, \theta_{r_i})}{\sigma_{r_n(d)}^2 + \sigma_{r_i(d)}^2} \right) \cdot \left(\frac{(\mu_{r_n(d)} - \mu_{r_i(d)})^2}{\sigma_{r_n(d)}^2 + \sigma_{r_i(d)}^2} - 1 \right) \right.$$
$$\left. - \sum_{r_t=1}^{R_t} \left(\frac{P(\theta_{r_t})\mathcal{G}(\theta_{r_t}, \theta_{r_i})}{\sigma_{r_t(d)}^2 + \sigma_{r_i(d)}^2} \right) \cdot \left(\frac{(\mu_{r_t(d)} - \mu_{r_i(d)})^2}{\sigma_{r_t(d)}^2 + \sigma_{r_i(d)}^2} - 1 \right) \right], \tag{7}$$

where $r_i = 1, 2, \ldots, R_i$ and $d = 1, 2, \ldots, D$. D is the dimension of the feature space. For each iteration, the prior probability is updated as follows:

$$P_{new}(\theta_{r_j}) = \frac{P_{old}(\theta_{r_j}) \cdot \sum_{t=1}^{N} \left[\sum_{r_i=1}^{R_i} P(\theta_{r_i}) \cdot \mathcal{G}(\theta_{r_j}, \theta_{r_i}) - \sum_{r_t=1}^{R_t} P(\theta_{r_t}) \cdot \mathcal{G}(\theta_{r_j}, \theta_{r_t}) \right]}{\sum_{t=1}^{N} [F_{ii} - F_{it}]} \quad (8)$$

where $r_i = 1, 2, \ldots, R_i$ is the index of the cluster in the class model, $r_t = 1, 2, \ldots, R_t$ is the index of the cluster in the data model and $t = 1, 2, \ldots, N$. N is the number of the data sample.

4   Experiments

In training phase,we designed a bag generator to get the positive bags and negative bags as follows: First,we selected ten classes as same as the [5] described, and each class contains 100 images(bags). Then, we using the 100 images/bags as initial positive bags set,and randomly selected one image/bag from the other 9 classes as initial negative bags,respectively.

When a desired image class is to be trained, the positive and the negative exemplar images sets are first generated by our bag generator as described above. Then, the system learns the desired image class using the proposed Neural Network based Multiple-Instance learning algorithm. When the optimal model for desired class is trained,each Gaussian mixture gets its own mean ,variance and prior to approximate the model.

In this prototype system, we have trained ten classes : "Africa", "Beach", "Building", "Buses","Dinosaurs", "Elephants", "Flowers", "Horses" and "Mountains". The results of our experiment results are shown in Table 1. The results indicate that using the Multiple Instance Learning method and refined the retrieval process by the user relevance feedback can effectively retrieve the user's conceptual related images.

Table 1. Performance of retrieving results from (1) query by ($GPDNN$), (2) IRM, on the different categories of images. The precision performance of $GPDNN$ was presented by its precision rate.

category	Africa	Beach	Building	Buses	Dinosaurs	Elephants
IRM	0.475	0.325	0.33	0.36	0.981	0.4
GPDNN	0.58	0.28	0.55	0.49	0.99	0.41

category	Flowers	Horses	Mountains	Food	Average
IRM	0.402	0.719	0.342	0.34	0.468
GPDNN	0.63	0.84	0.47	0.58	0.494

5   Conclusions

In this paper, we propose a *Generalized Probabilistic decision based Neural Network* (GPDNN) for content-based image retrieval. The GPDNN is trained to

learn a desired concept via the Multiple-Instance learning techniques. The result indicates that using the Multiple Instance Learning method and refined the retrieval process by the user relevance feedback can effectively retrieve the user's conceptual related images.

References

1. Smith, J.R., Chang, S.F.: Safe: A general framework for integrated spatial and feature image search. In: IEEE 1997 Workshop on Multimedia Signal Processing. (1997)
2. Dietterich, T.G., Lathrop, R.H., Lozano-Perez, T.: Solving the multiple-instance problem with axis-parallel rectangles. Artificial Intelligence (1997) 89(1–2): 31–71
3. Maron, O., Ratan, A.L.: Multiple-instance learning for natural scene classification. In: Proceedings of the 15th International Conference on Machine Learning. (1998) 341–349
4. Fu, H.C., Xu, Y.Y.: Multilinguistic handwritten character recognition by bayesian decision-based neural networks. IEEE Transactions on Signal Processing **46** (1998) 2781–2789
5. Chen, Y., Wang, J.: A region-based fuzzy feature matching approach to content-based image retrieval. IEEE Transactions on Pattern analysis and machine intelligent **24** (2002) 1252–1267

Shot Type Classification in Sports Video Using Fuzzy Information Granular

Congyan Lang, De Xu, Wengang Cheng, and Yiwei Jiang

Department of Computer Science,
Beijing Jiaotong University, Beijing 100044, P.R. China
`gltree@263.net`

Abstract. In this paper, we present a new method for classifying shot type in sports video using fuzzy information granular. The problem is important for applications such as video structure analysis and content understanding. In particular, two-stage off-line learning processes perform knowledge extraction of semantic concepts and automatic shot classification, respectively. In the first stage, the extracted prominent regions are used as a good pattern in semantic concept level. Then a number of global features are defined as efficient input of the shot type classifier in the second stage. The identification of semantic concepts and classification of shot are based on soft decisions. Hence, this framework can adequately capture the uncertainty or ambiguity of scales of a shot. Experimental results show the excellent performance of the approach.

1 Introduction

Lately, there is an emerging need for efficient media management including browsing, filtering, indexing and retrieval with the current advance of video database technique. Specially, sports video has received increasing interest in recent years due to its tremendous commercial potentials and well-defined syntactic and semantic structures. Many approaches in the literature of sports video analysis and classification have been developed for soccer and various sports games[1]. In this task, shot type classification is an important and fundamental work for video structure analysis and content understanding.

For sports video analysis, most existing approaches focus on the integration and evaluation of domain specific knowledge and pattern recognition techniques. A number of methods have been proposed in the literature to classification event of sports video, such as the work of Gong et al. [2] on view and event classification for soccer video indexing. The method depends on camera view type of the shot as fundamental characters, which illustrated importance of shot type classification for semantic structure analysis. Therefore, distinguishing the view types of shot is important and meaningful. Motivated by these observations, Peng Xu et al.[3]classified shots into long, medium and close-up classes by computing dominant color ratio. However, due to variations of lighting and field conditions, a high error rate occurs in classification for medium shot and long shot. Ekin, A. et al.[4] propose a fully automatic framework for analysis and summarization of soccer videos using cinematic and object-based features, where authors define regions using Golden Section spatial composition rule to classify shot into three types.

Although these systems have been successful in their respective applications, shot type classification for sports video is still an unsolved problem. One reason is that low-level features are often not enough to classify video shot without the incorporation of high-level and perceptual information into the classification process. Another reason for the problems is discrimination subjectivity for long-medium-close-up scales. Different people can differ in their perception of shot view scale.

Given a sports video, our primary goal in this paper is to classifying shot into main types at semantic concept level. The domain of interest here is soccer video. Compared to existing work, our system is distinctive with several important advantages: (1) To bridge semantic gap, we use fuzzy granular to represent the human concepts, and then semantic concepts are constructed as content pattern instead of the low-level features; (2) Two-stage off-line learning processes perform knowledge extraction of semantic concepts and automatic shot classification, respectively; (3) The identification of semantic concepts and classification of shot are based on soft decisions.

In cinematography, shots are generally classified into four main types in soccer video production according to the scale of the shot, namely Long shot (LS), In-field medium shot (MS), Close-up shot (CS) and Out-of-field shot (OS). Motivated by this observation, we classify sports video shots into four classes described in Figure1. Rather than processing every frame of the video shot, our technique analyses only key frames since there is typically little difference among consecutive frames.

Fig. 1. Four types of shot: long shot(LS), medium shot(MS), close-up shot(CS), out-of field shot(OS)

The rest of the paper is organized as follows: In Section 2, we present algorithm for prominent region extraction including feature granular and fuzzy granular classifier; In Section 3, global features definition and shot classification are presented. The effectiveness of the proposed approach is validated by experiments over real-word soccer video in Section 4. Concluding remarks are given in Section 5.

2   Prominent Region Detection for Shot Content Representation

2.1   Feature Extraction and Selection

Firstly, each sampled frame is initially segmented into homogeneous regions. Then in order to get suitable content pattern of region, a number of perceptual features denoted as {*CSr*, *OCr, SIr and Mir*} are extracted to represent *Contrast*, *Orientation Conspicuity*, *Shape Indicator* and *Motion Prominent*. More details can be consulted in our prior work[5].

After perceptual features are extracted, we map raw numeric feature space to fuzzy granule space in order to simulate human perception process. Information granulation is an important endeavor of information processing. The use of fuzzy granular to

partition universes has significant advantages over more traditional approaches, and fuzzy granular based classifier tends to be more robust and less sensitive to small changes in attribute values near partition boundaries.

Mapping the original data to the fuzzy granular domain is preformed by labeling each data with linguistic variables. A linguistic variable is a variable which takes fuzzy values and has a linguistic meaning. In this paper, we defined three linguistic variables *Low, Medium* and *High* for each feature X_i. The results can be represented by a 3-D structure, denoted as (X_i, F_j, MF) where F_j is a pre-defined linguistic variable and *MF* is mapping function. In this paper, hypertraperzoidal membership function is ideal for our system because it is a convenient mechanism for representing and dealing with multidimensional fuzzy sets. The definition of these functions is based on the clusters[6]. To illustrate this, consider a quantitative attribute X_i and has three corresponding fuzzy sets, three fuzzy sets $F_j = \{f_{jL}, f_{jM}, f_{jH}\}$ are exploited which has been validated to represent our raw data by our experiment. Given a value x_i representing an instance of a feature X_i in raw feature space, by using $MF_{F_j}(x_i)$ the membership function transforms the numerical value x_i into the fuzzy value   as follows:

$$MF_{F_j}(x_i) = \{\mu_{f_{jL}}(x_i), \mu_{f_{jM}}(x_i), \mu_{f_{jH}H}(x_i)\} \tag{1}$$

2.2   Prominent Region Extraction Based on Fuzzy Granular Classifier

The aim of the classifier is to find the regions of each frame to which a human observer is mostly attracted. The theory of mass assignments developed by Baldwin[7] provides a formal mapping between probability distributions and linguistic descriptions. Due to its capability of manipulating both probabilistic and fuzzy uncertainty, fuzzy granular classifier based on mass assignments is ideal for our problem. Here, we briefly present the concepts for mass assignments and procedures of classification, the interested reader may consult[7] for more details. The bi-directional transformation proposed by Baldwin provides possibility to construct a fuzzy granular classifier combining fuzzy granular and Bayes theorem.

Suppose we have a set of feature vector $X = (x_1, x_2, ..., x_n)$ along with a set of class label $C = \{C_1, C_2, ..., C_l\}$. For feature vector X_i, the fuzzy granular is denoted as $F_j = \{f_{j1}, f_{j2}..., f_{jm}\}$. In order to predict the class to which the variables belong we need use label mass assignments to estimate classification probabilities by calculating the probability $\Pr(C_k | F_1, ..., F_n)$. To simplify the model, we assume region features extracted are all conditionally independent, which is intuitively reasonable. According to Bayes theorem, the posterior probability is given by:

$$\Pr(C_k | F) = \frac{\Pr(C_k) \prod_{j=1}^{n} p(F_j | C_k)}{p(F)} \quad k = 1, ..., l \tag{2}$$

$P(F_j \mid C_k)$ can be estimated from the linguistic prototype for C_k. From this we can obtain an estimate for $\Pr(C_k \mid F)$.

For the prominent region detection problem each segmented region was described using a set of four dimension features, comprising of the perceptual features described before. Let $X = (x_1, x_2, x_3, x_4)$ $(n=4)$ denote the perceptual features *CSr*, *OCr, SIr and MIr* of region R_i. By mapping the features into fuzzy granular domain *(m=3)*, a number of fuzzy granular corresponding to each original feature can be obtained. Then prominent region classifier receives a total of $3n$-dimensional input data, and outputs the class label: *PR* and *NPR (l=2)*. After this stage each key frame can be represent as $kf = \{\mathrm{P}R_i, NPR_j\}$, which offer semantic concept representation in higher level. Most importantly, they transform diverse input data structures into a granular form which facilitates global feature extraction for a robust shot type classification.

3   Shot Type Classification

In this section we describe the process of classifying shots into predefined set of views. Semantically, four shot type differ in their shooting scale. Most existing approaches[3] for soccer shot classification use the ratio of green grass area as main feature. Due to variations of lighting and field conditions, medium shots with high pass colored ratio will be mislabeled as long shots. To overcome these limitations, for each shot we defined a number of global features which both computationally low and robust to variations of lighting and field conditions within the sequences.

1) Grass Area Ratio (GR):
By studying the results of prominent region map, we find that for the strong illuminated frames, there is a distinct prominent region map with other frames.

Fig. 2. Prominent region detection for two illuminated frames

Form the fig.2, we can see that rare prominent region can be detected for the frames corresponding to the variations of lighting. Hence, we improve the measure of grass color computation by discarding the frames where the number of prominent region N_{PR_i} is low ($N_{PR_i} < T_n$). After detection of prominent region, the histogram of the hue component of non-prominent region is added up over these frames. The grass color is represented as the peak of this cumulative hue histogram.

*2) **Average Relative Size of Prominent Region(AP_{size}):***

For a prominent region of key frame, relative prominent size PS is defined as fol-lows:

$$PS_i = N_{pixel}(PR_i) \bullet e^{-\alpha D(PR_i)} \tag{3}$$

where $N_{pixel}(PR_i)$ denotes the number of pixels in the prominent region PR_i , $D(PR_i)$ presents the distance between the location of the region and the center of corresponding frame. The parameter α balances the two concerns in the prominent size computation and is chosen so as to control an extent to which the location of region impacts the relative prominent size. Finally, average size of prominent region AP_{size} is obtained by computing average value of all the prominent regions whose size is smaller than pre-defined threshold T_s .

*3) **Average Shape of Prominent Region(AP_{shape}):***

The shape of prominent region is defined as:

$$PD_i = \frac{N_{edge}(PR_i)}{N_{pixel}(PR_i)} \tag{4}$$

where $N_{edge}(PR_i)$ is number of edge pixels in the region. Then, we compute aver-age value of all the prominent region shape as the third global feature AP_{shape} .

Then for classification view types of shot, the global features are used as input of shot classifier to classify shot which relies on the second off-line learning process based on fuzzy granular classifier proposed in section 2.2. After the Global features are granulated, feature granular is fed to the system inputs, four class labels (LS, MS, CS and OS) are output (*l=4*) of shot type classifier which can extract and use rules automatically with superior performance and provide a richer and more accurate classification than crisp classifiers.

4   Experimental Results

In this section, we describe the experiments aimed at evaluating the proposed ap-proach, which was tested using several soccer video sequences. The parameters are determined by experimental tests, more reasonable results of our method could be achieved by setting threshold T_n =0.15, T_s =0.3 and balance parameter α =0.5 in global feature extraction phase. In order to perform training, three video sequences including 1769 frames are randomly selected. And then 181 key frames are extracted from the training data. For each key frame, we label the ground truth manually. Fig.3 shows the prominent region detection results of four key frames taken from different type of shot.

Fig. 3. Prominent region detection in soccer video shots
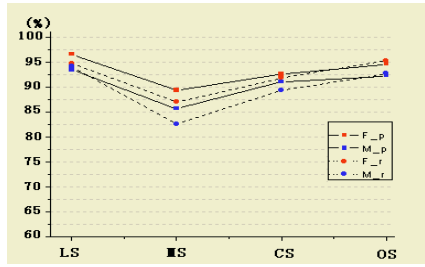
Fig. 4. PR Graph for two methods

Specifically, Fig.4 shows the original frames, corresponding results of homogenous region segmentation, and prominent region detection results from top to bottom row. Prominent regions are located and used to obtain a binary image that contains white and black pixels to represent prominent region and non-prominent region, respectively.

Table 1. Performance of the proposed shot type classification for two soccer video sequences

Types	Truth	Miss	False	Precision (%)	Recall (%)
LS	1872	21	18	99.04	98.89
MS	699	39	32	95.62	94.71
CS	576	3	46	92.60	99.31
OS	245	2	11	95.70	99.19

Soccer video sequences including 226 shots are used in the experiments to train shot type classifier. Table.1 shows performance of the classification for one soccer video sequence taken from Test Data. From the experiment results, we can see that our approach achieves very high precision and recall in classifying long shot (LS) and Out-of fields shot (OS). By comparison, precision for Medium shot (CS) is relative lower due to the errors created in prominent region detection. For instance, in some frames there are many players stand together. Therefore, corresponding regions would be mistakenly detected as one prominent region, which lead to classifier mistakenly medium shots as close-up shots. However, we expect this drawback can be improved by using temporal context analysis, which is one of our future works.

We also compared our results with the results of the crisp classification algorithm (we refer to this methods as "*CCA* method" throughout the rest of this section) which employed naïve bayes classifier using grass-color-ratio described in [3] considering the assumption of feature independence. A sports video clip, which has strong variations of lighting, is selected to perform comparison. Fig.6 shows the comparison results using PR graph for two methods, where the X-axis denotes types of shot and Y-axis denotes the precision and recall. Precision and recall of the proposed approach and *CCA* method is denoted as F_p ,F_r, M_p and M_r , respectively. As we can see, the results given in fig.4 show a significant increase in classification especially for medium shot.

5   Conclusions

A fuzzy granular based approach for classifying view types of shot is presented in this paper. Instead of low-level feature, prominent regions are determined as semantic concepts to represent each key frame, which facilitates global feature extraction for a robust shot type classification. Two-stage off-line learning processes perform knowledge extraction of semantic concepts and automatic shot classification, respectively. Moreover, Usage of fuzzy information granular has advantages over numeric form, because of its proximity to human understanding of multimedia information. One advantage of the proposed approach is its ability to robust classify shot view, this information can be used to perform a further analysis on the sports video sequences. Although the experimental results show the encouraging performance, the conducted research also shows that there is plenty of room for improvement. Future work will be focused on the utilization of temporal context and multi-modal information in the process of the event analysis, such as goal detection, high-level semantic event classification, etc.

References

1. Niall Rea, Rozenn Dahyot, Anil Kokaram, Modeling High Level Structure in Sports With Motion Driven HMMs: Proc. Of ICASSP (2004) 621-624
2. Y.Gong, L.T.Sin, C.H.Chuan, H.J.Zhang, M.Sakauchi: Automatic Parsing of TV Soccer Programs. Proc. Of IEEE ICME (1995) 165-174
3. Peng Xu, et al: Algorithms and System for Segmentation and Structure Analysis in Soccer Video.Proc. Of IEEE ICME (2001) 928-931
4. Ekin, A , Tekalp, A. M.: Automatic Soccer Video Analysis and Summarization. IEEE Trans. On Image Processing 12(7), (2003) 796-807
5. Lang Congyan, Xu De: Perception-Oriented Prominent Region Detection in Video Sequences Using Fuzzy Inference Neural Network. Lecture Notes in Computer Science 3497 / 2005 (2005) 819-828
6. W.Kelly, J.Painter, Hypertrazoidal Membership Functions. Proc. Of IEEE Conference On Fuzzy Systems. (1996)
7. J.F.Baldwin, J.Lawry, T.P.Martin: A Mass Assignment Method for Prototype Induction. International Journal of Intelligent Systems. 14(10), (1999) 1041-1070

Method for Searching Similar Images Using Quality Index Measurement

Chin-Chen Chang[1,2] and Tzu-Chuen Lu[2]

[1] Feng Chia University, Taichung, Taiwan, 40724, R.O.C.
`ccc@cs.ccu.edu.tw`
[2] National Chung Cheng University, Chiayi, Taiwan, 621, R.O.C.

Abstract. Searching for similar images is an important research topic for multimedia database management. This paper uses a quality index model to search for similar images from digital image databases. In order to speed up retrieval, the quality index model is partitioned into three factors: loss of correlation, luminance distortion, and contrast distortion. The method is performed on three different image databases to test for retrieval accuracy and category retrieval ability. The experimental results show that the proposed method performs better than the color histogram method, the color moment method, and the CDESSO method.

1   Introduction

In image processing, the most commonly used measurements for estimating the difference between two images are mean squared error (MSE), peak signal-to-noise ratio (PSNR), mean absolute error (MAE), and so on. These measurement methods are usually easy and have lower computation complexity. Nevertheless, according to Wang and Bovik's experiments, most of the methods cannot be accurately applied in strict testing conditions or in a different image distortion environment [10]. In addition, the methods require human vision to incorporate perceptual quality measurements. Therefore, Wang and Bovik proposed another image quality measurement - a universal image quality index (quality index) to estimate the difference between two images. The method was independent of images and human vision. According to their experiments, the scheme outperformed the MSE significantly. The concept of the quality index can be applied to various image processing applications. One application is used to search for similar images from an image database. Because of the rapid advance in information technology has enabled us to access a large number of images in an instant from every corner of the world. However, the huge number of images available, finding the right image from a large number of image databases is difficult. Hence, searching for images has become an important research issue [1, 4, 5, 10]. In the beginning, keywords were used to search for desired images. Each image in an image database annotated with keywords. However, annotation-based methods have become impractical and inefficient, since more and more images are generated from the Internet and most of them do not include keywords. In

addition, annotation is a subjective operation in that different people may use different keywords for the same image [8]. Therefore, some researchers have developed other similarity search methods to accurately access desired images. One popular method is content-based image retrieval (CBIR), which automatically extracts features, such as color, texture, shape, moment, distance, and so on, from an image [2, 6, 7].

In this paper, we adopt the concept of the quality index to represent the features of an image for searching similar images. In addition, in order to speed up retrieval in terms of comparing the similarity between two images, the quality index is partitioned into three different factors, loss of correlation, luminance distortion, and contrast distortion. The organization of the paper is as follows. Section 2 briefly presents Chan and Liu's CDESSO image retrieval method and Wang's quality index model. Section 3 describes the proposed method in more detail. Section 5 demonstrates the experimental results, and Section 6 provides the conclusions.

2   Related Works

In 2003, Chan and Liu proposed a CDESSO-based method to characterize color complexity and color differences among adjacent pixels for 24-bit full color image retrieval [1, 3]. First, they used the K-mean algorithm to reduce the color space of an image. They divided all pixels of database images into 64 clusters. Each cluster had its own bin to record the difference between two adjacent pixels. Second, each pixel in an image was fitted into the closest cluster. Then they scanned the reduced image in spiral order and computed the difference between any two neighboring pixels. The difference is then added to the corresponding bin of the current pixel. Chen and Liu used the final 64 bins, called a color histogram, as the features to represent an image. In terms of image retrieval, the histogram of a query image was compared to all the histograms of images in the image database using the Euclidean distance. According to their experimental results, their method not only provided a high accuracy rate for finding database images, but also resisted scale variants, such as shifting and rotation, of images. The CDESSO image retrieval method, as well as other color-based CBIR methods, uses the Euclidean distance to estimate the difference between two images. However, measurements such as MSE, PSNR, and Euclidean distance cannot be accurately applied in strict testing conditions or in a different image distortion environment. Therefore, Wang and Bovik proposed a mathematically defined universal image quality index to measure the quality of an image.

The definition of Wang and Bovik's quality index model is given below. Let $\alpha = \{\alpha_i, 1 \leq i \leq N \times N\}$ be an original image of size $N \times N$, where δ_i is the i-th pixel of α and $\beta = \{\beta_i, 1 \leq i \leq N \times N\}$ is a test image of size $N \times N$. The difference between α and β is measured by

$$Q = \frac{4 \times \sigma_{\alpha\beta} \times \mu_\alpha \times \mu_\beta}{(\sigma_\alpha^2 + \sigma_\beta^2) \times (\mu_\alpha^2 + \mu_\beta^2)}, \tag{1}$$

where μ_α is the mean of α that is given by $\mu_\alpha = \frac{\sum \alpha_i}{N \times N}$, μ_β is the mean of β that is given by $\mu_\beta = \frac{\sum \beta_i}{N \times N}$, σ_α^2 is the variance of α, σ_β^2 is the variance of β, $\sigma_{\alpha\beta}$ is the correction between α and β that is given by $\sigma_{\alpha\beta} = \frac{\sum(\alpha_i - \mu_\alpha) \times \sum(\beta_i - \mu_\beta)}{N \times (N-1)}$. The range of Q is from -1 to 1. If the value of Q is close to 1, then we can say that the test image is almost the same as the original one. If the value of Q is equal to -1, then the two images are absolutely different. In the paper, we adopt the concept of quality index to search for similar images from an image database.

3   The Proposed Method

Suppose A is a digital image of size $N \times N$ in an image database. Most digital images are represented in RGB color space, such that each pixel can be interpreted as a value in the 3-dimension, red, green, and blue, color space. However, the RGB color space is unsuitable and inconvenient for image analysis. Therefore, the first step of the proposed method is to transform a 3-dimension color space into a 1-dimension gray space. Let $\alpha = \{\alpha_i, 1 \leq i \leq N \times N\}$ be the transformed image of A. Let B be a query image. The first step in searching for images similar to B from an image database is to transform B into 1-dimension gray space image $\beta = \{\beta_i, 1 \leq i \leq N \times N\}$. The next step is to use Equation 1 to calculate the quality index between β and each image α in the database. The image with the largest quality index is the image most similar to the query image. Since the range of Q is from -1 to 1, when the value of Q between β and α is 1, then we can say that images A and B are the same image. In the quality index model, the value of Q is computed using means and variances of β and α, and the correlation between β and α. The means and variances can be preprocessed independently. Nevertheless, the correlation still needs to be dynamically computed in terms of images comparison. That may requires a lot of comparison. Therefore, we recombine the formula of the quality index in Equation 1. The new formula can filter out non-matching images in advance to speed up image retrieval.

According to the definition of the quality index model, the quality index can be recombined by three factors: loss of correlation, luminance distortion, and contrast distortion. The recombined equation of the quality index is given by

$$Q = \gamma \times \ell \times \partial, \tag{2}$$

where $\gamma = \frac{\sigma_{\alpha\beta}}{\sigma_\alpha \times \sigma_\beta}$ is the correlation coefficient, $\ell = \frac{2 \times \mu_\alpha \times \mu_\beta}{\mu_\alpha^2 + \mu_\beta^2}$ is the luminance distortion, $\partial = \frac{2 \times \sigma_\alpha \times \sigma_\beta}{\sigma_\alpha^2 + \sigma_\beta^2}$ is the contrast distortion. The range of γ is from -1 to 1; the range of ℓ is from 0 to 1; and the range of ∂ is from 0 to 1. We can see that only the value of γ is concerned with each pixel of α and β. The parameters $\mu_\alpha, \mu_\beta, \sigma_\alpha, \sigma_\beta$ can be preprocessed before image comparison. Let $\theta = \ell \times \partial$ be the product of ℓ and ∂. Hence, the quality index Q can be represented by $Q = \gamma \times \theta$. Since the maximum value of γ, any value of γ would decrease the value of Q except $\gamma = 1$. Hence, we can predefine a threshold T to filter out unlikely images

whose value of θ is lower than T. In other words, only the candidate image whose value of θ is greater than T needs to be compared. The features used to represent an image in an image database are mean, variance, and the set of central pixels, $C = \{\alpha_k, k \in Centralregion\}$, of the transformed image α. Since the most important objects in an image are usually located in the central region, we only record the transformed pixels in the central region of an image for further correlation coefficient analysis to save storage space.

4   Experimental Results

Two different experiments were carried out to test the performance of the proposed method. The first experiment examined the retrieval accuracy, and another examined the category query ability. Three existing programs, color histogram, color moment, and CDESSO, were used to benchmark the proposed method. All the methods were implemented on an Intel Pentium III 500 MHz PC with 256 MB Ram in Java language. The methods were tested with three image databases. The first image database contained two image sets, D_1 and Q_1. Each image set contained 410 full color animations in JPEG format with different sizes [3]. Every image in Q_1 had one corresponding image in D_1 in pairs. The second image database contained two image sets, D_2 and Q_2. Each image set contained 235 full color images with size 384×256 or 256×384 [3]. Every image in Q_2 had one corresponding image in D_2. The third image database D_3 contained 10,235 full color images, which were collected from [5, 7, 9]. The 235 images in D_2 were also embedded into D_3. Let $\|D_i\|$ and $\|Q_i\|$ be the total numbers of images in the image sets D_i and Q_i, respectively.

4.1   The Retrieval Accuracy Experiments

The first experiment tested the retrieval accuracy of the proposed method with D_1 and Q_1. The benchmark used in this experiment was the CDESSO method. We used each image in the sets Q_1 and Q_2 as the query image to retrieve N_R images from D_1 and D_2. The retrieved images were ranked according to the quality index values, which are given by Equation 4, in descending order. If the corresponding image of a query image is one of the N_R retrieved images, then we can say that the query image was accurately used to retrieve the target image. The measurement used to estimate the retrieval accuracy of a method using D_i is given by

$$R_i = \frac{C_i}{\|Q_i\|}, \tag{3}$$

where C_i is the total number of the query images in Q_i, which can be accurately used to retrieve the target image from D_i. Table 1 shows the retrieval accuracies of the CDESSO method and the proposed method using D_1 and D_2. For $N_R = 1$, the retrieval accuracies of the proposed method were 95.34% and 55.13% using Q_1 and Q_2, respectively, while the retrieval accuracies of the CDESSO method were 93.72% and 40% using Q_1 and Q_2, respectively. According to the results, as

Table 1. The retrieval accuracies from the first experiment.

	Q_1		Q_2	
N_R	CDESSO	Our's	CDESSO	Our's
$N_R = 1$	93.72	95.34	40.00	55.13
$N_R = 2$	97.09	97.67	47.23	63.68
$N_R = 3$	97.09	98.19	51.49	67.52
$N_R = 4$	97.76	98.45	52.77	68.80
$N_R = 5$	97.98	98.70	54.47	68.80
$N_R = 10$	99.33	99.48	60.43	71.37
$N_R = 20$	99.55	99.74	70.21	77.78

shown in Table 1, the proposed method would have had better retrieval accuracy than the CDESSO method using both Q_1 and Q_2.

The second experiment examined the retrieval accuracy of the proposed method with a large number of images from the image set D_3. In this example, we also used Q_2 as the query image set and retrieved N_R images from D_3. The benchmarks used in this experiment were the color histogram and the color moment methods. The experimental results are shown in Table 2.

Table 2. The retrieval accuracies from the second experiment.

Methods	$N_R \leq 10$	$N_R \leq 100$	$N_R \leq 500$	$N_R \leq 1000$	$N_R > 1000$
Color histogram	0.00	3.03	8.08	12.12	87.88
Color moment	46.46	70.71	85.86	88.89	11.11
Our method	61.80	74.68	89.27	94.85	5.58

4.2   The Category Retrieval Precision

The third experiment evaluated the categorizing ability of the proposed method. In this experiment, we formed the subset images of D_3 using ten image categories (architecture, city, dog, eagle, elephant, leopard, model, mountain, pyramid, and royal), each containing 100 images. Every image in these ten categories was used as the query images, and the corresponding images for each query image were the other images belonging to the same category. The measurement used to evaluate the precision of the category of a method is given by: $P = \frac{N_{Cr}}{N_R}$ where N_{Cr} is the number of N_R retrieved images that belong to the same category as the query image. The N_R set in this experiment was 30. Table 3 shows the experimental results of the proposed method, the color histogram, and the color moment. For the category Eagle, the number of target images found by the proposed method was 23, while no target image was found by the color histogram, and only 3 target images were found by the color moment. The corresponding precisions were 63%, 0%, and 10%, respectively. Fig. 1 shows the aggregated results over all ten categories. According to the results, we can see that our proposed scheme is indeed better than others in most cases in terms of the categorizing ability.

Table 3. The experimental results for the ten categories for $N_R = 30$.

Category	Color histogram		Color moment		Our method	
	N_{Cr}	P (%)	N_{Cr}	P (%)	N_{Cr}	P (%)
Architecture	1	3	2	6	7	22
City	0	0	2	6	20	66
Dog	1	3	1	3	13	43
Eagle	0	0	3	10	23	63
Elephant	0	0	2	6	21	69
Leopard	0	0	2	6	11	35
Model	1	3	1	3	12	41
Mountain	0	0	2	6	12	39
Pyramid	0	0	2	6	10	33
Royal	1	3	2	6	13	45
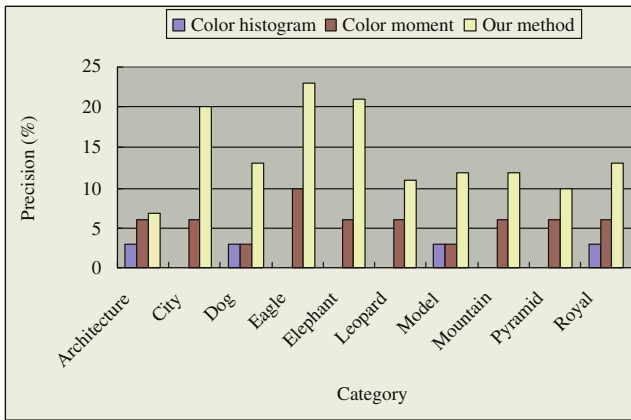
Fig. 1. Aggregated results over all ten categories.

5 Conclusions

In this paper, we proposed a method for searching for similar images based on the quality index measurement. In order to speed up image retrieval, we recombined the formula of the quality index for filtering impossible images. The proposed method was tested under different conditions to determine performance. The experiments were performed on three different databases. The results showed that the proposed method performs better than the color histogram, the color moment, and the CDESSO method, especially for retrieval of images in the city, eagle, elephant, leopard, mountain, pyramid, and royal categories. The total memory space for saving the image features of the proposed method is less than other methods.

References

1. Chang, C. C. and Chan, Y. K.: A Fast Filter for Image Retrieval Based on Color-Spatial Features. Proceedings of the Second International Workshop on Software Engineering and Multimedia Applications. **2**, Baden-Baden, Germany. (2000) 47–51

2. Chan, Y. K. and Chen, C. Y.: Image Retrieval System Based on Color-Complexity and Color-Spatial Features. Journal of Systems and Software. **71** (2004) 65–70

3. Chan, Y. K. and Liu, Y. T.: An Image Retrieval System Based on the Image Feature of Color Differences on Edges in Spiral Scan Order. International Journal on Pattern Recognition and Artificial Intelligence. **17** (2003) 1417–1429

4. Fuh, C. S., Cho, S. W. and Essig, K.: Hierarchical Color Image Region Segmentation for Content-Based Image Retrieval System. IEEE Transactions on Image Processing. **9** (2000) 156–162

5. Iqbal, Q. and Aggarwal, J. K.: CIRES- A System for Content-based Retrieval in Digital Image Libraries. Proceedings of the Seventh International Conference on Control, Automation, Robotics and Vision, Singapore. (2002) 205–210

6. Kokare, M., Chatterji, B. N. and Biswas, P. K.: Cosine-Modulated Wavelet Based Texture Features for Content-Based Image Retrieval. Pattern Recognition Letters. **25** (2004) 391–398

7. Li, J. and Wang, J. Z.: Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence. **25** (2003) 1075–1088

8. Schettini, R., Ciocca, G. and Zuffi, S.: A Survey of Methods for Colour Image Indexing and Retrieval in Image Databases. Color Imaging Science: Exploiting Digital Media, (MacDonald, L. W. and Luo, M. R. Eds.), Wiley, J. & Sons Ltd., Chichester, England. (2001)

9. Wang, J. Z., Li, J. and Wiederhold, G.: SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence. **23** (2001) 947–963

10. Wang, Z. and Bovik, C.: A Universal Image Quality Index. IEEE Signal Processing Letters. **9** (2002) 81–84

Edge Projection-Based Image Registration[*]

Hua Yan[1,2], Ju Liu[1], and Jiande Sun[1]

[1] School of Information Science and Engineering, Shandong University,
Jinan 250100, Shandong, China
{yhzhjg,juliu,jd_sun}@sdu.edu.cn
[2] National Laboratory on Machine Perception, Beijing 100871, China

Abstract. Image registration is a crucial step in all image analysis tasks. In this paper, an edge projection-based image registration method is proposed. First, Radon transform is used to project the edges of images along different directions, then cross-correlation-based approach is used to estimate global rotation and shift, finally polynomial fit is used to improve the accuracy of image registration. Simulation indicates that the proposed method can estimate global rotation and shift accurately and speedily.

1 Introduction

For obtaining more information about a certain scene, different images about the same scene are taken at different times, from different viewpoints, and/or by different sensors. One of the images is chosen as the reference image. Image registration is the process of determining mapping relations between the reference one and the others. Image registration is a crucial step in all image analysis tasks and extensively applied in remote sensing, medicine, cartography, computer vision and so on.

Image registration approaches can be sorted by characters (based on region and feature) or by four basic registration steps (feature detection, feature matching, transform model estimation, image resampling and transformation) [1]. But traditional image registration techniques are carried out in 2-D image space directly and have heavy computational cost. Therefore, in 1986 Alliney etc. [2] matched images by combining projection and a relative phase approach in which 2-D problem was converted to 1-D, thus computational complexity and cost were both reduced. Hereafter, Kim etc. [3] and Sauer etc. [4] used integral projection and correlation-based block matching algorithm to quicken video coding. Cain etc. [5] transformed images into two vector projections, and then matched images taking advantage of cross-correlation-based method. They also discussed the effectiveness of the algorithm in the presence of fixed-pattern noise. In terms of the properties of Radon transform such as shift, rotation and scaling, You etc. [6] discussed registration problem about image translation, rotation and uniform scaling. Tsuboi etc.[7] detected the position and orientation of planar motion objects using one-side Radon transform.

In this paper, an edge projection-based image registration method is proposed. In the method, the edges of images are projected along different directions using Radon

---

transform, then global shift and rotation are estimated based on maximum cross-correlation. In order to raise registration precision further, polynomial fit [8] is introduced. Simulation indicates that the proposed method can estimate global rotation and shift accurately and speedily.

2   Radon Transform and Its Properties

A projection of a 2-D function $f(x, y)$ can be seen as a linear integral in a certain direction, shown as fig.1.

(a)   (b)   (c)   (d)
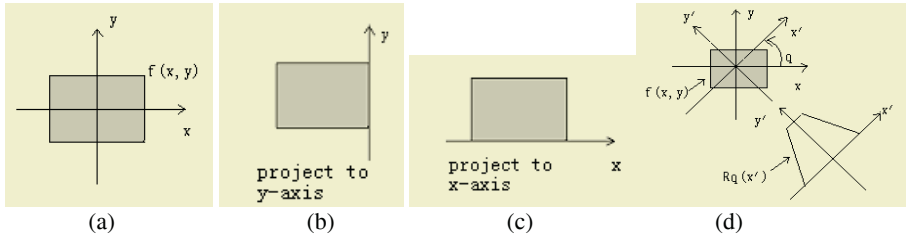
Fig. 1. The projections for a simple two-dimensional function $f(x, y)$ (a) are shown. (b) is the linear integral of $f(x, y)$ in the vertical direction which is the projection of $f(x, y)$ onto the x-axis; (c) is the linear integral of $f(x, y)$ in the horizontal direction which is the projection of $f(x, y)$ onto the y-axis. (d) is the projection of $f(x, y)$ onto either direction[9]

Image projection is a linear integral of the image along specified directions. To digital image, the projection is a linear accumulation along the directions.

Projections can be computed along any angle θ. In general, the radon transform of $f(x, y)$ is the linear integral of $f(x, y)$ parallel to the y′ axis, defined by

$$R_{\theta,x'}\left(f(x, y)\right)=\int f(x' \cos\theta - y' \sin\theta, x' \sin\theta + y' \cos\theta)dy' = p(x', \theta) \qquad (2.1)$$

To digital image, the function is expressed by

$$R_{\theta,x'}\left(f(x, y)\right)=\sum_{y'} f(x' \cos\theta - y' \sin\theta, x' \sin\theta + y' \cos\theta) = p(x', \theta) \qquad (2.2)$$

where
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Fig. 1(d) also can illustrate the geometry of the Radon transform.

Radon transform has shift-invariable and rotation-invariable properties [10].

Property1: If $f(x-x_0, y-y_0)$ is obtained by shifting $f(x, y)$, then the shift $x' = x_0 \cos\theta + y_0 \sin\theta$ of $p(x', \theta)$ denotes the shift of $f(x, y)$.

The projection of $f(x-x_0, y-y_0)$ along any angle θ can be expressed as $p\left[x' - (x_0 \cos\theta + y_0 \sin\theta), \theta\right]$.

Property 2: If $f_r(r, \theta + \theta_0)$ is obtained by rotating $f_r(r, \theta)$ (r denotes polar coordinate), then a variable θ_0 is added to the projection angle. The projection of $f_r(r, \theta + \theta_0)$ can be expressed as $p\left[x', \theta + \theta_0\right]$.

3 Edge Projection-Based Image Registration

Consider observing two successive frames $f_1(m, n)$ and $f_2(m, n)$ of a scene $g(m, n)$ at two distinct times by motion camera. The motion includes a shift (m_0, n_0) or an anticlockwise rotation θ_0 or both. The relations between two frames and the scene are shown as follow.

$$f_1(m, n) = g(m, n) + e_1(m, n) \tag{3.1}$$

$$f_2(m, n) = g(T(m, n)) + e_2(m, n) \tag{3.2}$$

$$T = \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & m_0 \\ \sin\theta_0 & \cos\theta_0 & n_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.3}$$

Where T denotes coordinate transform between $f_1(m, n)$ and $f_2(m, n)$. The size of the two frames is $M \times N$.

As is known, in some video acquisitions, the shading may vary from frame to frame even if there is no change in the external illumination. For example, if an object rotates its surface normal changes, the rotation results in a change in the shading. This change in shading may cause the intensity of the pixels along a motion trajectory to vary [11]. Then the projections in different directions using Radon transform lost their corresponding relations, such as property 2 in section 2. But the change in shading has little influence on some characters such as image edge.

In this paper, in order of estimating rotation between images about the same scene, image edges are extracted firstly according to (3.4) and (3.5). Then image edges $d_1(m, n)$ and $d_2(m, n)$ are projected along different directions using Radon transform and the corresponding projections $p_1(x', \theta)$ and $p_2(x', \theta)$ are obtained.

$$d_1(m, n) = edge(f_1(m, n)) \quad p_1\left(x', \theta\right) = R_{\theta, x'}\left(d_1(m, n)\right) \tag{3.4}$$

$$d_2(m, n) = edge(f_2(m, n)) \quad p_2\left(x', \theta\right) = R_{\theta, x'}\left(d_2(m, n)\right) \tag{3.5}$$

Before introducing the edge projection-based method, we review a common-used (based on the 2-D cross-correlation) approach for estimating the shift (m_0, n_0) between two frames $f_1(m, n)$ and $f_2(m, n)$ where the rotation is zero. The technique relies on maximizing the sample 2-D cross-correlation defined by

$$C(z, w) = \sum_{i=1}^{M} \sum_{j=1}^{N} f_1(m, n) f_2(z + m, w + n) \tag{3.6}$$

Where the shifts are assumed for convenience to be circular (Border effects resulting from the circular shift become negligible as N increases.). In particular, the traditional 2-D cross-correlation shift estimator selects the shift estimates $(\overline{m}_0, \overline{n}_0)$ by the rule

$$(\overline{m}_0, \overline{n}_0) = \arg\max_{z,w} C(z, w) \tag{3.7}$$

In terms of the properties of Radon transform, the shift and rotation between two frames can be seen as the shift of projection coordinates x' and θ. In contrast to (3.6), edge projection-based method also uses 2-D cross-correlation-based approach to estimate the shift of projection coordinates, defined by

$$C(z, w) = \sum_{x'} \sum_{\theta} p_2(x', \theta) p_1(z + x', w + \theta) \tag{3.8}$$

However, if we solve projection shift using (3.8), the translation along x'-axis will have the information about the rotation in image space. For avoiding the influence of this rotation, we change the location of projection direction in (3.8), shown as follow:

$$C(z, w) = \sum_{x'} \sum_{\theta} p_2(x', w + \theta) p_1(z + x', \theta) \tag{3.9}$$

But if we calculate cross-correlation to all the projections, the computational cost will be very large. Therefore, we only consider the case when θ is equal to zero or 90 degree.

$$C(z, w) = \sum_{x'} p_2(x', w) p_1(z + x', 0), when \quad \theta = 0 \tag{3.10}$$

$$C(z, w) = \sum_{x'} p_2(x', w + 90) p_1(z + x', 90), when \quad \theta = 90 \tag{3.11}$$

In the estimator the projection selects the shift estimates $(\overline{x}_0, \overline{\theta}_0)$ by the rule

$$(\overline{x}_0, \overline{\theta}_0) = \arg\max_{z,w} C(z, w) \tag{3.12}$$

We know that the definitions of coordinates in Radon transform and image space are different. In the vertical direction, y-axis of Radon transform can be seen as the rotation of 180 degree of n-axis of image space. Then the relation between the translation in image space and the translation of the projection is shown below.

$$m_0 = x_0, when \quad \theta = 0 \tag{3.13}$$

$$n_0 = -x_0, when \quad \theta = 90 \tag{3.14}$$

The estimator only adapts to pixel-level image registration which is thought as the precondition of the subpixel-level one. Here we can make use of polynomial fit to improve the accuracy of image registration. Take quadric fit for example. If we solve $(\overline{x}_0, \overline{\theta}_0)$ by (3.12) and find $C(\overline{x}_0, \overline{\theta}_0)$, its neighbours $C(\overline{x}_0 - 1, \overline{\theta}_0)$ and $C(\overline{x}_0 + 1, \overline{\theta}_0)$ also can be found. For gaining better fit curve, we can use more neighbours. Here the three neighbours are used to take the function of quadric fit, then the range $[\overline{x}_0 - 1, \overline{x}_0 + 1]$ is resampled. Cross-correlation corresponding to resampled nods is calculated and compared again to achieve more accurate projection translation.

4 Simulation

4.1 Results Using Simulation Sequence

In this section, we choose standard image 'cameraman' as reference image and simulate the estimation of global rotation, global rotation and shift respectively.

Firstly, the reference image is rotated in the range of 0~15º to create an image sequence. At the same time, Gaussian white noise of different SNR is added to the sequence.

At the beginning, the edges of the sequence are extracted. Then the edge of the reference is projected along the angle θ equal to zero. $p_1(x',0)$ denotes The projection. Then a range of projection direction is chosen experientially. The range is separated by 0.1º interval. The others among the edges are projected respectively along these angles in the range. $p_2(x',\theta)$ denotes these projections. We used (3.10) to calculate cross-correlation between $p_1(x',0)$ and $p_2(x',\theta)$ and estimate the rotation corresponding to different frame using (3.12). Fig. 2 shows the accuracy of the estimation and robustness to noise.

Next, we shift and rotate anticlockwise the reference and get motion frames. The shift and rotation is [5 5] and 5ºor [-1 1] and 3º. The edge projection-based method is implemented. We use (3.10) and (3.12) to get the rotation θ_0 and horizontal translation. We let w equal to θ_0 again and use (3.11) and (3.12) to get vertical translation. At last quadric fit is used to improve the accuracy of the shift estimation. The result of the estimator is shown by table 1.

4.2 Results Using Real Image Sequence

Real motion sequence is taken and the edge projection-based method is used to get global rotation and shift. The result of the estimation is shown by table 2. Fig. 3 is the part obtained by cutting motion sequence [12]. Fig. 4 is differential sequence. PSNR [11] is defined by (4.1). Fig. 4(b) and 4(d) demonstrate that edge projection-based method can accurately estimate global motion.

Table 1. Results using simulation sequence

Real motion	5º, [5 5]	3º, [-1 1]	*SNR* (dB)
Estimated motion	5.1º, [4.9 5.1]	3.2º, [-1 1]	infinite
	5.2º, [4.9 4.8]	3.1º, [-1 1]	20

Table 2. Results using real sequence

corresponding motion	rotation estimation	pixel-level shift estimation	subpixel-level shift estimation	*PSNR* (dB)
b to a	0.2º	[5 -12]	[4.7 -11.7]	57
c to a	1.6º	[7 -13]	[6.7 -13.1]	72
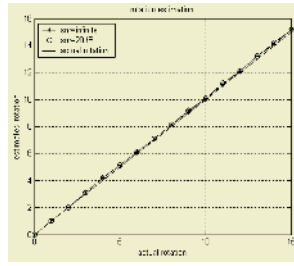
Fig. 2. Global rotation estimation. The line with * denotes rotation estimation when noise is free; the line with o denotes rotation estimation when snr is 20 db; the real line is actual rotation

| (a) | (b) | (c) |
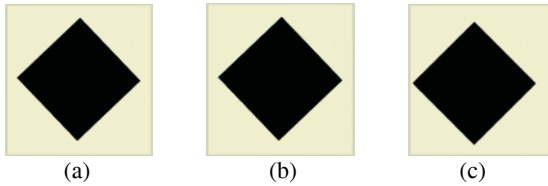
Fig. 3. Motion sequence

| (a) | (b) | (c) | (d) |
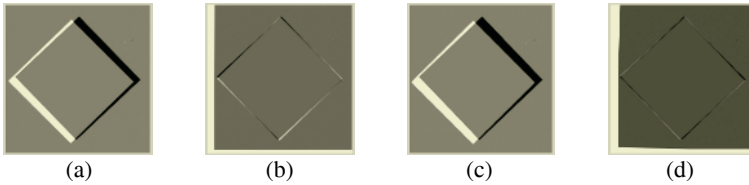
Fig. 4. Differential image. (a) is the differential between fig. 3(a) and 3(b). (b) is the differential between fig. 3(a) and the image gotten by motion-compensating the image shown by fig. 3(b). (c) is the differential between fig. 3(a) and 3(c). (d) is the differential between fig. 3(a) and the image gotten by motion-compensating the image shown by fig. 3(c)

$$PSNR = -10\log_{10}\frac{\sum\sum[I(m,n)-I_r(m,n)]^2/(K\times L)}{255^2} \qquad (4.1)$$

5   Conclusions and Future Work

This paper uses radon transform to project the edges of image sequence along some angles and takes advantage of cross-correlation-based method to estimate global motion between frames. Polynomial fit is used to improve the accuracy of global shift further. Simulation shows that the edge projection-based method can estimate global rotation and shift effectively. Though projection-based method for affine motion estimation had been proposed in [13], [14], projection-based motion estimation is still an opening problem. The proposition of more effectively projection-based scheme and the application to super-resolution image reconstruction will be our future work.

References

1. Zitova B., Flusser J.: Image registration method: a survey, image and vision computing, vol. 21 (2003) 977-1000
2. Alliney S., Morandi C.: Digital image registration using projections, IEEE Trans. Pattern Anal. Machine Intell., vol. 8 (1986) 222–233
3. Kim J.-S., Park R.-H., A fast feature-based block matching algorithm using integral projections, IEEE Journal on Selected Areas in Communications, vol. 10 (1992) 968–971
4. Sauer K., Schwartz B.: Efficient block motion estimation using integral projections, IEEE Trans. Circuits Syst. Video Technol., vol. 6 (1996) 513-518
5. Cain S. C., Hayat M. M., Armstrong E. E.: Projection-based image registration in the presence of fixed-pattern noise, IEEE Transactions on Image Processing, vol. 10 (2001) 1860–1872
6. You J., Lu W., Li J., Gindi G., Liang Z.: Image matching for translation, rotation, and uniform scaling by the Radon transform, Proceedings International Conference on Image Processing, vol. 1 (1998) 847–851
7. Tsuboi T., Masubuchi A., Hirai S.: Video-frame rate detection of position and orientation of planar motion objects using one-sided Radon transform, Proceedings IEEE Conference of Robotics and Automation, vol. 2 (2001) 1233–1238
8. Wang L. Q., Wei H. C., Zhou X. S.: Engineering numerical analysis, Shandong University Press, (2002)
9. Matlab 6.5, image processing toolbox.
10. http://pst.nst.pku.edu.cn/teaching/nuclear_medicine/DIPnotes/chapter4/chapter4_4.pdf.
11. Tekalp A. M.: Digital video processing, Prentice Hall PTR, (1998)
12. http://lcavwww.epfl.ch/reproducible_research/VandewalleSV04/
13. Milanfar P.: A model of the effect of image motion in the Radon transform domain, IEEE transactions on image processing, vol. 8 (1999) 1276-1281
14. Robinson D., Milanfar P.: Fast local and global projection-based methods for affine motion estimation, Journal of Mathematical Imaging and Vision, vol. 18 (2003) 35-54

Automated Information Mining
on Multimedia TV News Archives[*]

P.S. Lai[1], S.S. Cheng[1], S.Y. Sun, T.Y. Huang, J.M. Su, Y.Y. Xu[1], Y.H. Chen[1],
S.C. Chuang, C.L. Tseng[1], C.L. Hsieh[1], Y.L. Lu[1], Y.C. Shen[1], J.R. Chen[1],
J.B. Nie[1], F.P. Tsai[1], H.C. Huang[3], H.T. Pao[2], and Hsin-Chia Fu[1]

[1] Department of Computer Science and Information Engineering
National Chiao-Tung University, Hsinchu 300, Taiwan
[2] Department of Management Science
National Chiao-Tung University, Hsinchu 300, Taiwan
[3] Department of Electronics Engineering
National Chiao-Tung University, Hsinchu 300, Taiwan

Abstract. This paper addresses an integrated information mining techniques for multimedia TV-news archive. The utilizes techniques from the fields of acoustic, image, and video analysis, for information retrieval on news story title, newsman and scene identification. The goal is to construct a compact yet meaningful abstraction of broadcast news video, allowing users to browse through large amounts of data in a non-linear fashion with flexibility and efficiency. By using acoustic analysis, the system can classify video into news versus commercials, with 90% accuracy on a data set of 400 hours TV-news recorded off the air from July 2003 to August of 2004. By applying speaker identification and/or image detection techniques, each news stories can be segmented with an accuracy of 96%. On screen captions or subtitles are recognized by OCR techniques to produce the text title of each news stories. The extracted title words can be used to link or to navigate more related News contents on the WWW. In cooperation with facial and scene analysis and recognition techniques, OCR results can provide users with multimodality query for specific news stories. Some experimental results are presented and discussed for the system reliability and performance evaluation and comparison.

1 Introduction

Among the major sources of news program, TV has clearly had the dominant influence at least since the 1960s. Yet it is easy to find the old newspaper in microfilm in any public library, but it is impossible to find the old footage of television news in the same library. TV news archive has existed in the United States for 35 years. Paul C. Simpson founded the Vanderbilt University Television News archive in 1968. In [1], a team in the

---

University of Missouri-Columbia decided to do a content analysis of the three US network coverage of the 1989 Tiananmen Massacre, they located these news items in the Vanderbilt Archive Index. The Vanderbilt archive promptly provided the 11-hour video clips all related to the Tiananmen Massacre. At the same time, the Missourian team also planned to do a comparable study of Taiwanese reportage on Tiananmen Massacre. But the equivalent material of the Vanderbilt archive did not exist in Taiwan then. Therefore, that study only contained the US perspective of the Tiananmen Massacre. This paper proposes an integrated methodology for the information mining on a multimedia TV news archive in Taiwan. As described in [2, 3], A Fully Automated Web-Based TV-News System were implemented to achieve the following goals:

1. Academic and applied aspects: This archive will greatly improve the quality of TV news. As Dan Rather, the CBS anchorman, once mentioned that he lives with two burdens -*the ratings and the Vanderbilt Television News Archive* Therefore, once the archive is there, the researchers and the public will do some content analysis on the TV news. And the journalists will be more careful in what they report.
2. Timing factor: Vanderbilt archive started its project with Betacam video-tapes in 1968. There will be a problem of preservation because these tapes deteriorate along the years. Today, we can save all the TV news in hard disc, VCD or DVD.

Infomedia[4] is an integrated project launched in Carnegie Mellon university. Its overall goal is to use modern AI techniques to archive video and film media. VACE-II[5], a sub-project of Informedia, automatically detects, extracts, and edits highly interested people, patterns, and story evolves and trends in visual content from news video.

This paper proposes a systematic methodology that can automatically generate semantic labels from news video, and statistical methods to discover hidden information. We intend to expect that the following significance will come to exist.

– Although web-news provides another efficient way to access news, watching TV-news already becomes habit of many people. Beside this, most of web-news system can only provide text-based news.
– There are so many channels providing TV-news. People need more information for searching like-minded channel.
– Although almost every channel announced that they are dispassion, real dispassion is hard to archive with human editing. We need some evaluation to check if the channel is really dispassion.

The rest of this paper is organized as follows. In section 2, methods of generating necessary semantic labels from the recording TV news video are presented. Section 3 focus on describing the information mining from these semantic labels. Finally, summary and concluding remarks are given in section4.

2   News Information Tree Generation

The most important things in news story writing are that journalists commonly refer to as the 5 W's: who, what, when, where and why. These questions are crucial for catching a reader's attention and introducing the essential facts of the story. Standing on this basic rules, the news archive system introduced in our previous work[3] are further improved to extract more information from a recorded news video. A *news information tree* is suggested to structure the contents of recorded video clips for helping the user focus on specific news information, and information that is a little more general.

2.1   News Information Tree

The news information tree contains five types of video information records: (1) Date (when), (2) Channel (where), (3) Title (what), (4) Content (how), and (5)Commercial. The title record contains the starting time, length, and brief description of the corresponding video clips. The content record can also be further divided into the following sub-records: (a) on-site locations, (b)interview, and (c) tables or quoted word.

2.2   Analysis Units

Usually, a TV-news program contains the following items: news stories, commercials and weather reports. Complete description of shot detection and scene segmentation can be found in [6]. TV-news program segmentation is briefed as following: Among various scene shots, anchor video clips are detected first. In general, anchor segments are the most appeared video clips, thus we propose to use BIC[7], an unsupervised method, to cluster anchor segments from the other clips. This method contains the following procedures:

1. The MFCC audio feature sequence X is generate from input audio at first.
2. BIC segments X into segments $X_1, X_2,...X_n$.
3. These segments then are clustered as several clusters $C_1, C_2,...,C_3$.
4. The cluster containing most clip segments is the set of anchor clip.

   After locating each anchor shot, a SVM model based video classifier[8] is used to detect weather report shots. Finally, commercials are detected and separated from on-site news stories. Four feature detecting techniques - 1) The variation rate of zero crossing rate, 2) Short time energy, 3) Shot change rate, and 4) Clip length, are integrated to achieve a high performance commercial detector.
   At this stage, anchor's briefing, weather reports, commercial, and the background stories are all separated and identified. For each on-site stories, we further segment and classify each on-site scene into three categories: locations, interview and tables or quoted words (what).
   In order to segment storis into more detailed clips, the interview scenes are extracted first. Then the table or quoted words scens are extracted from the

rest and the rest scenes are locality scens. In general, on-site narration is not active during the interview scene, the interview scene can be distinguished from location scene. By using its special characteristics of the character regions, tables or quoted words scene can be distinguished from location scenes.

2.3   Semantic Labels of Units

This section describes how to assign each segmented unit with semantic labels. Basically, the text words for each label are extracting from text streams in close-caption.

Usually, a TV-news program often provides audience a quick overview of each news story in on-screen captions, such as names of location, people, and keywords of events, ... etc. In general, these texts are quiet enough to give enough information for labeling each segmented units.

The information tree establishing process contains two phases: story and scene phases.

In story phase, all on-screen characters are recognized by video-OCR first. Then, the recognized characters or words are used to match with text-based news documents, which are usually retrieved from Internet. The title and contents of best matched text-news document not only fill out the story information record of news information tree, but also used to picking label candidates, including locations, people names, event words, quoted phrases, and tabular data, up for scene phase processing.

In scene phase, picking semantic labels up from label candidates for each scene is done in this phase. As shown as figure 1(a), the on-screen captions of locality scene provide location name and event descriptions. Therefore, in locality scene, the location and event words of label candidates are searched from on-screen captions to find which ones are exactly appeared.

Fig. 1. An example of (a)locality, (b)interview, and (c)data chart scene frames.

Figure 1(b) is an example of interview scene. In interview scene, the intervie-wee's name and their points are always given by on-screen captions. Therefore, in interview scene, we search people name, and events word of label candidates instead.

The example of data chart scene is shown as figure 1(c). In general, on-screen captions fill data chart scene. These captions may present quoted sentence or tabular data. Therefore, searching for quoted sentence or tabular data in data chart scene is the major task.

3   Data Mining on the News Information Tree

This section presents how and what to mine from a news information tree (NIT). In Section 3.1, we propose to mine the favored or preferred news contents of a TV-station. The news information tree can also be used to track the evolution of a series of news stories (see Section 3.2). In addition, the mining results from the NIT and the realtime ratings can be combined to provide TV-news commercial buyers a very useful guidance.

3.1   Mine the News Preference of a TV Station

Generally speaking, a TV-station arranges the broadcasting sequence of each story in a news program according to their impact and attractiveness to audience. In fact, a preferred news story often gets more time on the air. By analyzing the sequence order and the length of stories, the preferred or the favored news stories of a TV-station can be roughly estimated or judged. Mining the NIT to extract favored or preferred types of news story from a TV station will help audience to find favor news channel.

The proposed news mining method is described as follows: Given N sets of keywords, $K_1, K_2, ..., K_i, ..., K_N$, which correspond to N news topics (or subjects), let the following delta function $\delta(k, K_i)$ define the relations between a keyword k and a keyword set K_i:

$$\delta(k, K_i) = \begin{cases} 1, \text{if keyword } k \in K_i \\ 0, \text{otherwise.} \end{cases}$$

1. Extract keywords $\{k_{s_j}^l;\ l = 1, \cdots, L_j\}$ from a scene unit s_j in a news program.
2. For each scene units s_j, compute its association frequency $F(K_i|s_j)$ with respect to a subject K_i,

$$F(K_i|s_j) = \frac{\sum_{l=1}^{L_j} \delta(k_{s_j}^l, K_i)}{\sum_{i=1}^{N} \sum_{l=1}^{L_j} \delta(k_{s_j}^l, K_i)}$$

3. Compute the the association frequency of news program $F_d(t|K_i)$ at time t and day d:

$$F_d(t, K_i) = \begin{cases} F(K_i, s_1), & \text{for } s_1.start \leq t \leq s_1.end \\ F(K_i, s_2), & \text{for } s_2.start \leq t \leq s_2.end \\ \vdots & \vdots \\ F(K_i, s_M), & \text{for } s_M.start \leq t \leq s_M.end, \end{cases}$$

where $s_j.start$ and $s_j.end$ are the start and the end time of a scene unit s_j in a news program at time t of the day d.

The associated frequency distribution from one segment of news program is not enough to represent the overall preference or trend of a news channel, thus long term statistics is needed.

By accumulating a longer period (say one month) of associated frequency of news subjects, the preference of a channel can be discovered. Figure 2 is an example of such station trend statistics. In this example, two months data are collected and test with three categories of keywords. From this example, we can find the trends of target station. More data would help us discover more.
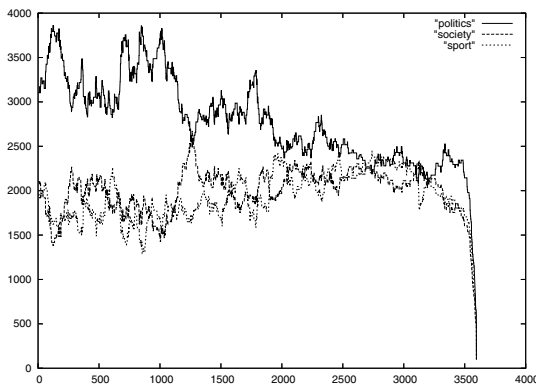
Fig. 2. Three sets of (representative) keywords are used to associate the appearing frequency of *social*, *political* and *sport* news in a news channel.

3.2   The Evolution of a Series of News Stories

The evolution of a news story can also be mined from the news information tree. By associating the keywords of a specific event with recorded news scenes over a period of days, then the accumulated association frequency of matched scene units presents an overall developing and progressing of the specific news stories.

3.3   The Mining on TV Commercial

Beside background stories, commercial records are also valuable information. Huang et al., [6] proposed commercial detecting and identifying methods in TV video clips. When a commercial frame contains image keywords in a video frame, video OCR techniques can be used to extract keywords to label the corresponding video clips. Otherwise, keyblock-based image retrieval methods [9] may be utilized to represent and to identify each commercial clips. However, manual annotation is needed to label the keyblock. By gathering statistical information

of these labels and keywords in news programs, cross relationship between TV commercials, realtime ratings, and news stories can be observed and analyzed to achieve a useful marketing database.

4   Conclusion

This paper addresses techniques and possible applications of fully automated information mining on a multimedia TV-news archive. The proposed information mining techniques contain the following processes: (1) segmenting a TV-news program video recording into scene clips, (2) using video OCR to extract and recognize close-caption and/or image characters into keywords for each scenes, (3) using keywords to generating semantic labels for each scenes, and (4) segmenting commercial video clips from news clips. These labels and scene information (e.g., the starting and ending time of a scene) are stored in the proposed *news information tree*. Statistics of news information tree shows many hidden information, like favor of channel and evolution of stories. These information helps people find their desired news-channel, searching most interviewed people, track story, ... , and so on.

References

1. Huffman, S., Yang, T.E., Yan, L., Sanders, K.: Genie out of the bottle: Three u.s. networks report tiananmen square. In: Proceedings of the annual meeting of Association for Education in Journalism and Mass Communication, Minneapolis, Minnesota, USA (1990)
2. Xu, Y., Chen, Y., Tseng, C., Lai, P., Hsieh, R., Lu., Y., Shen, Y., Fu, H.C.: Multimedia tv news browsing system. In: Proceedings. IEEE International Conference on Multimedia and Expo, Taipei, Taiwan, ROC (2004)
3. P.S.Lai, L.Y.Lai, T.C.Tseng, Y.H.Chen, Fu, H.C.: A fully automated web-based tv-news system. In: Proceedings of PCM2004, Tokyo, Japan (2004)
4. Informedia: http://www.informedia.cs.cmu.edu/ (2005)
5. VACEII: http://www.informedia.cs.cmu.edu/arda/vaceii.html (2005)
6. Tzu-Yang Huang, P.S.L., Fu, H.C.: A shot-based video clip search method. In: Proceedings. of CVGIP2004, Taipei, Hualien, ROC (2004)
7. Fraley, C., Raftery, A.E.: How many clusters? which clustering method? answers via model-based cluster analysis. Computer Journal **41** (1998) 578–588
8. Sun, S.Y., C.L.Tseng, Y.H.Chen, S.C.Chuang, H.C.Fu: Cluster-based support vector machine in text-independent speaker identification. In: Proceedings of International Joint Conference on Neural Networks IJCNN 2004, Budapest, Hungary (2004)
9. Zhu, L., Rao, A., Zhang, A.: Theory of keyblock-based image rerieval. ACM Trans. on Information Systems **20** (2002) 224–257

An Emergency Model of Home Network Environment Based on Genetic Algorithm

Huey-Ming Lee and Shih-Feng Liao

Department of Information Management, Chinese Culture University
55, Hwa-Kung Road, Yang-Ming-San, Taipei (11114), Taiwan
hmlee@faculty.pccu.edu.tw, hf_liao@mail2000.com.tw

Abstract. In this paper, we proposed an emergency model of home network environment based on genetic algorithm. This model can not only adapt the home network environment by using genetic algorithm but also detect the emergency events automatically. There are four modules in this model, saying, training knowledge base (TKB), genetic operator module (GOM), emergency knowledge base (EKB), and emergency early warning module (EEWM). TKB receives the messages from the environment and provides them for GOM to train EKB. GOM trains the EKB to fit the real situation by using genetic algorithm. EKB includes the database and the rule base which can provide messages for EEWM to infer. EEWM determines the emergency situations by fuzzy inferences and sends the caution messages to the users by mobile devices. Via this model, our home network environment will become more reliable and safer.

1 Introduction

Since information appliances (IAs) have been available to all in recent years, there are more and more varied IA products appeared. IA plays an important role in home environment. In home network environment, an IA control mechanism can provide the fine control capability of IA devices.

Lee and Huang [9] proposed an IA controlling model (IACM) that can control IA devices through home management broker. Lee et al. [8] came up with the idea of IAs intelligent agent model (IAIA) that makes home environment more comfortable and convenient. Lee et al. [14] proposed an intelligent control model of information appliances (IACMIA) that can keep the home environment more comfortable. Lee et al. [11] proposed an adaptive exception process model of information appliances that can adjust the membership functions of the home environment variables to make the reasoning more approximate to real situation. If there is a mechanism that can do the active response of emergency, then we can prevent the serious accident in home network environment.

In this study, we proposed an emergency model of home network environment based on genetic algorithm. There are four modules in this model, saying, training knowledge base (TKB), genetic operator module (GOM), emergency knowledge base (EKB), and emergency early warning module (EEWM). TKB receives the messages from the environment and provides them for GOM to train EKB. GOM trains the EKB to fit the real situation by using genetic algorithm. EKB includes the database and rule base that can provide the messages for EEWM to infer. EEWM determines

the emergency situations by fuzzy inferences and sends the caution messages to the users by mobile devices. Via this model, our home network environment will be more reliable and safer.

2   Framework of the Proposed Model

In this section, we presented an emergency model of home network environment based on genetic algorithm (EMHNE) under the supervision of IAIA [8], as shown in Fig. 1. There are four modules in this model, saying, training knowledge base (TKB), genetic operator module (GOM), emergency knowledge base (EKB), and emergency early warning module (EEWM), as shown in Fig. 2.
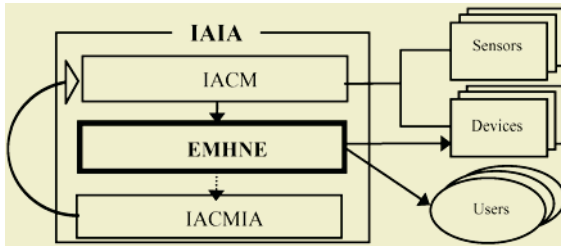
Fig. 1. Proposed model of the home network environment (Dotted line is normal event)
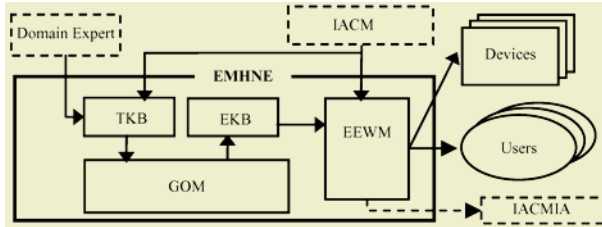
Fig. 2. Framework of EMHNE (Dotted line is normal event)

   The functions of these modules are as the follows:

- ◆ TKB: It can receive the messages from the home network environment and provide the messages for GOM to train EKB.
- ◆ EKB: It includes the database and rule base that can provide messages for EEWM to infer.
- ◆ EEWM: By the inference results, it will send the warning messages to the users or pass by.
- ◆ GOM: It trains the EKB to fit the real situation by using genetic algorithm.

2.1   EEWM

After receiving the messages from IACM, EEWM can determine the emergency events by fuzzy inferences and send the warning message to the users by mobile devices. There are three components in this module, saying, fuzzy inference engine (FIE), device control component (DCC) and emergency cautioner component (ECC).

FIE can reason the emergency events by fuzzy inference. If the emergency events occurred, DCC would start the devices up and ECC would send the warning messages to the remote users. If the events are not emergent, FIE will let the messages pass by, as shown in Fig. 3.
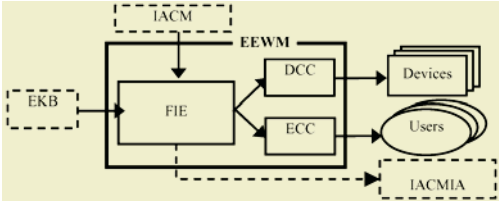
Fig. 3. Framework of EEWM (Dotted line is normal event)

2.2 TKB

There are two components in TKB, saying, updated rules base (URB) and updated database (UDB). The domain experts may update the URB. UDB can be updated by the messages from IACM. Both of them provide GOM to train the EKB, as shown in Fig. 4.
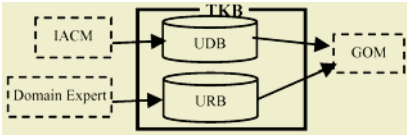
Fig. 4. Framework of TKB

2.3 EKB

EKB comprises of two components, saying, genetic tuning database (GTDB) and genetic tuning rule base (GTRB). While GOM receives the training data from TKB, it can tune the scaling factors and train rules in the EKB by using genetic algorithm for EEWM to reason, as shown in Fig. 5.
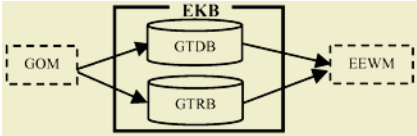
Fig. 5. Framework of EKB

2.4 GOM

GOM comprises of three components, saying, encode component (EC), optimize component (OC) and decode component (DC). EC can encode environment messages and rules as chromosomes. OC can optimize these chromosomes by using crossover and mutation. DC can decode these chromosomes to update EKB, as shown in Fig. 6.

Fig. 6. Framework of GOM

(Ⅰ) Rule Training Process

For the sake of getting comprehensive rule base, we can train the rule base by using genetic algorithm [1-4]. The encode approach is shown in Fig. 7.

Fig. 7. Initial population encode approach

After encoding the rule base, we can train the rule base by using genetic algorithm.

(Ⅱ) Adjusting Scaling Factors Process

In order to generate more suitable database, we can adjust the scaling factors by using genetic algorithm [1-4]. The adjusting of scaling factors can be regarded as global tuning procedures. The reasoning steps are as shown in Fig. 8.

Fig. 8. Fuzzy inference with scaling factors [6]

Nomenclature:

SF_{i1} : Input scaling factor of parameter 1.

SF_{i2} : Input scaling factor of parameter 2.

SF_{o3} : Input scaling factor of parameter 3.

We set the initial scaling function as Eq. (1)

$$f(x) = \alpha x + \beta \tag{1}$$

If $\alpha = 1$ and $\beta = 0$, then the tuning fits the real situation.

3   Model Implementation

For the purpose of ease manipulation, cross-platform, and remote-control capability, we constructed this model by Java Server Page (JSP), Java Servlet written Web Server structure, Java 2 Platform, Standard Edition, v1.4.2 API Specification and MATLAB 7.0. Above mentions are done with Pentium 4 1.6G desktop that is powered by O/S Windows Professional and Microsoft Access 2002.

We take two parameters as environment variables and one parameter as consequence. We set on the rule base training as shown in the following Table 1:

Table 1. Rule setting

Parameter	Labels	Type
Temperature	3	environment variable
Visibility	5	environment variable
Emergency degree	5	consequence

- Population size: 20
- Probability of crossover: $P_c = 0.7$
- Probability of mutation: $P_m = 0.2$
- Selection: roulette wheel
- Elite: enable

The training results are shown in Table 2, Fig. 9 and Fig. 10:

Table 2. Rule training result

Parameter	Rules
Before Training	3
After Training	13
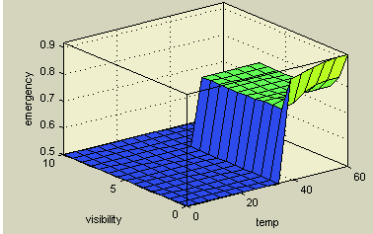
Fig. 9. Rule base before training
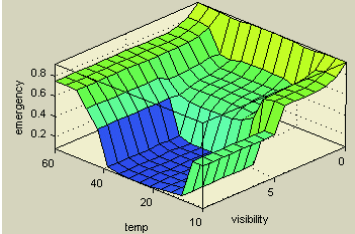
Fig. 10. Rule base after training

After training, we have that the rule base is more robust.

Next stage, we tune the scaling factors with multiple inputs single output (MISO) system. We set the scaling factors tuning as shown in Table 3:

Table 3. MISO system

Parameter	Labels	Type
Parameter 1	3	input
Parameter 2	5	input
Parameter 3	5	output

- Population size: 200
- Probability of crossover: $P_c = 0.8$
- Probability of mutation: $P_m = 0.1$

◆ Selection: roulette wheel
◆ Elite: enable

The tuning results are shown in Table 4:

Table 4. Parameter 1 scaling factors tuning result

	Parameter 1		Parameter 2		Parameter 3	
Scaling factor	α	β	α	β	α	β
Before Training	1	0	1	0	1	0
After Training	1.21	0.25	1.12	0.66	1.15	0.03

After tuning, the scaling factors are more suitable for real situation.

4   Conclusion

At present, there are more and more varied IA products appeared in home network environment. The emergency process mechanism is the most significant in home network environment. In this study we proposed an emergency model of home network environment based on genetic algorithm (EMHNE). This model can not only adapt the home network environment by using genetic algorithm but also detect the emergency events automatically. Via this model, our home network environment will be more reliable and safer.

References

1. Arslan, A., Kaya, M: Determination of fuzzy logic membership functions using genetic algorithm. Fuzzy sets and systems Vol. 118, (2001) 297-306
2. Cordón, O., Herrera, F.: A Three −Stage Evolutionary Process for Learning Descriptive and Approximate Fuzzy-Logic-Controller Knowledge Bases From Examples. International Journal of Approximate Reasoning, Vol.17 (1997) 369-407
3. Cordón, O., Herrera, F., Hoffmann, F., Magdalena, L.: Genetic Fuzzy Systems. World Scientific Publishing Co. (2001)
4. Gurocak, H.B.: A genetic-algorithm-based method for tuning fuzzy logic controllers. Fuzzy sets and systems Vol. 108, (1999), 39-47
5. Herrera, F., Lozano, M., and Verdegay, J.L.: Tuning Fuzzy Logic Controller by Genetic Algorithm. International Journal of Approximate Reasoning, Vol.12 (1995) 299-315
6. Ju, M.-S., Yang, D.-L.: Design of adaptive fuzzy controls based on natural control laws. Fuzzy sets and systems Vol. 81 (1996) 191-204
7. Jung, C.-H., Ham, C.-S., Lee, K.-I.: A real-time self-tuning fuzzy controller through scaling factor adjustment for the steam generator of NPP. Fuzzy sets and systems Vol. 74 (1995) 53-60
8. Lee, H.-M., Chen, Y.-C., Chen, J.-J.: The Intelligent Agent Design of Information Appliance. JCIS, 2003, Proceeding of the 7th Join Conference on Information Sciences, Cary. NC. USA, (2003) 1681-1684
9. Lee, H.-M., Huang, J.-H.: The study of IA devices monitoring model, The sixth seminar of the research and practices of information management, (2002) 430-437
10. Lee, H.-M., Liao, H.-F., Lee, S.-Y.: A Remote Authentication Model of Information Appliances. WSEAS TRANSACTIONS on INFORMATION SCIENCE & APPLICATIONS, Issue 2, Vol. 1 (2004) 728-732

11. Lee, H.-M., Liao, H.-F., Lee, S.-Y.: An Adaptive Exception Process Model of Information Appliances. WSEAS TRANSACTIONS on INFORMATION SCIENCE & APPLICATIONS, Issue 2, Vol. 1 (2004) 778-783
12. Lee, H.-M., Mao, C.-H.: A Fuzzy Clustering Model of Information Applicance. Third International Conference on Electronic Business (ICEB2003) Singapore, (2003) 241-243
13. Lee, H.-M., Mao, C.-H., Lee, S.-Y.: A Fuzzy Neural Network of Information Appliance. International Workshop on Fuzzy System & Innovation Computing 2004 (FIC2004), Kitakyushu, Japan, (2004)
14. Lee, H.-M., Mao, C.,-H., Lee, S.-Y.:Intelligent Control Model of Information Appliances. Negoita, M.-G, Howlett, R.-J., Jain, L.-C. (Eds.) Knowledge-Based Intelligent Information and Engineering Systems, Part 3, 123-128 (2004)
15. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)
16. Siarry, P., Guely, F.: A genetic algorithm for optimizing Takagi-Sugeno fuzzy rule bases. Fuzzy sets and systems, Vol.99 (1998) 37-47

A Distributed Backup Agent
Based on Grid Computing Architecture

Huey-Ming Lee and Cheng-Hsiung Yang

Department of Information Management, Chinese Culture University
55, Hwa-Kung Road, Yang-Ming-San, Taipei (11114), Taiwan
hmlee@faculty.pccu.edu.tw, a4269385@ms38.hinet.net

Abstract. In this paper, we proposed a distributed backup agent based on grid computing architecture. In grid environment, this agent can assist user to backup the data files in different nodes storages, update replicas synchronously, and access the replicas if necessary. There are two modules in this agent, saying, grid consistency service module (GCSM) and grid transaction module (GTM). GCSM can monitor the grid backup nodes status, e.g. storages space, on-line available, etc. GTM can maintain the consistency and integrity of backup files. This agent makes the distributed backups of the data files be more convenient and secure to access.

1 Introduction

Recently, files in storages are possibly corrupted by some faults, such as noise, human errors, and hardware failures, etc. In order to ensure the data security and prevent the original files from corrupting, we often backup the full files in others locations [7]. However, it takes much time and cost to backup full files, especially in volumes of files. In order to reduce the backup cost, Nakamura et al. [7] proposed a stochastic model to achieve an optimal interval for a full backup that minimizes the expected cost and compute.

The term "Grid" was coined in the mid 1990s to denote a proposed distributed computing infrastructure for advanced science and engineering [3]. In grid environment, users may access the computational resources at many sites [2, 4]. The functions of information systems based on grid computing architectures are resources (e.g., CPUs, storages, etc.) sharing, collaborative processing, reliable and secure connection, etc. Data grids focus on providing secure accessing to the distributed, heterogeneous pools of data [5]. Due to the advantages of grid computing, Lee et al. [6] proposed a grid distributed backup of data files model. The functions of this model are capable to divide a data file into some sub-files, replicate, connect with each others to access the backup sub-files, and provide the secure mechanism for data distributed backup.

To process the backup files is repeatable and scattered, so it needs an automatic mechanism to do the backup. The "agent" is autonomous objects created for dynamic and distributed applications that are responsible for executing designated tasks [8].

In this study, we proposed a distributed backup agent based on grid computing architecture. In order to achieve secure transmission and data integrity, we can divide the file into some replicas, and store them in the different nodes storages on demands. Even if an unauthorized user wants to get these replicas, he/she can not obtain a com-

plete file. When any one of the replicas is updated, then the others related replicas would be updated synchronously. The proposed agent can assist user to backup the data files in different nodes storages, update replicas synchronously, and access the replicas if necessary. This agent makes the distributed backups of the data files be more convenient and secure to access.

2   Framework of the Proposed Agent

In this section, we presented a distributed backups agent (DBA) based on grid computing architecture. This agent was built on the grid node, saying M_1, as shown in Fig. 1.

where   SS denotes source storage,
         M_i denotes backup server node,
         S_i denotes the server node's storage
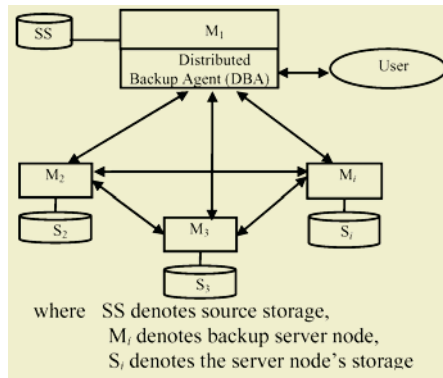
Fig. 1. Context of the proposed agent (DBA)

The proposed agent can assist user to backup the data file by dividing it into some sub-files and store them in the different nodes storages, update them synchronously, and access them conveniently if necessary. There are two modules in this agent, saying, grid consistency service module (GCSM) and grid transaction module (GTM), as shown in Fig. 2. Grid consistency service module (GCSM) can transfer the monitoring grid nodes status (e.g., storages spaces, on-line available, etc.) and the user's processing requirements (e.g., dividing, accessing, or updating, etc.) to GTM. GTM can do the dividing, accessing, or updating the distributed backups depending on user's processing requirements, and return the backup analysis messages (e.g., index files and the backup related analysis results, etc.) to GCSM. GCSM will divide, access, or update synchronously depending on the backup analysis messages.
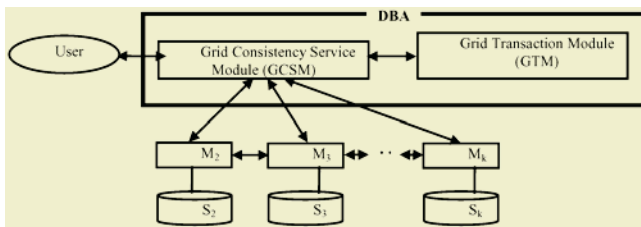
Fig. 2. Framework of the proposed agent (DBA)

♦ Grid consistency service module (GCSM): It can monitor the grid nodes status and do the dividing, accessing, or synchronously updating files.
♦ Grid transaction module (GTM): It can maintain the consistency and integrity of backup files.

2.1 Grid Consistency Service Module

Grid consistency service module (GCSM) can monitor the status of backup nodes. Also, it can manipulate the backup files dynamically, such as, dividing a file, accessing distributed replicas, and updating the related replicas, etc. There are two components in GCSM, saying, distributed message interaction component (DMI) and distributed backup manipulation component (DBM), as shown in Fig. 3.
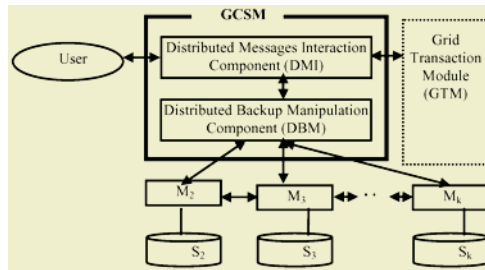
Fig. 3. Framework of the grid consistency service module (GCSM)

The functions of these two components are as follows:
♦ Distributed messages interaction component (DMI): DMI interacts with others components in security. It can receive the user's processing requirements (e.g., dividing, accessing, or updating, etc.) and grid nodes status (e.g., storage spaces, on-line available, etc.) from DBM, then send these messages to GTM for analyzing the backup files depending on these messages. After analyzing, GTM would produce the backup analysis messages (e.g., index files, backup related analysis results, etc.), and send them to DMI. Finally, DMI can send the messages to DBM for practical execution
♦ Distributed backup manipulation component (DBM): DBM dispatches replicas, divides files, accesses replicas, and updates all related replicas synchronously depending on the analysis messages. In addition, it can monitor the grid nodes status.

In security, DBM follows GSI [1] standard and has the capability of authentication, authorization and delegation. Via the GSI X.509 Proxy Certificates, DBM can make the connection between clients and gird environment more secure, and achieve the feature of single sign on.

In collaborative computers connecting, DBM follows monitoring and discover system (MDS) [1] to access the static and dynamic resources messages. For more detail of MDS, please refer to [4].

For archiving file consistency, DBM has a locked mechanism to hold and release the replicas in others node storages, also, it can update all the related replicas synchronously depending on the analysis messages.

During the executing, DBM can detect the backup grid nodes status depending on the default or pre-set timing. Moreover, if one of the collaborative computers crashed, DBM would dispatch the replica from source storage to others nodes.

2.2   Grid Transaction Module

Grid transaction module (GTM) can divide, access, or update the distributed files depending on user's demand, and record the status of processing the replicas. There are three components, saying, distributed backup transferring module (DBTM), replica consistency catalogue management component (RCCM), data file backup management (DFBM), and one repository, saying, index repository (IR) in GTM, as shown in Fig. 4.
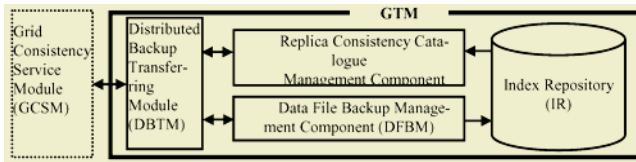
Fig. 4. Framework of the grid transaction module

The functions of these components are as follows:

♦ Distributed backup transferring module (DBTM): DBTM can divide, access, update the distributed files depending on user's demands, receive the analysis message from DFBM or RCCM, and transfer these messages to GCSM.
♦ Data file backup management (DFBM): It records the messages of dividing replica into index repository (IR). When the user wants to access the portion of required files, he should access the related replicas via IR.
♦ Index repository (IR): It stores the messages of the dividing and accessing files, and provides these messages to DFBM.
♦ Replica consistency catalogue management component (RCCM): RCCM accesses the messages of the dividing from index repository (IR), and records the process of updating replicas. Via the proposed distributed backup transaction algorithm in this agent as shown in Fig. 5, RCCM can prevent the backup files from losing and disappearing.

```
Grid distributed backup transaction algorithm {
While (file received) {
    Save as f[i] the times will be added (i++);
    Save as f[i] to temp;
    If (f[i] was modified)
      Save as s[i], and overwrite temp to f[i]
      If (s[i] was modified)
        Save as t[i], and overwrite temp to s[i],f[i]
      } }
```

Fig. 5. Grid distributed backup transaction algorithm

When any one of the replica was updated then the others related replicas would be updated synchronously.

3   Agent Implementation

In this section, we implemented the proposed agent in Section 3.1, and presented the applications of this agent in Section 3.2.

3.1   Implementation

For easy to manipulating, crossing platform and remote controlling, we applied Java Server Page (JSP) and Java Servlet written Web Server structure, Java 2 Platform Standard Edition v 1.4.2 API, globus toolkit to construct the proposed agent. The mentioned above is done by grid computing architecture with three Pentium 4, 2.4GHz computers that are powered by OS Linux Red Hat 9.0, 256MB RAM, a 100-Mbps network interface card, and 100-Mbps Ethernet. Each of computers is connected to HUB which is ZOT IA700.

The network environment was run through a LAN that reduced delay and loss of data files. The data files were transferred based on Gridftp.

We inputted data via user interface of DBA, as shown in Fig. 6, as examples.
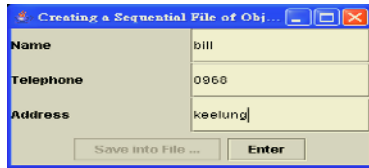
Fig. 6. The user interface of the input

Also, we took each field as a single replica. The contents of the original file were shown in Table 1. The backup files were composed of three replicas, saying, name, telephone and address, which were stored into different nodes storages as shown in Table 2.

Table 1. Contents of the original file

IP:192.0.0.10		
cloud	0920	Taipei
bill	0921	Keeling
john	0922	China

Table 2. Contents of the distributed backup in grid nodes

IP:192.0.0.1	IP:192.0.0.2	IP:192.0.0.3
cloud	0920	Taipei
bill	0921	Keeling
john	0922	China

3.2 Applications

The applications of this agent are as follows:

 (1) Recovery: During the executing, DBM detects the status of backup nodes depending on the default or pre-set timing. Moreover, if one of the collaborative computers crashed, DBM would dispatch the replica from source storage to others nodes. For example, if the node of IP in 192.0.0.1 was crashed, DBA would recover it from source storage.
 (2) Updating: If any one of the replicas was updated by user, then the other related replicas would be updated synchronously.
 (3) Accessing: Via the functions of grid, users can easily access the related replicas from the related grid nodes.
 (4) Dividing: This agent can divide data files into some replicas, and store them in the distributed nodes storages.

4 Conclusion

In this study, we proposed a distributed backup agent based on grid computing architecture. In order to achieve capabilities of secure transmission and data integrity, we can divide the file into some replicas on demands, and store them in the different grid nodes storages. Even if an unauthorized user wants to get these replicas, he/she can not obtain a complete file. When any one of the replicas is updated then the others related replicas will be updated synchronously. Moreover, via the proposed agent, it is easy to divide one file into some distributed backup replicas and store them in others grid nodes storages, access and update the replicas conveniently; also, via the proposed agent, security of the distributed backup files can be improved.

References

 1. Ferreira, L., Berstis, V., Armstrong, J., Kendzierski, M., Neukoetter, A., Takagi, M., Richard, R.-W., Amir, A., Murakawa, R., Hernandez, O., Magowan, J., Bieberstein, N.: Introduction to Grid Computing with Globus, IBM, (2002).
 2. Foster, I., Kesselman, C.: Globus: A Metacomputing Infrastructure Toolkit, International Journal of Supercomputer Application, Vol. 11, No. 2, (1997), 115-128
 3. Foster, I., Kesselman, C. (ed.): The Grid2: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (2004)
 4. Globus: http://www.globus.org (2005)
 5. Huang, C.-Y., Ceroni, J.-A., Nof, S.-Y.: Agility of networked enterprises-parallelism, error recovery and conflict resolution, Computers in Industry, Issue 2-3, Vol. 42, (2000), 275-287
 6. Lee, H.-M., Yang, C.-H., Hsu, C.-C.: A Grid Distributed Backups of Data Files Model, Proceeding of 11th Compiler Techniques for High-Performance Computing, Taichung, Taiwan, (2005), 174-177
 7. Nakamura, S., Qian, C., Fukumoto, S., Nakagawa, T.: Optimal backup policy for a database system with incremental and full backups, Mathematical and Computer Modelling, Issue 11-13, Vol. 38, (2003), 1373-1379
 8. Wooldridge, M.-J., Jennings, N.-R.: Agent theories, architectures and languages: a survey, Lecture Notes in Computer Science, Vol. 890, (1995), 1-39

A Dynamic Supervising Model
Based on Grid Environment

Huey-Ming Lee, Chao-Chi Hsu, and Mu-Hsiu Hsu

Department of Information Management, Chinese Culture University
55, Hwa-Kung Road, Yang-Ming-San, Taipei (11114), Taiwan
hmlee@faculty.pccu.edu.tw, jawchii@gmail.com,
hsu_clay@yahoo.com.tw

Abstract. In this paper, we proposed a dynamic supervising model based on grid environment. This model can utilize the grid resources, e.g., CPU, storages, etc., more flexible and optimal. There are three modules in this model, saying, connecting authentication module (CAM), trusting module (TM), and resource monitor module (RMM). CAM can authenticate the demander. TM can adjust trust degrees of the other collaborators by fuzzy inferences, and provide these trust degrees for RMM to schedule the works process. RMM can discover the resources, schedule and distribute works to others nodes. Once the demander passes the authentication, RMM will schedule and distribute works depending on the trust degrees from TM. RMM will dynamically monitor the resource variation of these collaborators. This model not only can make the collaboration more flexible and reliable, but also can optimize the grid resources.

1 Introduction

At present, the requirement of high performance computing is increasing at high speed. A single computer can not satisfy with large-scale problem solving. In this situation, grid computing is arisen in recent years and constructed high performance computing infrastructure. In order to make the cooperation go smoothly, it is necessary to have a monitor mechanism [1, 3, 4]. Miller et al. [5] proposed Paradyn parallel performance measurement tools which can identify the heaviest loading process by heuristic method and find out the bottleneck point, e.g., I/O, memory and CPU, etc. Wolski et al. [7 - 9] proposed Network Weather Service that can detect the related parameters in network on a regular time by the sensor spread and predict the coming situation via mathematical model. Ribler et al. [6] proposed Autopilot model that can tune the performance of heterogeneous computational grids based on closed loop control [2]. However, if there is a mechanism that can provide trust assessments of the collaborators' performances in the past and dynamically monitor resources, then the monitor mechanism will be more complete and the utilizing resources will be more flexible and optimal.

In this study, we proposed a dynamic supervising model based on grid environment. This model can make the collaboration more flexible and reliable, and optimize the grid resources.

In Section 2, we presented the framework of the proposed model based on grid environment. We implemented the proposed model in Section 3. Finally, we made the conclusions in Section 4.

2   Framework of the Proposed Model

In order to make the monitor model more complete and the utilizing resources more flexible and optimal, we introduced the trust degree in this model. The trust degree is a rating of collaborator executed performance in the past. By the trust degree, we can settle the priority of collaborators in the work schedules.

In this section, we presented the framework of the proposed dynamic supervising model based on grid environment (DSMGE). It was built on the grid node (node plays a role of resource consumer, resource producer role or both of them), as shown in Fig. 1. There are three modules in this model, saying, connecting authentication module (CAM), trusting module (TM), and resource monitor module (RMM), as shown in Fig. 2.

Fig. 1. Context of the proposed model
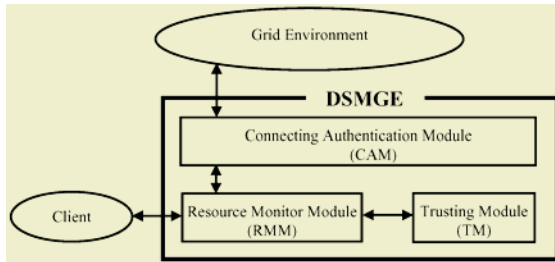
Fig. 2. Framework of the proposed model (DSMGE)

The functions of these modules are as follows:

- Connecting authentication module (CAM): It can authenticate the demander.
- Trusting module (TM): According to the workload and executed time, TM can adjust trust degrees of others collaborators by fuzzy inferences, and provide these trust degrees for RMM to schedule the works processes.
- Resource monitor module (RMM): It can discover the resources, schedule and distribute works depending on the trust degrees from TM, and dynamically monitor the resources variation of collaborators.

2.1   Connecting Authentication Module (CAM)

Connecting authentication module (CAM) can construct security connection and authenticate the demanders. It comprises of two components, saying, security connection component (SCC) and authentication component (AC). SCC receives authentic messages from demanders and sends these messages to AC for authenticating the demanders. Once the demander passes the authentication from AC, SCC will send the demander's requests to RMM; otherwise, AC will reject the demander's requests, as shown in Fig. 3.
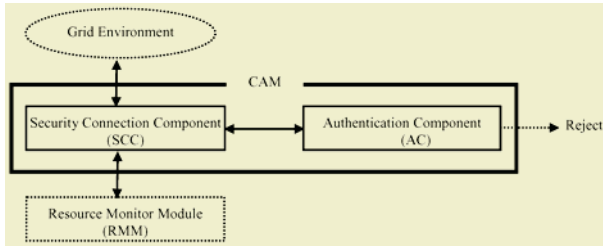
Fig. 3. Framework of the connecting authentication module (CAM)

2.2   Resource Monitor Module (RMM)

RMM can discover the resources, schedule and distribute works to others nodes. Once the demander passes the authentication, RMM will schedule and distribute works depending on the trust degrees from TM. When the distributed work was finished, RMM would send the finished work messages, e.g., work executed time, workload, etc., to TM for adjusting the trust degrees. RMM comprises of three components, saying, work receiving component (WRC), resources discovering and detecting component (RDDC), and works arranging component (WAC), as shown in Fig. 4.
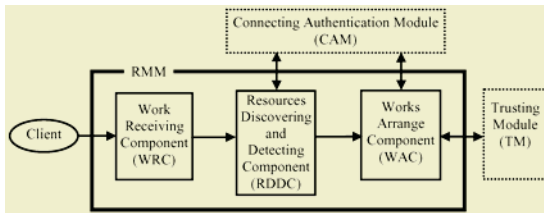
Fig. 4. Framework of the resource monitor module (RMM)

The functions of these components are as follows:

- Work receiving component (WRC): It can analyze work requirements from clients' demands, and send the analyzed requiring resources, e.g., CPUs, storages, etc., to RDDC.
- Resources discovering and detecting component (RDDC): After receiving the analyzed requiring resources from WRC, RDDC can discover resources in the grid environment and send available resources messages to WAC for scheduling the works processes. When the work was executing, RDDC would dynamically monitor resources variation of grid environment.
- Works arranging component (WAC): According to the messages from RDDC and trust degrees from TM, WAC can schedule and distribute works to proper grid nodes. When the work was finished, WAC would receive the related messages from the proper grid nodes and send them to TM for adjusting the trust degrees.

2.3   Trusting Module (TM)

According to the workload and executed time, trusting module (TM) can adjust trust degrees of others collaborators by fuzzy inferences, and provide these trust degrees for RMM to schedule the works processes.

The criteria ratings of workload are linguistic variables with linguistic values W_1, W_2, W_3, W_4, W_5, where W_1 = very few, W_2 = few, W_3 = middle, W_4 = large, W_5 = very large. These linguistic values are treated as fuzzy numbers with trapezoid membership functions as shown in Fig. 5.
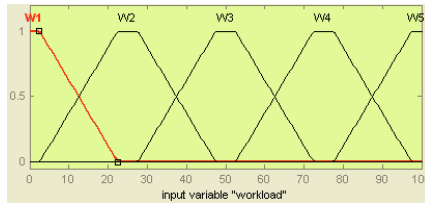
Fig. 5. Membership functions of the set of the criteria rating of workload

The criteria ratings of executed time are linguistic variables with linguistic values E_1, E_2, E_3, E_4, E_5, where E_1 = very short, E_2 = short, E_3 = middle, E_4 = long, E_5 = very long. These linguistic values are treated as fuzzy numbers with trapezoid membership functions as shown in Fig. 6.
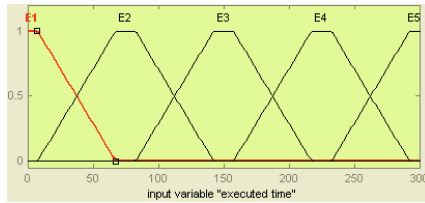
Fig. 6. Membership functions of the set of the criteria rating of executed time

The criteria ratings of trust degree are linguistic variables with linguistic values T_1, T_2, T_3, T_4, T_5, where T_1 = very low, T_2 = low, T_3 = middle, T_4 = high, T_5 = very high. These linguistic values are treated as fuzzy numbers with triangular membership functions as shown in Fig. 7.
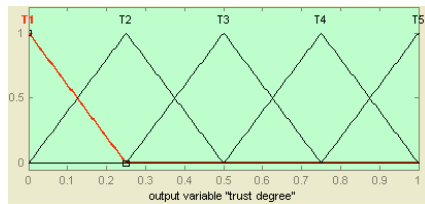
Fig. 7. Membership functions of the set of the criteria rating of trust degree

This module (TM) comprises of three components, saying, fuzzy trust judge component (FTJC), trust rules base (TRB) and trust degrees base (TDB). TDB includes trust degrees of collaborators. TRB can provide rules for FTJC to do the fuzzy inferences. When the work was finished, RMM would send the related messages, e.g., work executed time, workload, etc., to FTJC. When receiving these messages from RMM, FTJC will adjust the trust degree using fuzzy inference, and store this value to TDB, as shown in Fig. 8.
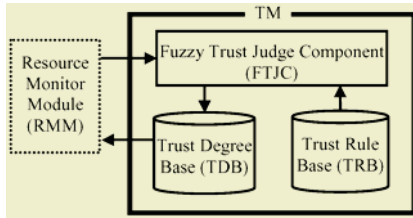
Fig. 8. Framework of trusting module (TM)

(I) Trust Degree Adjustment

We set zero as an initial value of trust degree of the new collaborated nodes. There are twenty-five rules in our rule base, and we take five rules for examples as follows.

if (workload is W_5) and (executed time is E_1) then (trust degree is T_5)
if (workload is W_4) and (executed time is E_2) then (trust degree is T_4)
if (workload is W_3) and (executed time is E_2) then (trust degree is T_3)
if (workload is W_2) and (executed time is E_4) then (trust degree is T_2)
if (workload is W_1) and (executed time is E_5) then (trust degree is T_1)

Fig. 9 shows that our rule base is reasonable and robust. We may adjust the trust degree by the rules.
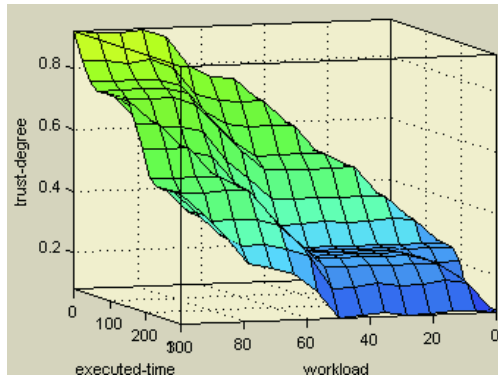
Fig. 9. Surface of the trust rule base (TRB)

(II) Trust Degree Application

We can schedule the priority of demanders depending on trust degrees.

3   Model Implementation

For the purpose of easy manipulation, cross-platform, and remote-control capability, we applied Java server page (JSP) and Java Servlet written Web server structure, Java 2 platform standard edition v 1.4.2 API. Above mentioned is done by grid computing architecture with three Pentium 4, 2.4GHz computers that are powered by OS Linux Red Hat 9.0, 256MB RAM, a 100-Mbps network interface card (NIC), and 100-Mbps Ethernet. Each computer is connected to HUB that is ZOT IA700.

We implemented the model without trust and with trust degrees respectively by matrix multiplying testing. First, we allotted the same workload to every proper grid node, as shown in Table 1.

Table 1. Allot work without trust degree

Node	A	B	C	Total time
Executed time (sec)	126	229	176	531

According to Table 1, we adjusted trust degrees to every proper node. The shorter executed time, the higher trust degree is. The higher trust degree, the more workload is. Because of more workload, the efficient nodes have to spend more executive time (e.g., node A), but we can shorten the total executive time via the trust degree allotted mechanism as shown in Table 2.

Table 2. Allot work with trust degree

Node	A	B	C	Total time
Trust	0.9	0.2	0.6	
Executed time (sec)	216	118	154	488

Table 2 shows that we can get shorter total executed time by trust degree allotted under the same total workload.

4   Conclusions

Nowadays the requirement of high performance computing is increasing at high speed. Grid computing is getting more and more significantly. Monitor model is one of the most important mechanisms in grid environment.

In this study, we proposed a dynamic supervising model based on grid environment. This model not only can make the collaboration more flexible and reliable, but also can optimize the grid resources. Via this model, we can make the monitor mechanism more complete and useful.

References

1. A Grid Monitoring Architecture: http://www.ggf.org/documents/GFD/GFD-I.7.pdf (2005)
2. Autopilot: http://www.renci.unc.edu/Project/Autopilot/autopilotoverview.htm (2005)
3. Foster, I., Kesselman, C. (ed.): The Grid2: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (2004)
4. Foster, I., Kesselman, C., Tuecke, S.: The Anatomy of the Grid: Enabling scalable virtual organizations. Int J High Performance Comp Appl Vol. 15, No. 3, (2001), 200-222
5. Miller, Barton P., Callaghan, Mark D., Cargille, Jonathan M., Hollingsworth, Jeffrey K., Irvin, R. Bruce., Karavanic, Karen L., Kunchithapadam, Krishna., Newhall, Tia.: The Paradyn Parallel Performance Measurement Tool. *IEEE Computer* Vol. 28, No. 11, (1995), 37-46

6. Ribler, Randy L., Vetter, Jeffrey S., Simitci, Huseyin, Reed, Daniel A.: Autopilot: Adaptive Control of Distributed Applications. Proceedings of the 7th IEEE Symposium on High-Performance Distributed Computing, Chicago, IL, (1998), 172-179
7. Wolski, Rich, Spring, Neil, Hayes, Jim: The Network Weather Service: A Distributed Resource Performance Forecasting Service for Metacomputing. Journal of Future Generation Computing Systems, Vol. 15, No. 5-6, (1999) 757-768
8. Wolski, Rich: Experiences with Predicting Resource Performance On-line in Computational Grid Settings. ACM SIGMETRICS Performance Evaluation Review, Vol. 30, No. 4, (2003) 41-49
9. Wolski, Rich: Dynamically Forecasting Network Performance Using the Network Weather Service. Journal of Cluster Computing, Vol.1, (1998), 119-132

An Intelligent Extracting Web Content Agent on the Internet

Huey-Ming Lee, Pin-Jen Chen, Yao-Jen Shih, Yuan-Chieh Tsai, and Ching-Hao Mao

Department of Information Management, Chinese Culture University
55, Hwa-Kung Road, Yang-Ming-San, Taipei (11114), Taiwan
hmlee@faculty.pccu.edu.tw

Abstract. The purpose of this study is to propose an intelligent extracting web content agent on the internet. This agent can automatically collect the web pages, generate several kinds of web pages templates, extract and simplify the appropriate web pages, provide the headlines for remote users via the remote devices. Via the proposed agent, the remote users can easily get the web information without large network bandwidth and read them by the remote devices conveniently.

1 Introduction

As the popular of the internet, we can get more and more information from the internet, especially in the World Wide Web (WWW). The web provides us a convenient media to disseminate information [6]. Due to the evolution of automatic generation of web pages, the numbers of web pages grow explosively [1, 2]. However, the useful information on the web is often accompanied by volumes of noises, e.g., banner advertisements, navigation bars, copyright notices, etc. Although such information items are functionally useful for human viewers and necessary for the web site owners, they often hamper automated information gathering and web data mining [7]. It is hard for us to get information by mobile devices.

Yi et al. [7] introduced an information based measure to determine which parts of the site style tree represent noises and which parts represent the main contents of the site and can improve the mining results significantly. Lin and Ho [5] proposed a new approach to discover informative contents from a set of tabular documents (or web pages) of a web site. Kao et al. [4] proposed a mechanism to automatically discover informative structure of a web site and contents of pages to reduce intra-site redundancy. However, if there is a simple mechanism that can reduce the web noises and provide the headlines for the remote users via remote devices, then it will be more effectively.

Due to the hyperlink often appeared in web pages, it is important to determine the web attributes. In [3], Fisher and Everson have shown that user's interests may be represented by a latent space model of a hyper-linked document collection and have given preliminary results demonstrating that it can be used to improving information retrieval tasks. As mentioned above, the hyperlink is important for extracting the web page contents.

Therefore, we proposed an intelligent extracting web content agent on the internet. Furthermore, in this agent, we proposed two webs content cluster factors for web

template generators, saying, hyper-link HTML tag distribution degree (HTDD), and none-text HTML tag distribution degree (NTTD), to cluster the web pages on demands. This agent not only can automatically collect the web pages and generate several kinds of web page templates from the internet webs, but also can extract and simplify the appropriate web pages for user preferences, provide the headlines for remote users via the remote devices. Via the proposed agent, the remote users can easily get the web information without large network bandwidth, and read them by the remote devices conveniently.

2   Framework of the Proposed Agent

In this section, we presented an intelligent extracting web content agent (IEWA) on the internet environment. The proposed agent is constructed between the web information service providers and the remote users, as shown in Fig. 1. IEWA not only can automatically generate the number of templates from the internet's web pages, but also can retrieve the appropriate web pages contents by meaningful templates.
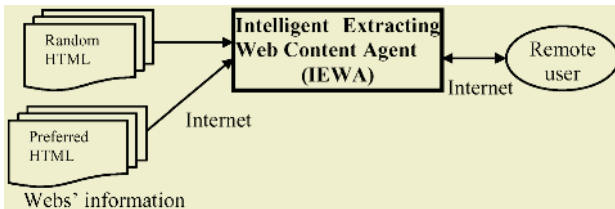
Fig. 1. Architecture of intelligent extracting web content agent

There are three modules in IEWA, saying, web template generator module (WTG), web content retrieving module (WCRM) and remote user interface module (RUI). In addition, there are three databases in the proposed agent, saying, template database (TDB), retrieved content database (RCDB) and user requirement database (URDB). The framework of the proposed agent is shown in Fig. 2.

The functions of these modules are illustrated as follows:

- ♦ Web template generator module (WTG): It can automatically collect the web pages via the internet, and generate their content templates.
- ♦ Web content retrieving module (WCRM): It can recognize the web content's templates of the web pages, and retrieve the headlines with contents from the appropriate web pages.
- ♦ Remote user interface module (RUI): It can provide the interfaces for remote users to set their preferring web sites and interesting content issues and provide the extracted web contents for remote users.
- ♦ Template database (TDB): It stores the web content templates which are generated from WTG.
- ♦ Retrieved content database (RCDB): It stores the retrieved web contents which are extracted from WCRM.
- ♦ User requirement database (URDB): It stores the user's preferring web sites and content issues for the RUI customizing dispatching the retrieved web contents.
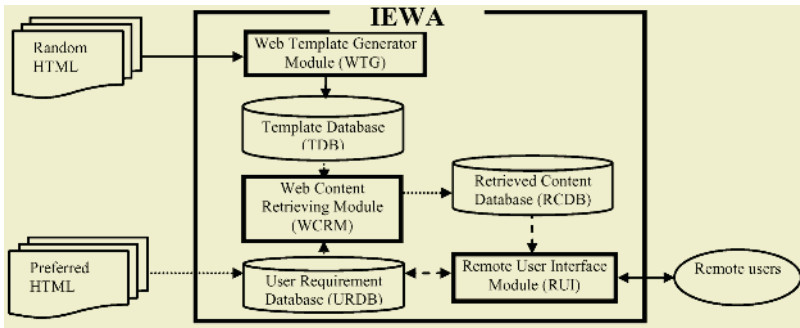
Fig. 2. Framework of the proposed agent (IEWA)

2.1 Web Template Generator Module

Web template generator module (WTG) can automatically collect the web pages via the internet, and generate their content templates. There are two components in this module, saying, web page spider (WPS) and fuzzy cluster engine (FCE). Also, there is a web page database (WPDB) in this module. The framework of this module is shown in Fig. 3
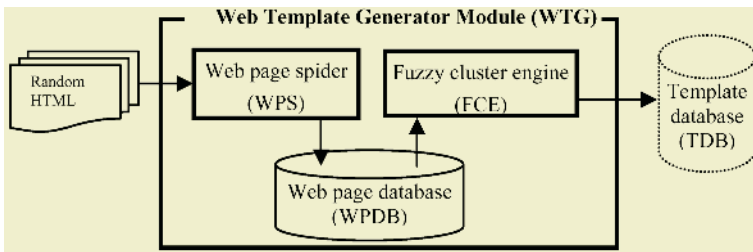
Fig. 3. Framework of web template generator module

The functions of these components are illustrated as follows:

♦ Web page spider (WPS): It can automatically collect all kinds of web pages via the internet, especially in HTML format web pages. Via the hyperlink, WPS can automatically trace the webs, and assign the serial number to the received page.
♦ Web page database (WPDB): It stores the web pages which are collected by WPS.
♦ Fuzzy cluster engine (FCE): It clusters the web pages by the two proposed weighting factors, saying, hyperlink HTML tag distribution degree (HTDD) and none-Text HTML tag distribution degree (NTTD), and automatically generates the different kinds of templates.
♦ We briefly introduced these two weighting factors as follows:
　(1) Hyperlink HTML tag distribution degree (HTDD): The hyperlink tag means that there are further messages or other pages related to this page. However, the higher hyperlinks, the lower web page contents are. Therefore, the HTDD factor can assist us to determine what kinds of web templates should be retrieved.

(2) None-text HTML tag distribution degree (NTTD): The webs have many kinds of multimedia objects, such as Flash, Java Script, image, etc. If multimedia occupied volumes of the web pages, then the text contents of the web might be insufficient. Therefore, the NTTD factor can assist us to find out the different kinds of web page templates.

(3) As stated in (1) and (2), it is the best way for us to extract contents only if both of HTDD and NTTD are lower.

Web page spider (WPS) can collect the web pages automatically, and store them in the web page database (WPDB). Fuzzy cluster engine (FCE) will proceed to cluster the web pages by the characteristics from the WPDB, and generate the web pages templates. Via this module, we can get several web pages templates and select the appropriate templates for web templates recognition.

2.2 Web Content Retrieving Module

Web content retrieving module (WCRM) can recognize the web contents templates of the web pages, and retrieve the contents from the applicable web pages. There are three components in this module, saying, web page collector (WPC), template recognizing (TR), and content extracting (CE), as shown in Fig. 4.
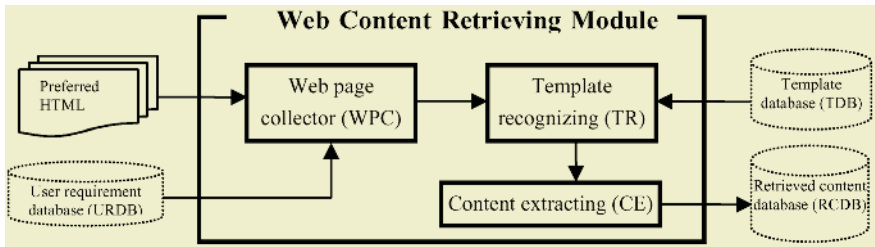
Fig. 4. Framework of web content retrieving module

The functions of these components are illustrated as follows:

♦ Web page collector (WPC): It can automatically collect the user preferred web page's html documents from the web pages, and send them to TR for template recognizing.

♦ Template recognizing (TR): TR will recognize the appropriate web templates of the preferred pages, and filter the redundant pages which are not suitable for extracting the web contents according to TDB.

♦ Content extracting (CE): It can extract, simplify the web pages contents, store the results in the RCDB, and provide the split web headlines with contents for the remote users easy to read via remote devices.

2.3 Remote User Interface Module

The remote user interface module not only can provide the interfaces for the remote users to set their preferring web sites and content issues, but also can provide the split web headlines for remote users via the remote devices.

3   Examples Implementation

In this section, we illustrate the implementation of the agent, and adapt twenty web sites to illustrate and analyze the operation of this agent as examples.

For the purpose of easy manipulation, crossing platform, and remote controlling, we have adopted Java Server Page (JSP) and Java Servlet written Web Server structure, as well, Java 2 Platform, Standard Edition, v 5.0 API Specification is utilized for constructing the proposed agent. We also apply Java 2 Micro Edition for implementing remote devices. This agent is constructed upon Tomcat 5 that employs browsers as its interface. The mentioned above are done with an Intel Celeron 2.4 GHz PC that O/S is Windows Professional and database is mysql-4.0.21.

Example 1
We take two web sites, saying, CNN and BBC, as examples to implement this agent. We use web page spider to collect the web pages from the web sites as shown in Table 1.

Table 1. Random webpage collection

Web site name	Number of web pages	Domain Name
CNN	408	www.cnn.com
BBC	391	www.bbc.co.uk

Via the FCE, we can get the four types of clusters, as shown in Table 2.

Table 2. Web pages clusters

Cluster	HTDD	NTTD	Number of web pages
#1	High	High	41
#2	High	Low	132
#3	Low	High	153
#4	Low	Low	473

The types of these clusters are shown in Fig. 5 – Fig. 8, respectively As mentioned in Section 2.1, and from Fig. 5 – Fig. 8, we can have that the collected web pages in the cluster#1 - #3 can be ignored, and the web pages in cluster#4 is appropriate for remote users to read.
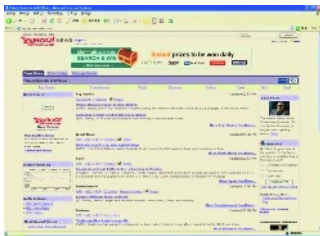
Fig. 5. Web type of cluster #1

Fig. 6. Web type of cluster #2
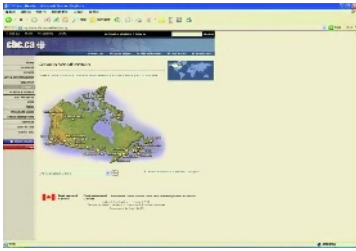
Fig. 7. Web type of cluster #3

Fig. 8. Web type of cluster #4

Example 2

If we select the preferring website to the agent via remote device, as shown in Fig. 9, then we can easily get the headlines by remote devices, as shown in Fig. 10. Also, we can drill down the headline, and get the contents if necessary, as shown in Fig.11.

Fig. 9. Querying interface
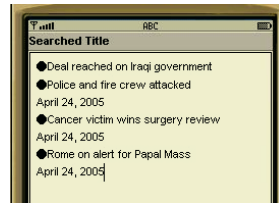
Fig. 10. Responding headlines
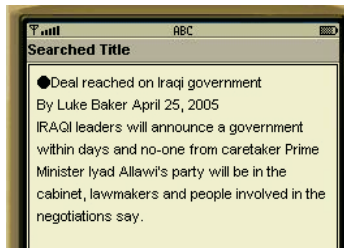
Fig. 11. Headline with content

4   Conclusion

In this study, we proposed an intelligent extracting web content agent on the internet. Furthermore, we proposed two webs content cluster factors for web templates generators in this agent, namely, hyperlink HTML tag distributed degree (HTDD) and none-text HTML tag distributed degree (NTTD). This agent can automatically collect the web pages, generate several kinds of web pages templates, extract and simplify the appropriate web pages, provide the headlines for remote users via the remote devices. Via the proposed agent, the remote users not only can read the web information by remote devices conveniently, but also can easily get the web information without large network bandwidth.

References

1. Broder A., Kumar R., Maghoul F., Raghavan, P., Rajagopalan S., Stata R., Tomkins A., Wiener J.: Graph structure in the Web, Proc. of the 9th World Wide Web Conference, (2000),309-320
2. Chen M.-S., Park J.-S., Yu P.-S.: Efficient Data Mining for Path Traversal Patterns, IEEE Transactions on Knowledge and Data Engineering, Vol. 10, No. 2, (1998), 209-221
3. Fisher M., Everson R.: Representing Interests as a Hyperlinked Document Collection, Conference on Knowledge Discovery in Data, Proc. of the 9th ACM SIGKDD,(2003), 378-385
4. Kao H.-Y., Lin S.-H., Ho J.-M., Chen M.-S.: Mining Web Informative Structures and Contents Based on Entropy Analysis, IEEE Transaction on Knowledge and Data engineering, Vol. 16, No. 1, (2004), 41-55
5. Lin S.-H., Ho J.-M.: Discovering Informative Content Blocks from Web Documents, Proc. of the 8th ACM SIGKDD, (2002), 588-593
6. Kobayashi M., Takeda K.: Information Retrieval on the Web, ACM Computing Surveys, Vol. 32, No. 2, (2000), 144-173
7. Yi L., Liu B., Li X.: Eliminating Noisy Information in Web Page for Data Mining, Proc. of the 9th ACM SIGKDD, (2003), 296-305

A Rendering System for Image Extraction from Music and Dance

Chihaya Watanabe and Hisao Shiizuka

Informatics, Kogakuin University Graduate School
Nishishinjuku 1-24-2, Shinjuku-ku, Tokyo 163-8677, Japan
`shiizuka@cc.kogakuin.ac.jp`

Abstract. The correspondence between precision dance movements and the accompanying musical arrangement is imperative to the reception of the performance. Whether the choreography is excellent or not becomes irrelevant if the performer's movements do not coincide with the score. This paper proposes an evaluation method based on an audience's impression of the correspondence between the music and the dancer's movements. This evaluation method is called "*Kansei* (Sensibility) Matching." For the matching method developed in this research, the term *Kansei* refers to a set of terms that express, verbally, the sentiment of music and movement.

1 Introduction

In recent years, the development of the Internet and various other applications has led to an increase in the number of scenes that utilize digital content, especially regarding the entertainment industry. Particular attention has been given to the automatic generation of choreographic movements to suit music. A great deal of research on body motion and body expression has been done in the fields of physical education and dance. Moreover, there is a great body of work on the relationships between a person's body movements and their mental state. However, research into movement, expression, and feeling is still developing, particularly in the field of creative dance. This paper proposes a system for automatically creating dance movements based on *Kansei* sensibility.

Dance can roughly be divided into two elements, "Music" and "Body Movement" (hereafter called dance). In arts like dance, it is considered that one "feels" rather than "contemplates" them; moreover, it is often difficult to present the intended image, or expression, clearly.

Music has a large part in contemporary society. There are even people, particularly among the younger generations, who may feel that it would be impossible to live without music. Music is often used to create moods. For instance, BGM (Background Music) is utilized in movies and television programs to convey the moods and images of particular scenes. Music helps the creator of the work to express the image that she/he originally intended. Dance is also a method of expression, albeit one that uses the body to transport the meaning or the image. However, when a person does not have a thorough grounding in dance techniques, their ability to express their intent successfully (this would entail the precise matching of their *Kansei* sensibility to dance, for the purpose of transmitting their *Kansei* feelings through body movements) would be greatly diminished. One of the most critical factors in advancing this re-

search field is the doubt surrounding how to extract the intended *Kansei* images, or features, from the music and dance pieces.

The basic impetuous behind the original conception to create a system to extract and match Kansei features from music and dance was that when the images formed by the music coincide with the images generated by the motions of the dancers, then the impression of the performance upon the audience members would be greatest. The crucial first step is that the *Kansei* information on music be matched to dance movements; then the movements of the dance can be choreographed. A rendering system that displays this choreography in visual actuality is described in this paper. Section-2 of this paper describes the system configuration utilized in this research. In Section-3, information relevant to the musical aspects of this paper and their relation to Kansei language is described. Section-4, explains the methodology behind the questionnaire used to investigate the correspondence between the musical aspects and Kansei language. Then in Section-5, the method, based on the discussions of Sections 2-4, of generating the dance images is described. The operation analysis of the dance and the extraction method of *Kansei* information are described in Section-6, respectively. In Section-7, the experimental results of the actual system are analyzed. Section-8 offers some concluding remarks upon the system.

2   System Configuration

In this research, a system that would be easy to use, fully automated, and available for use to visual entertainment industries was developed with the aim of presenting the expressions of dance in a manner that would be suitable to the images generated by the accompanying music. The proposed name of the system is Choreographer1.0. The configuration of the system is described below. The overall structure is shown in Figure 1, with its function being to present images of dance motions that will be suitable to music input based upon *Kansei* sensibility. The flow of the operating procedure of the system is as follows:

 (a) Music is input.
 (b) Assigning of *Kansei* words from music input
 (c) Correlation of dance motions to the *Kansei* words
 (d) Simultaneous output of dance image and music

The input medium for the music is limited to MIDI data for the current phase of this research. In the first stage, the user inputs the music, through MIDI data, for which she/he wants to generate a dance. The tempo information and tonality (through binary analysis: major key or a minor key) are obtained from the inputted MIDI data. The tempo information will change accordingly with the various inputted MIDI data. However, in the research conducted for this paper, it was decided to test only the tempos that are used most frequently. Within the database of the Choreography Part of the system (Figure 1), there are many partial dance images that can be associated with a *Kansei* word. The system will determine *Kansei* words based on the tempo information and the tonality obtained in the Music Part of the system. Then a dance image that corresponds to the Kansei word will be called. Matching the music input to a *Kansei* word, the system will extract, at random, partial images from the appropriate folders in the Choreography Part and generate a complete dance image. The

aim behind making the system utilize a random method of appropriate image selection and display was to insure variation between the completed images. The dance images were created with the motion capture CG software "Krops," this is described in Section-7. The final output is an image of the CG generated character dancing simultaneously with the inputted music.
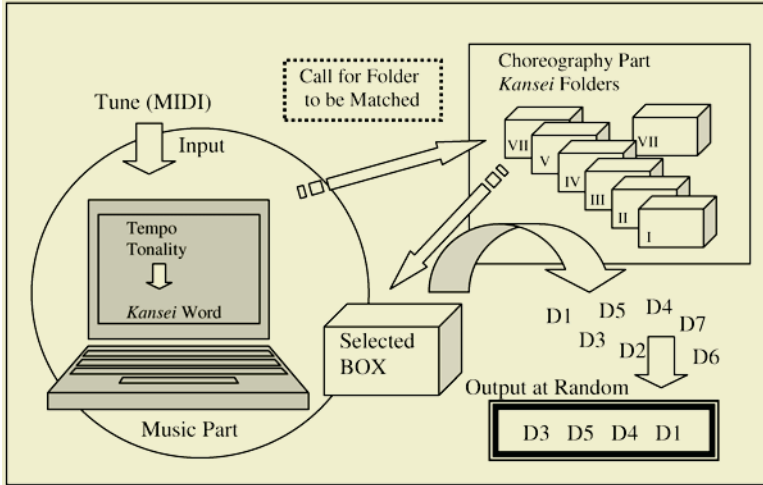
Fig. 1. System Configuration

3   Music Information and *Kansei* Language

Music creates its own images. The best way to extract images from a given piece of music would be by utilizing *Kansei* language. However, given the arbitrary nature of *Kansei* language, there are many different choices that can be made dependent upon each individual's sensibility. Therefore, in this research it is necessary to determine what particular *Kansei* words are usually associated with particular musical elements. The first step is to determine what about music forms images. Generally, it is held that major keys generate images of brightness and minor keys images of darkness. Tempo also influences the images created by a piece of music. In this research, in order to examine the influence of tempo, all of the sample tunes were in four-four time (four beats per measure). Then the research would yield data about the influences of various tempos upon musical images. The *Kansei* words utilized in this study, regarding the musical aspect, were taken from The Psychology Measurement Standard Collection II [1].

4   Assigning *Kansei* Words

4.1   Investigation by Questionnaire

The first step in this stage was to examine which *Kansei* words would be associated with the particular musical elements under investigation. A questionnaire based on the SD method was utilized to perform this analysis.

First, two tunes in minor keys and two tunes in major keys were prepared as samples. Each of the four tunes was formatted into the common tempos of 80, 120, 160, and 200, so that there were a total of 16 samples. It was believed that the musical image would be influenced by variations in tempo, and that varying the tempo of the same tune would be the best way to determine if this were the case. The test subjects listened to sample tunes and evaluated them in five stages using the prepared *Kansei* words. To avoid any contamination that might occur by having a test subject listen to the same tune twice, with the only difference being in tempo, a matrix of 4×4 was made.

A total of 40 subjects were tested. They were divided into four groups of ten. The age dispersion was from teens to seventies, with the majority of the subjects being university students in their early twenties.

Table 1. Content of Questionnaire

	Strong	Slight	Neutral	Slight	Strong	
Weak						Powerful
Lively						Lonely
Violent						Calm
Fresh						Gloomy
Cool						Warm
Happiness						Unhappiness

4.2   Principal Component Analysis

The principal component analysis (PCA) is a technique that uses a mathematical procedure to transform a number of possibly correlated variables into a smaller number of uncorrelated variables, which are termed principal components. The first principal component takes in to account as much of the variability of the data as it can, with each successive component accounting for as much of the remaining variability as possible [2]. Table 2 shows the amount of absorption for the three principal components of this research. Figure 2 and Figure 3 are scatter charts that show the variability between the 1st principal component and the 2nd principal component for *Kansei* words and tonality, respectively.

Figure 2 is a variable plot. Each feature of the first principal component and the second principal component can be analyzed from this variable plot. The *Kansei* words "unhappiness," "cold," and "gloomy" show high scores in the positive area of the horizontal axis (the first principal component), and the *Kansei* words "Fresh," "warm," and "happiness" show high scores in the negative direction. Therefore, it can be said that the first principal component will show the feature of warmth and the brightness of music from this feature. The *Kansei* words "lively," "powerful," and "violent" obtained high scores in the positive direction of the spindle of the second principal component, while "calm," "weak," and "lonely" obtained high scores in the negative direction. The second principal component analyzed the strength of the music.

Table 2. Amount of Absorption of the Principal Components

Total amount of information	96		
	First	Second	Third
Eigenvalue	8.43163	3.13722	0.19598
Absorption of information	67.45308	25.09775	1.56787
Ratio	0.70264	0.26143	0.01633
Cumulative Ratio	0.70264	0.9402	0.98040

Table 3. Coefficient of the Principal Components

Tune	First	Second	Third
Major 80	-2.5819	-1.3458	-0.5397
Major 120	-2.9627	0.6468	-0.4909
Major 160	-3.3631	-0.2506	0.6959
Major 200	-2.6563	1.0793	0.2882
Minor 80	2.7674	-3.1885	-0.2224
Minor 120	2.7746	-1.2932	0.4577
Minor 160	2.7333	2.5830	-0.4038
Minor 200	3.2886	1.7690	0.2150

Table 4. Variable Scorebook

	First	Second	Third
Powerful	0.21131	0.43156	0.19449
Weak	- 0.21131	- 0.43156	- 0.19449
Lively	- 0.24095	0.39397	- 0.07504
Lonely	0.24095	- 0.39397	0.07504
Violent	0.25352	0.36642	- 0.18647
Clam	- 0.25352	- 0.36642	0.18647
Fresh	- 0.32706	0.10234	0.54216
Gloomy	0.32706	- 0.10234	- 0.54216
Warm	- 0.33604	0.08941	- 0.30416
Cool	0.33604	- 0.08941	0.30416
Happiness	- 0.33635	0.07626	- 0.18791
Unhappiness	0.33635	- 0.07626	0.18791

After completing the principal component analysis, then the plots from the scatter charts, which showed the results for the *Kansei* words and the music elements, were coordinated to form Table 5.

Considering the above results it can be held that the musical elements of tone and tempo play a role in the image that the music generates. Thus, an image can be extracted from a piece of music by using the principal component analysis technique. The musical elements of a particular tune can then be analyzed and used to extract *Kansei* information on the particular tune. The system under development in this research will perform this function automatically upon input of a piece of music.
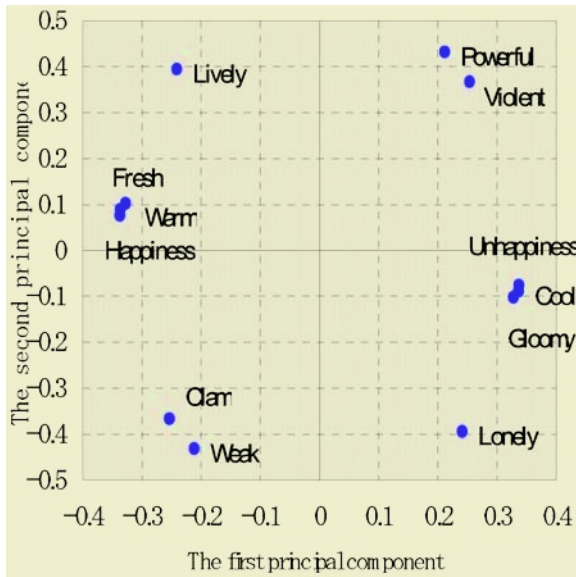
Fig. 2. The first principal component and the second principal component of *Kansei* words (p=12)
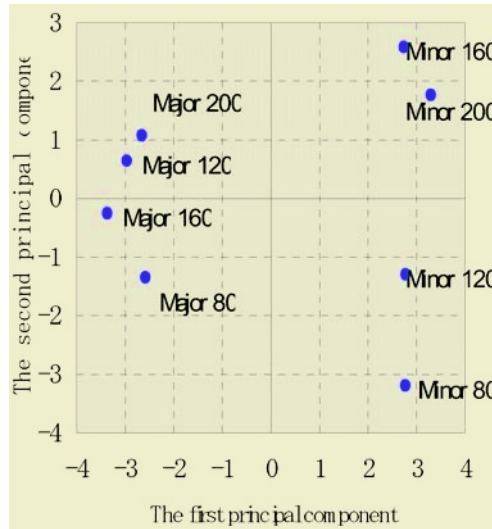
Fig. 3. The first principal component and the second principal component of the music sample (n=8)

Table 5. *Kansei* Words/Music Elements

Minor 200	Dark, violent image	Major 200	Happy image
Minor 160	Powerful, violent image	Major 160	Warm image
Minor 120	Dark, weak image	Major 120	Bright, warm image
Minor 80	Weak, lonely image	Major 80	Calm image

5 Dance Image and Animation Analysis

The body movements utilized in this research to generate the output dance patterns were based on the video The Basics of Ballet and the book The Basic Movements of Dance [4].

The motion capture software "Krops" was used to capture movement patterns (bitmap form; one image for each step) from existing sources (for example, video or DVD). This is a very useful method, even though no numerical analysis can be performed, because one can enter the movements into the system without having to actually go through the process of extrapolating information from a live performer.

We used a VHS video, mentioned above, to capture existing ballet patterns, and choreographic patterns were prepared. These patterns can be associated with *Kansei* words to form suitable images of the dance that will match the music. The bitmap forms, which were accumulated in the database, can then be processed and strung together to create the continuous movement of the animated dancer of the system. Moreover, the reproduction speed is set to the tempo of the particular music that was used as input. The dance movements are stored in the database according to their associative *Kansei* word, as shown in Figure 1. When music is input and analyzed and assigned a *Kansei* word, the corresponding dance images are automatically extracted from the database and are then reproduced at random.
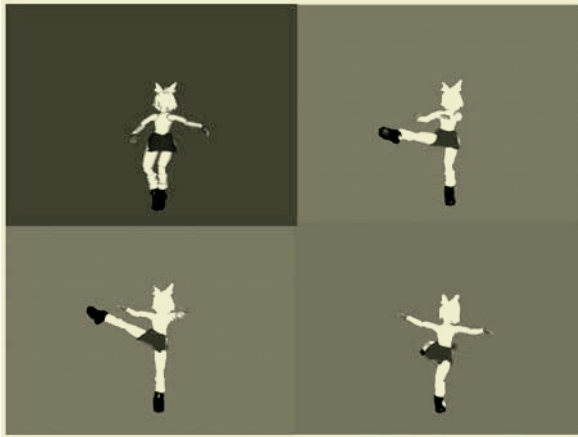
Fig. 4. Example of Dance Images

6 Dance and *Kansei* Word

In associating particular dance movements with *Kansei* words, the principal component system was not utilized. To make the associations necessary, the research team was aided by individuals with dance experience. While there is not a proven analytical method to accurately measure the correctness of these associations, it is expected that the input from experienced dancers is reliable. The information needed for the operation of the system was extracted by the following original technique.

In dance, the main parts of the body (arms, legs, head and torso) all play an important role. However, since the seat of balance rests in the torso, and since it is the central aspect, or base from which, all movements originate, it was decided to focus on examining only this area of the body during dance movements.

Table 6. An Example of the Data Analysis

Time/axis	Y	Time/axis	X
0:00	17	0:00	-12
0:11	22	0:11	-15
0:17	52	0:17	-15
1:06	4	1:06	-25
1:20	26	1:20	-18
1:28	15	1:28	-18
2:11	56	2:11	22
2:20	16	2:20	23
3:00	-3	3:00	-16

All the choreographic movements were analyzed using the aforementioned method and the data was then associated with the appropriate *Kansei* words.

The extraction examination of image words for dance was done via questionnaire. However, regarding dance, assigning *Kansei* words would be problematic since the spectators' choices may not necessarily relate to *Kansei* sensibility. For instance, should "amorous" and "graceful" be held as *Kansei* words when they may merely be the result of the spectators' actually witnessing the dance rather than a result of the particular aspects of the dance. Moreover, expressions such as "It is quiet," "It is lovely," or "It is elegant" are often used when referring to dance, but do not hold *Kansei* sensibility.

To eliminate the problems mention in the preceding paragraph it was decided that we would create a questionnaire using words from the Color Image Scale [5]. The language of this scale is also used to describe things such as shapes, patterns, and interior and fashion designs. The number of words contained in this scale is great; however, for this research, it was decided to limit the words to only those that would be appropriate regarding dance. 23 students who possessed dance experience aided in the selection of the words. The image words concerning dance are given in Table 7.

In the next stage of this research, a questionnaire based on the specified image words will be given. The results will then be correlated in order to draw associations between the dance image words and the *Kansei* words from the data on music. Through this process it is hoped that system being developed will not only be broadened but also improved for use in future applications. For instance, if the image word for dance and the *Kansei* word for music are the same, then there is only a one to one relationship and this will limit the output options for the system. However, if the image word for dance is different than the *Kansei* word for music that is stored in the database, the output options for the system will be broadened. Thus, it is hoped that the results of the new questionnaire will be able to allow the system to generate the choreographed output by associating two or more image types.

Table 7. Image Words Concerning Dance

Cute	Pretty	Sinuous	Flexible	Fresh
17	5	10	6	7
Quite	Freely	Smoothly	Pleasant	Funny
6	9	8	16	8
Lively	Brilliant	Showy	Kind	Delicate
7	14	5	6	5
Elegant	Moist	Refined	Youthful	Speedy
7	8	8	7	6
Quick	Modern	Sharp	Silent	Calm
5	7	7	7	8
Dandy	Classic	Wild	Strong	Gorgeous
7	6	8	5	9
Luxurious	Voluptuous	Bold	Violent	Natural
6	13	8	13	5
Elegant	Steady	Dynamic		
12	5	7	Average=4.518072289	

The structure of the system may look something like that expressed in Figure 5.
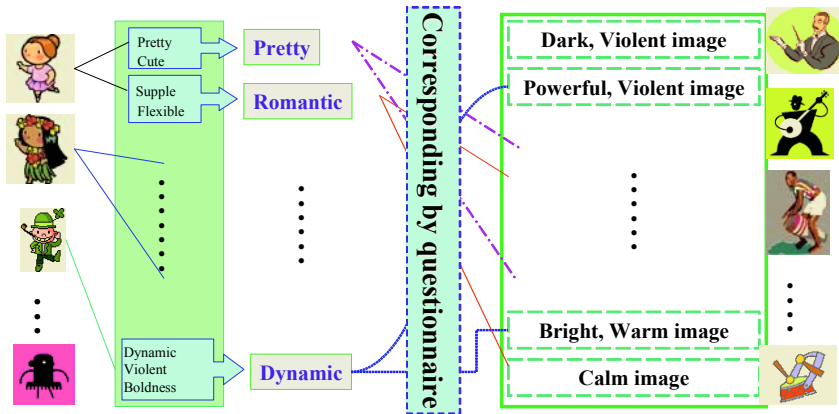
Fig. 5. Music/Dance Relationship

7   Experiment and Consideration

The system described here was designed for trial purposes. Experiments to verify the reliability of the system were conducted, and the accuracy was verified. When submitting your camera-ready manuscript to the volume editors, please make sure you include the following:

7.1   Experiment

The system was started; a test subject listened to the music sample in conjunction with watching the system's output image. No information about the music was disclosed to the test subject, who was then asked to complete the questionnaire below, Table 8.

The terms that were used in the questionnaire shown in Table 8, coincided with the terms that were used in Table 5, the matching of *Kansei* words and music elements. By doing this, it is possible to cross reference the two tables to determine whether or not there are any discrepancies between the image perceived from the music by the test subject (when listening to the music and viewing the dance output in conjunction) and the correlation between tempo/tone and *Kansei* words that was established and presented in Table 5.

7.2   Results and Consideration

Table 9 shows the results of this experiment. The test subjects' responses to the questionnaire about the image of the music matched the tempo/tone relationship to *Kansei* words presented in Table 5 rather accurately, 87.5% of the time. Thus, it has been established that the extraction of sensibility words from tempo and tone is roughly accurate. However, there are many other aspects of music that need to be considered in regard to the link between music image and *Kansei* words. This is a factor that will receive consideration in future aspects of this research. One area in particular that will be examined will focus on more subtle degrees of tempo variation.

Table 8. Experimental Questionnaire Form

- What image is the music?	
Weak, lonely image	Dark, weak image
Calm image	Bright, warm image
Powerful, violent image	Dark, violent image
Warm image	Happy image
- Is the dance image suitable for music?	
- It is.	
- It is not.	

Table 9. Experimental Results

The whole	Extraction of image information	Match level with animation
%	87.5	68.75

For the second question of the questionnaire, whether the dance image matched the music or not, a positive response was given 68.75% of the time. Under further analysis of these responses, it was found that for tempos where the BMP equaled 80, 120, 160, and 200 the test subjects felt that the dance image matched the music exactly. However, if the BMP equaled 190, for example, the test subjects felt that the dance image did not match the music. These types of responses were most likely due to the fact that when the BMP equaled 190, there was no exact dance image at that BMP (only dance images of 80, 120, 160 and 200 BMP exist in the system database): the output dance image would then be at the closest approximation, in this instance at 200, and therefore, the test subjects felt it did not match the music. To alleviate this problem, a method that would generate dance images that fit more closely with tempo should be given consideration.

8   Problems and Future Consideration

In this section, the problems encountered and future aspects of this research project, based on the results explained in Section7, are expressed. First of all, it will be necessary to match more accurately the rate of the dance movements with the tempo of the music. One possible solution to this problem is to make the playback speed of the existing bitmap forms of the animated dancer already in the database more flexible so that it can more closely match the input music tempo. It is also hoped that a closer association between the actual aspects of body motion in dance and the image words used for dance in Table 7 can be formed. The method that will most likely be used in order to achieve this closer relationship will be a questionnaire, based on some of the concepts outlined in "Body Expression-Body and Live Feeling It" [6].

In addition, it is held that by adding more image words that are directly related to dance, greater associations between dance perception, music perception, and *Kansei* sensibility can be achieved.

9   Conclusion

In this paper, the main focus was on creating an automated system for generating dance images that would coincide with the *Kansei* images of music input into the system. The first step was to assign *Kansei* words to particular music tempos and tones. This was achieved by processing the data gathered from the questionnaire with the principal component analysis. It was also discovered that there is a greater need for the speed of the output dance image to match the tempo of the music input. At this stage of the research, success has been achieved in inputting MIDI files into the system, generating continuous reproduced animation, and playing the reproduction simultaneously with the music.

In the future, it is hoped that the system will obtain a higher degree of accuracy in its extraction and matching capabilities, so that people will be able to enjoy music along with visual output. Also in the future, the possibilities of making the system portable will be examined.

References

1. H. Hori and F. Yoshida: Psychology Measurement Standard Collection , SAIENSU-SHA (2003)
2. K.Hasegawa: "Multivariate Analysis that Understands Really," Kyoritsu Shuppan Co., Ltd. (1998)
3. About Standard MIDI format: http://www2s.biglobe.ne.jp/~yyagi/material/smfspec.html
4. N.Tagawa et al: "The Basic Movements of Dance," KOTOKUSHA (1997)
5. S. Kobayashi: "Color Image Scale," Kodansha, (1990)
6. M. Shiba: "Body Expression-Body and Live Feeling It," Tokyo SHOSEKI (1993)

Trend of Fuzzy Multivariant Analysis in Management Engineering*

Junzo Watada

Waseda University
Graduate School of Information, Production & Systems
`junzow@osb.att.ne.jp`

Abstract. In 1965 L.A. Zadeh presented a fuzzy set. After then, fuzzy clustering method was proposed in early stage. L.A. Zadeh presented the concept of similarity for a fuzzy set in 1981. Ruspini proposed a clustering method based on fuzzy decomposition. Dunn and Bezdek wrote a clustering based on IDODATA algorithm in terms of fuzzy concept. Many methods for fuzzy multivariate analysis were presented .
For example, there are several approaches to multivariant analysis such as dynamic programming, M. Sugeno's measure, similarity concepts and clustering. M. Roubens and M.P. Windham discussed about the validity of clustering. Regarding hierarchical clustering, N. Osumi et al discussed it using fuzzy concepts. J. Watada et al proposed a heuristic hierarchical clustering which is employing similarity concepts of L.A. Zadeh's. There are many papers which discuss fuzzy clustering.
In 1979 H. Tanaka et al proposed fuzzy linear regression model. This development broke through fuzzy multivariate analysis because no methods are proposed except clustering and hierarchical clustering till then. There are also various applications of fuzzy multicariant analyses. Furthermore, various data analyses were proposed based on the fuzzy linear regression model. For example, J. Watada et al developed Fuzzy time-series analysis and I. Hayashi developed GMDH method.
This paper summarizes the 40 year history of Fuzzy Multivariate Methods.

Keyword: Fuzzy Multivariate Analysis, Fuzzy Statistics, Fuzzy Linear Programming.

1 Introduction

In 1965 L.A. Zadeh presented a fuzzy set. After then, fuzzy clustering method was proposed in early stage. L.A. Zadeh presented the concept of similarity for

---

Table 1. Fuzzy Multivariate Analysis.

Teaching Data Teaching Data	Method	Objective
No	(Fuzzy Decomposition)	Classification
	Fuzzy Clustering (Fuzzy)	Classification
	(Graph Theory)	Classification
	Fuzzy Hierarchical Clustering	Classification
	Fuzzy Dual Scaling	Meaning
		Quantitative method of multi-attribute
Yes	Fuzzy Discriminant of Qualitative Data	Meaning
		Quantitative method of multi-attribute
	Fuzzy Discriminant Analysis	Meaning

a fuzzy set in 1981. E.H. Ruspini proposed a clustering method [44-46] based on fuzzy decomposition. J.C. Dunn [14-17] and J.C. Bezdek [3-5,7] wrote a clustering based on IDODATA algorithm in terms of fuzzy concept. Many methods for fuzzy multivariate analysis were presented till now.

For example, there are several approaches to multivariant analysis such as dynamic programming [18], M. Sugeno's measure [38], similarity concepts and clustering [23,30,32]. M. Roubens [42-43] and M.P. Windham discussed about the validity of clustering. Regarding hierarchical clustering, N. Osumi et al discussing using fuzzy concepts [41]. J. Watada et al proposed a heuristic hierarchical clustering which is employing similarity concepts of L.A. Zadeh's. Papers [8,13, 16, 41] discuss fuzzy clustering.

In 1979 H. Tanaka et al proposed fuzzy linear regression model. This development broke through fuzzy multivariate analysis because no methods are proposed except clustering and hierarchical clustering till then. There are also various applications fuzzy multivariant analyses [51-75, 71]. Furthermore, various data analyses were proposed based on the fuzzy linear regression model. For example, J. Watada et al developed Fuzzy time-series analysis [29,66,74] and I. Hayashi developed GMDH method [28].

Analytical Hierarchy Process Model developed by T.L. Saaty [47, 49] was fuzzified by Buckley [9] and Ichihashi [80] in 1984.

The quantification model prevailed in Japan is to deal with qualitative data. The model consists of regression model, discriminant model, pattern classification and multi-dimensional scaling. J. Watada applied fuzzy probability and fuzzy statistic concepts in the quantification model [60-61, 68, 72, 73]. The model is employed in various field [60, 73].

On the other hand fuzzy integral based on fuzzy measure is widely used to analyze data [40].

Fuzzy discriminant analysis is developed by J. Watada et al [64]]

Tables 2 and 1 illustrate the fuzzy methods of multivariate analysis. These methods are classified according to whether the classification is given or not. In other words, we can classified according no indication of teaching classification and with indication of teaching classification.

Table 2. Fuzzy Multivariant Analysis.

Name of Method	Method of Fuzzy Set Theory
Fuzzy Clustering	Fuzzy Decomposition
	ISODATA
	Fuzzy Graph
Fuzzy Hierarchical Clustering	Fuzzy Similarity. Fuzzy Relation
Fuzzy Regression Analysis	Fuzzy Number
Fuzzy Time-series Analysis	Fuzzy Number
Fuzzy GMDH	Fuzzy Number
Fuzzy Discriminant Analysis	Fuzzy Number
Regression Analysis of Qualitative Data	Fuzzy Probability, Fuzzy Statistics
Discriminant Analysis of Qualitative Data	Fuzzy Probability, Fuzzy Statistics
Pattern Classification of Qualitative Data	Fuzzy Probability, Fuzzy Statistics
Multi Dimensional Scaling of Qualitative Data	Fuzzy Probability, Fuzzy Statistics
Fuzzy AHP	Fuzzy Number
Fuzzy Measure Analysis	Fuzzy Integral

2   Trend of Researches on Fuzzy Multivariant Analysis in Japan

In Japan, the researches on fuzzy sets and systems were started from early days in 1960s. The past researches can be summarized in the conference *Fuzzy Systems Symposiums* started from 1985 before these symposiums two research groups worked from quit early days before 1985. The number of presented papers and sessions are listed in Table 3. Table 3 shows also the number of papers relating multivariant analyses. Both trends are quite increasing. In Japan regarding the research on multivariate analyses, fuzzy clustering and fuzzy regression analyses are widely pursued. The reason is because J.C. Bezdeck and E.H. Raspini contributed on the fuzzy clustering from the very early days of Fuzzy Systems, and H. Tanaka contributed fuzzy regression analysis by inventing the formulation using linear programming in 1979.

Table 4 shows the distribution of the presented total papers in whole fuzzy systems symposiums from 1985 to 2004. We can include quantification method, discriminant analysis, time-series analysis with both regression and auto regression types, principal component analysis, fuzzy statistics, GMDH, AHP, fuzzy graph as well as fuzzy clustering and fuzzy regression analyses. Fuzzy regression analysis includes linguistic regression analysis proposed by J. Watada that deals with natural words. We can say that the research on fuzzy multivariant analyses is much fertile.

3   Concluding Remarks

In this paper we reported the trend of fuzzy multivariant analysis in this 40 years since the presentation of fuzzy sets by L.A. Zadeh. Also, we summarized the

Table 3. Fuzzy Systems Symposium in Japan.

No	Year	Place	No. of Sessions	No. of Papers	Multivariant Analysis
1	1985	Kyoto University	6	15	4
2	1986	Gakushuin University	19	35	3
3	1987	Osaka Electro-communication University	13	37	10
4	1988	Meiji University	18	60	9
5	1989	University of Marketing and Distribution Sciences	27	81	14
6	1990	Meiji	9	139	5
7	1991	Nagoya City Univeristy	56	167	15
8	1992	Hiroshima University	59	177	10
9	1993	Hokkaido University	72	199	10
10	1994	Kansai University	70	145	26
11	1995	Ryukyu University	71	251	23
12	1996	Waseda University	95	293	24
13	1997	Toyama University	87	282	31
14	1998	Nagara International Conference Hall	92	290	24
15	1999	Osaka Institute of Technology	98	259	15
16	2000	Akita Municipal Junior College of Arts and Crafts	73	192	15
17	2001	Nihon University	67	245	20
18	2002	Nagaya University	54	199	21
19	2003	Osaka Prefecture University	72	234	29
20	2004	Kyushu Institute of Technology	87	322	9

Table 4. Researches on Fuzzy Multivariate Analysis in Japan.

Method	Total No. of Papers
Clustering	95
Regression Analysis	58
Discriminant Analysis	11
Quantification Analysis	8
Graph	30
AHP	15
Time-series Analysis	14
Principal Component Analysis	7
GMDH	9
Multivariant Analysis	15
Fuzzy Statistics	23
Discribution	3
Measurement	3
Evaluation	15
Sociometry	6
DEA	9
Canonical Correlation Analysis	3
Variance Analysis	3
Total Number	327

trend of researches based on the number of papers presented at Fuzzy Systems Symposium started from 1985. The symposium also covers half of the whole history of Fuzzy systems. The research of fuzzy multivariate analyses is much fertile and developing still now.

References

1. J.C. Bezdek and J.C. Dunn: Optimal fuzzy partitions: a heuristic for estimating the parameters in a mixture of normal distributions, *IEEE Transaction on Computer*, 24, 835-838, 1975.
2. J.C. Bezdek and J.D. Harris: Fuzzy partitions and relations: An axiomatic basis for clustering, *Fuzzy Sets and Systems*, 1, 111-127, 1978.
3. J.C. Bezdek: Numerical taxonomy with fuzzy sets, *Journal of Mathematical Biology*, 1, 57-71, 1974.
4. J.C. Bezdek: Cluster validity with fuzzy sets, *Journal of Cybernetics*, 3, 3, 58-73, 1974.
5. J.C. Bezdek: A physical interpretation of fuzzy ISODATA, *IEEE Transaction on Systems, Man, and Cybernetics*, May, 1976.
6. J.C. Bezdek: Prototype classification and feature selection with fuzzy sets, *IEEE Transaction on Systems, Man, and Cybernetics*, SMC-7, 2, 1977.
7. J.C. Bezdek: A convergence theorem for the fuzzy ISODATA clustering algorithms, *IEEE Tran. on Pattern. Anal. Machine Intell.*, 2, 1-8, 1980.
8. J.C. Bezdek: *Pattern Recognition With Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.
9. J.J. Buckley: Fuzzy Hierarchical Analysis, *Fuzzy Sets and Systems*, 17, 233-247, 1985.
10. M.R. Civanlar and H.J. Trussell: Constructing membership functions using statistical data, *Fuzzy Sets and Systems*, 18, 1-13, 1986.
11. C.H. Coombs: *A Theory of Data*, Wiley, New York, 1964.
12. D. Dimitrescu: Hierarchical pattern classification, *Fuzzy Sets and Systems*, 28, 145-162, 1988.
13. D. Dubois and H. Prade: *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.
14. J.C. Dunn: A fuzzy relative of the ISODATA process and its use in detecting compact and well-separated clusters, *Journal of Cybernetics*, 3, 32-57, 1974.
15. J.C. Dunn: A graph theoretic analysis of pattern classification via Tamura's fuzzy Relation, *IEEE Transaction on Systems, Man, and Cybernetics*, May, 1974.
16. J.C. Dunn: Some recent investigations of a new fuzzy partition algorithm and its application to pattern classification problems, *Journal of Cybernetics*, 4, 32-57, 1974.
17. J.C. Dunn: Indices of partition fuzziness and the detection of clusters in large data sets, in: M. Gupta and G. Saridis, eds., *Fuzzy Automata and Decision Processes*, American Elsevier, New York, 1976.
18. A.O. Esogbue: Optimal clustering of fuzzy data via fuzzy dynamic programming, *Fuzzy Sets and Systems*, 18, 283-298, 1986.
19. I. Gitman and M.D. Levine: An algorithm for detecting unimodal fuzzy sets and its application as a clustering techniques, *IEEE Transaction on Computer*, C-19, 7, 1970.

20. I.R. Goodman: Fuzzy sets and equivalence classes of random sets, in: R.R. Yager, Ed., *Fuzzy Set and Possibility Theory: Recent Developments* , Pergamon, New York, 1982.
21. T. Gu and B. Dubuisson: Similarity of classes and fuzzy clustering, *Fuzzy Sets and Systems*, 34, 213-221, 1990.
22. D.E. Gustafson and W.C. Kessell: Fuzzy clustering with a fuzzy covariance matrix, in: M.M. Gupta, Ed., *Advances in Fuzzy Set Theory and Applications*, North-Holland, Amsterdam, 605-620, 1979.
23. J. Halpern: Set adjacency measures in fuzzy graphs, *Journal of Cybernetics*, 5, 4, 1975.
24. C. Hayashi: On the quantification of qualitative data from the mathematico-statistical point of view, *Annals of the institute of Stat. II*, 1950.
25. C. Hayashi: On the prediction of phenomena from qualitative data and the quantification of qualitative data from the mathematico-statistical point of view, *Annals of the institute of Stat. Math.*, 3, pp.69–98, 1952.
26. Chikio Hayashi: *Quantification Method*, Tokyo Keizai Shinposha 1972.
27. Chikio Hayashi, Isao Higuchi, Tsutomu Komazawa: *Statistics of Information Processing*, Sangyo Tosho, 1970.
28. I. Hayashi and H. Tanaka: The fuzzy GMDH algorithm by possibility models and its application, *Fuzzy Sets and Systems*, 36, 245-258, 1990.
29. B. Heshmaty and A. Kandel: Fuzzy linear regression and its applications to forecasting in uncertain environment, *Fuzzy Sets and Systems*, 15, 2, 159-191, 1985.
30. Hidetomo Ichihashi: "Fuzzy AHP," *Application of Fuzzy Systems* edited by Toshiro Terano, Kiyoji Asahi, Michio Sugeno, Ohm Publication, 1989.
31. J. Jacas: Similarity relations-The calculation of minimal generating families, *Fuzzy Sets and Systems*, 35, 151-162, 1990.
32. R.A. Jarvis and E.A. Patrick: Clustering using a similarity measure based on shared near neighbors, *IEEE Transaction on Computer*, C-22, 11, 1973.
33. A. Kandel: *Fuzzy Technique in Pattern recognition*, John Wiley & Sons, New York, 1982.
34. M. Kochen and A.N. Badre: On the precision of adjectives which denote fuzzy sets, *Journal of Cybernetics*, 4, 49-59, 1974.
35. P.J.M. van Laarhoven and W. Pedrycz: A fuzzy extension of Saaty's priority theory, *Fuzzy Sets and Systems*, 11, 229-241, 1983.
36. L. Lebart, A. Morineau and K.M. Warwick: *Multivariate Descriptive Statistical Analysis*, John Wiley & Sons, 1984.
37. E.T. Lee: An application of fuzzy sets to the classification of geometric figures and chromosome images, *Information Sciences*, 10, 1976.
38. K. Leszczynski, P. Penczek and W. Grochulski: Sugeno's fuzzy measure and fuzzy clustering, *Fuzzy Sets and Systems*, 15, 147-158, 1985.
39. G. Libert and M. Roubens: Non-metric fuzzy clustering algorithms and their cluster validity. in: M. Gupta and E. Sanchez, Eds., *Approximate Reasoning in Decision Analysis*, North-Holland, Amsterdam, 417-425, 1982
40. T.Onisawa, M.Sugeno, Y.Nishiwaki, H.Kawai and Y.Harima, Fuzzy measure analysis of public attitude towards the use of nuclear energy, *fuzzy Setsa and Systems*, 20, 259-289, 1986.
41. Noboru Oosumi: "Fuzzy Clustering," *Mathmatical Science*, 191, MAY, 34-41, 1979.
42. M. Roubens: Pattern classification problems and fuzzy sets, *Fuzzy Sets and Systems*, 1, 239-253, 1978.
43. M. Roubens: Fuzzy clustering algorithms and their cluster validity, *European J. Oper. Res.*, 10, 294-301, 1982.

44. E.H. Ruspini: A new approach to clustering, *Information and Control*, 15, 1969.
45. E.H. Ruspini: Numerical methods for fuzzy clustering, *Information Sciences*, 2, 1970.
46. E.H. Ruspini: New experimental results in fuzzy clustering, *Information Sciences*, 6, 1973.
47. T.L. Saaty: A scaling method for priorities in hierarchical structures, *J. Math. Psychology*, 15, 234-281, 1977.
48. T.L. Saaty: Exploring the interface between hierarchies, multiple objectives and fuzzy set*Fuzzy Sets and Systems*, 1, 57-68, 1978.
49. T.L. Saaty: *The Analytic Hierarchy Process*, McGraw-Hill, New York, 1980.
50. Michael Smithson: Fuzzy set theory and the social sciences: the scope for applications, *Fuzzy Sets and Systems*, 26, 1-21, 1988.
51. H. Tanaka, S. Uejima and K. Asai: Fuzzy linear regression model, *Int. Congress on Applied Systems Research and Cybernetics*, Acapulco, Mexico, 1980.
52. H. Tanaka, S. Uejima and K. Asai: Linear regression analysis with fuzzy model, *IEEE Trans. Systems Man Cybernetics*, 903-907, 1982.
53. H. Tanaka, T.Shimomura, J. Watada and K. Asai: Fuzzy linear regression analysis of the number of staff in local government, *FIP-84*, Kauai, Hawaii, July 1984.
54. Hideo Tanaka, Junzo Watada, Isao Hayashi: "Three Formulations of Fuzzy Linear Regression Analysis," *Trans, Society of Instrument and Control Engineering*, 22, 10, 1051-1057, 1986.
55. H. Tanaka: Fuzzy data analysis by possibilistic linear models, *fuzzy Sets and Systems*, 24, 363-375, 1987.
56. H. Tanaka and J. Watada: Possibilistic linear systems and their application to linear regression model, *Fuzzy Sets and Systems*, 27, 1-15, 1988.
57. H. Tanaka, H. Hayashi and J. Watada: Possibilistic linear regression for fuzzy data, *European J. of Operarional Research*, 40, 3, 389-396, 1989.
58. M. Tenenhaus and F.W. Young: An Analysis and Synthesis of Multiple Correspondence Analysis, Optimal Scaling, Dual Scaling, Homogeneity Analysis and other Methods for Quantifying Categorical Multivariate Data, *Psychometrics*, 50, 1, 91-119, 1985.
59. E. Trauwaert: On the meaning of Dunn's partition coefficient for fuzzy clusters, *Fuzzy Sets and Systems*, 25, 217-242, 1988.
60. Junzo Watada, Hideo Tanaka, Kiyoji Asai: Analysis of Purchasing Behavior by Fuzzy Quantification Analysis Type II," *Trans, Japan Institute of Management Engineering*, 9, 5, 385-392, 1981.
61. Junzo Watada, Hideo Tanaka, Kiyoji Asai: "Fuzzy Quantification Analysis Type II," *Behavioral Metrics*, 9, 2, 24-32, 1982.
62. Junzo Watada, Hideo Tanaka, Kiyoji Asai: "Identification of Group Selection Structure by Heuristic Method," *System and Control*, 26, 2, 24-32, 1982.
63. J. Watada, H. Tanaka and K. Asai: A heuristic method of hierarchical clustering for fuzzy intransitive relations, in: R.R. Yager ed. *Fuzzy Sets and Possibility Theory: Recent Developments*, Pergamon Press, 215-220, 1982.
64. Junzo Watada, Hideo Tanaka, Kiyoji Asai: "Fuzzy Discriminant Analysis," *Journal, Japan Institute of Management Engineering*, 34, 2, 126-131, 1983.
65. Junzo Watada, Hideo Tanaka, Kiyoji Asai: "Identification of Structure of Evaluation and Selection on Multivariate Attributes by Fuzzy Sets," *Journal of Systems and Control*, 27, 6, 403-409, 1983.
66. Junzo Watada, Hideo Tanaka, Kiyoji Asai: "Fuzzy Time-series Model and Its Application to Forecasting Problems," *Journal, Japan Institute of Management Engineering*, 34, 3, 180-186, 1983.

67. J. Watada: *Theory of fuzzy multivariate analysis and its applications*, Ph.D. Dissertation, University of Osaka Prefecture, 1983.
68. Junzo Watada, Hideo Tanaka, Kiyoji Asai: "Fuzzy Quantification Model Type I," *Japanese Journal of Behavioral Metrics*, 11, 1, 66-73, 1984.
69. J. Watada, K.S. Fu, and J.T.P. Yao: Damage assessment using fuzzy multivariant analysis, *Technical Report no CE-STR–84-4,School of Civil Eng.*, Purdue University, 1984.
70. J. Watada, H. Tanaka and K. Asai: Fuzzy discriminant analysis in fuzzy groups, *Fuzzy Sets and Systems*, 19, 261-271, 1986.
71. J. Watada, H. Tanaka and T. Shimomura: Identification of learning curve possibilistic concepts, In: W. Karwowski & A. Mital eds., *Applications of Fuzzy Sets Theory in Human Factors*, Elsvier Science Publisher, 143-168, 1986.
72. J. Watada and H. Tanaka: Fuzzy quantification methods, *2nd IFSA Congress*, Tokyo 1987.
73. Junzo Watada: "Fuzzy Quantification Method," *Introduction to Fuzzy Systems*, eds. by Toshiro Terano, Kiyoji Asai, Michio Sugeno, Ohm Publisher, 1987.
74. J. Watada, H. Tanaka and K. Asai: Analysis of time series data by possibilistic model, *Int. Workshop on Fuzzy Systems Applications*, Fukuoka, 20-24, 1988
75. Junzo Watada: Multivariate Attribute Decision Making, *Introduction to Applied Fuzzy Systems*, eds. by Toshiro Terano, Kiyoji Asai, Michio Sugeno, Ohm Publisher, 1989.

Kansei Engineering
for Comfortable Space Management⋆

Motoki Kohritani[1], Junzo Watada[1],
Hideyasu Hirano[2], and Naoyoshi Yubazaki[3]

[1] Waseda University I.P.S.
[2] University of O.E.H. Medical School
[3] Mycom Inc Co.

Abstract. It is difficult to define a comfortable space for each individual. It is partly because comfortness relates to many attributes that specify a space, partly because each individual has different preference, and also because even the same person changes his/her preference according health states, body conditions, working states and so on. Various parameters and attributes should be controlled in order to realize such a comfortable space according the data-base of past usages. Information obtained from human bodies such as temperature, blood pressure, α brainwave and etc. can be employed to adjust the space to the best condition.

The objective of this paper is to achieve a comfortable space in terms of Kansei engineering. It is required to realize our comfortable environment suitable to feelings and customers often have their individual preference so that they feel comfortable, too. Kansei Engineering is a technology to offer better products to the customers by employing human sensitivity.

1 Introduction

In a today's stress society, comfortable space and life are important to cure strong stress on us. It is in such human science as psychology, human factors, physiology, anthropology, hygiene and so on to study feeling of comfortness, and it is in such technology as mechanical engineering, architecture, control engineering, system engineering, living science, material science, acoustics, social welfare and design science to build such a comfortable space. It is not sufficient to combine these separated research results. We should create a new total technology to realize a comfortable space.

In this paper, the comfortable space means one where a person feels comfortable when he/she stays there although all people have a different feeling about

---

being comfortable. It is hard to realize a comfortable space for all people. Therefore, we should build a comfortable space from a viewpoint of each individual's. It is expected that a machine or a system can recognize and evaluate the state of a person from his/her behavior, voice and measurement of a living body by gathering and recognizing data automatically. If we can measure an electroencephalogram (brainwaves), heart beats, sweat, saliva, etc. by an instrument, the obtained information can be employed to realize a comfortable space.

The objective of this paper is to show the possibility that the measurement of human senses enables us to realize a comfortable space for any person even if the person changes his/her feeling of comfort according the change of the conditions [6,7,9]. In the research, we employ the values of features obtained from electrocardiogram that a test object has in a comfortable environment. Fuzzy control and neural network are used to adjust a space to the most comfortable one as possible.

Recently, "Comfort," "Senses," and "Kansei" are used in order to express products in their advertising. The reason is because today we seek products more fitting to ourselves than in the days when manufacturing few kinds of many products. Furthermore, products should fit to each individual consumer. Such aspects are expressed using word "feeling" of each individual consumer. That is rewritten using Japanese word "Kansei." So "Kansei" is required to employ one attribute for expressing a product. That is, we may provide more acceptable products for us when "Kansei" will be measured and enumerated, and engineering in designing and manufacturing products.

Additionally, wearable computers enable us to handle such information on line and real time. Such techniques change and evaluate our society drastically. These techniques are essential in a ubiquitous age. Taking a computer like dressing makes it easy to measure the information of sensitivity. After measuring the information from senses using a wearable computer, it can be transferred to a control unit through wireless waves. Such systems realize a comfortable space. In such a comfortable space, it is not necessary to indicate or order the system directly by a person.

2   Kansei Engineering

Kansei is a Japanese word that means total information of human senses. Kansei Engineering is "a technology, method or theory to translate human Kansei or Image to production of real things or to design of objects." Although it is vague and uncertain that a customer has an image or expectation about some product, Kansei Engineering is a technology to build such Kansei or vague image in product design in some way [1]. There are many measurements from a human body which can be used for the control of a space. In this paper, electrocaridographes obtained from a person are employed to adjust the environment of a space to the most comfortable that he/she feels. For instance, heart beats, sweet, saliva, etc. are another alternative.
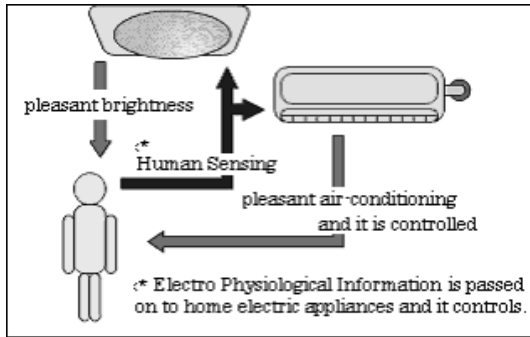
Fig. 1. Control based on Electro Physiological Information.

As mentioned above, Kansei Engineering is a theory and techniques which employ five senses and use such mechanism in developing products and services. If we successfully employ such sensitivity information in development of products, we can provide more personal-oriented products for individuals.

The most important issue in Kansei Engineering is has to measure "Kansei." There are methods of measuring human senses such as sensory test, measurement of a living body and cognitive measurement.

1) Sensory test:
   Generally, questionnaires can be included in the sensory test. Questionnaires have an influence from the objective patient or the environment. Therefore, we should manage such influences. Semantic differential, test electrocardiograph, and the rating scale method are included in this category.

2) Measurement of a living body:
   By measuring brainwaves, electrocardio- graph, heart rate, blood flow, sweating, the change of Kansei and emotion is detected. Furthermore, sensory test can result in the more accurate decision using measurement of a living body.

3) Cognitive measurement:
   Protocol analysis is widely employed in cognitive measurement. The method makes a test subject operate an test object such as a product, illustrate his impressions using words and at the same time his behavior is recorded using a videotape. After then, the information of the videotape and the descriptive impression of the test subject are scrutinized and analyzed.

2.1   Comfortable Space

There are various spaces in our environment. Also, the space of each room has different objectives even in a residential space. A tall building provides a different space from one in a house. Working space and space in a vehicle as well as residential space are included in an urban space. In future a universe space will be taken into consideration.

It is always essential that such a space provides us with comfort as well as safe and convenience. In the paper, information obtained from a living body is

employed to automate the adjustment of temperature, humidity, brightness and wind strength by home appliances. The technique realizes comfortable space for each individual in real time.

3   Electrocardiogram (ECG)

An electrocardiogram is a record that measures electric waves obtained from a heart muscle through electrodes placed on the surface skin of a chest, arms and legs. The wave causes the muscle to squeeze and pump blood from the heart. The activity of the heart is pursued by the Sino-atrial node in the right upper portion of the heart. The electric signal from the Sino-atrial node is transferred to a cardiac ventricle from an atrium and makes muscles electrically excite and shrink whose actions pump the blood to flow into blood vessels. The electrocardiogram records such electric actions of the heart as waves.
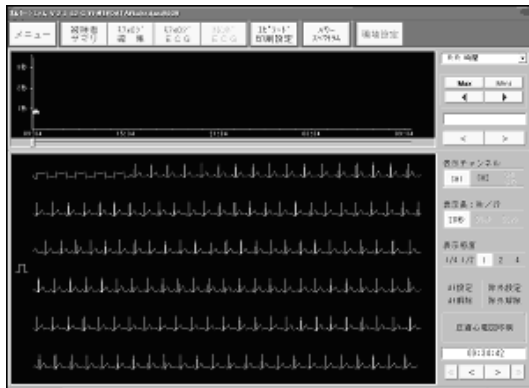
Fig. 2. HS1000Lite.

In this experiment a Wearable Holster Electrocardiogram is used to record long term, as the holster electrocardiogram can continuously measure and record electrocardiograph. The size of the holster electrocardiogram is light and small which does not weigh an test subject.

Figure 2 shows an example of a normal electrocardiograph. The electrocardiograph consists of five kinds of waves as PQRST. P waves show excites, QRS waves show excitement of entricles. T waves come out when the entricle cured its excitement. The excitement of an atrium can not be measured because their recover waves are quite small and overlapped with QRS waves.

In this experiment, R-R interval that is an interval between QRS waves is employed. One wave of QRS is produced when the heart gives one beat. That is, the R-R interval is one that the heart gives one beat.
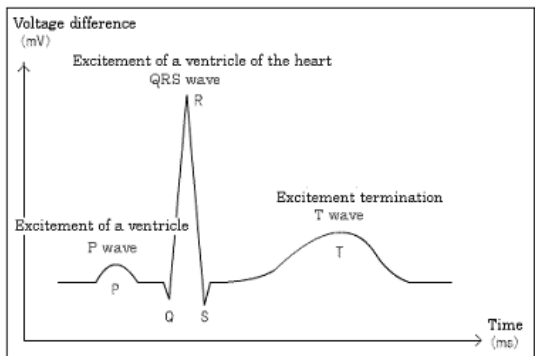
Fig. 3. Illustration of ECG.

Table 1. Names of Electrode Positions.

Code	Location of the leads	
Red	V5 [1]	1 Ch
Yellow	Intersection of right mid-clavicular line Sternum	CM_5 induce
Orange	xiphoid process of sternum	2 Ch
Blue	manubrium of sternum	NASA induce
Black	right V5: [2]	Earth

[1] V5: At the intersection of left anterior axillary line with a horizontal line through left V4 left V4: Intersection of left mid-clavicular line and fifth intercostal space

[2] right V5: At the intersection of right anterior axillary line with a horizontal line through right V4 right V4: Intersection of right mid-clavicular line and right fifth intercostal space

4   Information Abstract

4.1   Method of Abstracting Information

The experiment consists of a rest interval for 5 minutes and a work interval for 10 minutes. These intervals are repeated. In the work interval, test subjects are directed to typewrite some texts in order to give a stress on the test subjects. Figure 4 shows its results.

The horizontal axis shows QRS number numbered sequentially in time-series. The vertical axis shows voltage difference. The change between rest and work intervals can be detected clearly, as the figure of R-R intervals shows.

4.2   Results

Figure 4 shows the results.

- Work interval for 10 minutes is $1 \sim 725$ beats.
- Rest interval for 5 minutes is $726 \sim 1045$ beats.
- Work interval for 10 minutes is $1046 \sim 1743$ beats.
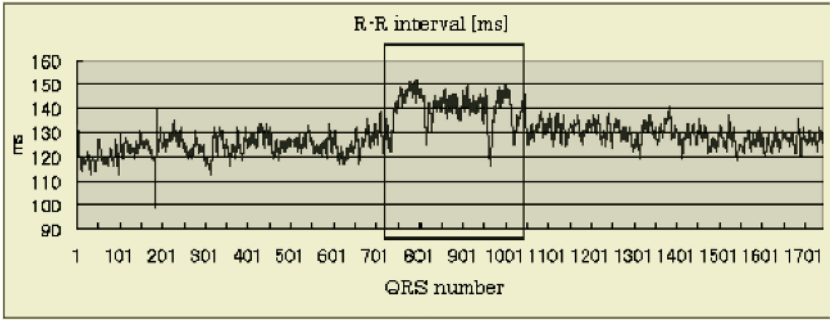
Fig. 4. R-R Interval.
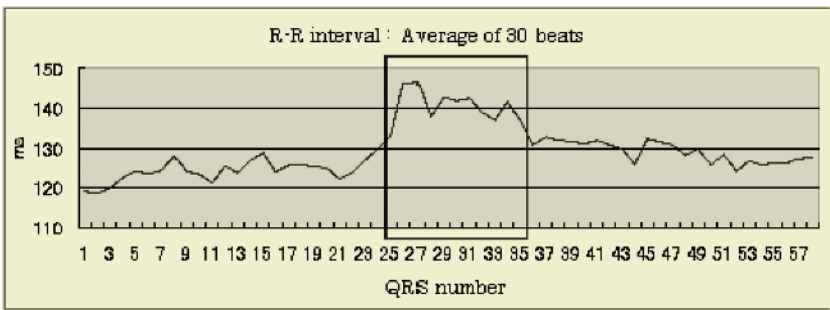
Fig. 5. R-R Interval (average of 30 beats).

We can recognize the difference between work and rest intervals in the figure. In order to clarify the difference, the average of 30 beats is figurized as in Figure 5.

The average of 30 beats resulted in the clear difference between work and rest intervals. Using this result, the threshold is decided and the result can be employed to control [1].

5   Conclusions

The temperature of the space is effectively controlled using the measurement of electrocardiographs of a test subject according his/her state such as in working or in rest. The proposed method can be extended to other parameters of a comfortable space such as humidity, smell and so on. And also we may use other parameters of a human body including an electrocardiogram, sweat analysis, saliva analysis and so on. This system enables us to control a human environment according the state of a person.

Nevertheless, there are several issues that should be solved. One issue is how the temperature should be set for plural persons in a room. One solution is to set the mean value for all persons there. In this case it might happen that all people would not be satisfied.

6   Instrument

Experimental instruments are provided by Fukuda ME Co.

- Holster electrocardiogram QR2100 which enables us to record electrocardiographs for 24 hours.
- HS1000Lite analyzing instument

References

1. Masato Takagi, Junzo Watada, Yubasaki Naoyosi, "Realization of Comfortable Space Based on Senses Information," IEEE SMC 2004, Session 77, pp. 6363-6364, (2004).
2. Junzo Watada, Masato Takagi, Naoyoshi Yubazaki, "Building a comfortable space by Human Senses," Proceedings, 17th Annual Conference of Biomedical Fuzzy Systems Association, pp.10-15 , Dec. 13-14, 2004 in Japanese.
3. Junzo Watada, Masato Takagi, Naoyoshi Yubazaki, Hideyasu Hirano, "Building a comfortable space by Human Senses," *Journal of Systems and Control Engineering Prat I*, Official Journal of Insitute of Mechanic ENgineering, 2005 to appear.

Fuzzy Logic Experience Model
in Human Resource Management*

Zhen Xu, Binheng Song, and Liang Chen

School of Software, Tsinghua University, Beijing, 100084, P.R. China
z-xu03@mails.tsinghua.edu.cn

Abstract. Job assignment is one of most important functions in human resource management. It presents a new model which optimizes the multi-objectives allocation problem by using fuzzy logic strategic. The fuzzy experience evaluation matrix indicates the score of certain employee on certain task. The values in the matrix are based on the employee's experience. Fuzzy appraisal decision-making method provides fuzzy synthesis appraisal matrix referring to individual experience value. Then Task-Arrange or Hungarian Algorithm provides the final solution with the help of our proposed experience matrix. As a numerical example demonstrated, it is helpful to make a realistic decision on human resource allocation under a dynamic environment for organizations.

1 Introduction

Experience Management (EM) is a new concept in the information systems field. It enables firms to effectively retain, share, and repurpose information in a way that improves client service, encourages client loyalty, and enhances business development. The management of organizations takes great advantages by using the concept of experience management. Since experience is a critical character to evaluate the quality of human resource, as well as an effective metric to estimate the employee's performance to the assigned tasks, we will discuss how to apply EM to Human Resource Management (HRM), especially to the problem of Human Resource Allocation in this paper.

Planning is the key feature of every information management system. In the process of tasks allocation, if the person who is in charge of planning certain task has previous experience for the company, he could assign weights and priority of goals quite accurately for a given situation. Considering the experience of the employee is very important factor for the managers to allocate upcoming projects.

Surveying the past research, there are two types of modeling about Human Resource Allocation. One is linear programming model [1]; the other is goal programming [2]. Linear programming is a single goal optimization technique; this is not the situation for a majority of firms with multiple goals. Goal programming can not deal with the organizational differentiation problems [3]. In addition, neither do the models studied before cover the consideration of employees' experience.

In the actual organizations, human resource allocation model not only identifies the current occupation situation of the employee, but also differentiates the performance value by employees' experience.

There are many algorithms on resource allocation. Distributed resource allocation [4] uses processors to solve it. But the number of processors is correlated with the number of tasks. The approaches proposed by [5-6] provide algorithms on task scheduling problem. But these algorithms do not consider the time constraints among tasks. The shortcomings limit the application to further research.

Fuzzy logic provides a way of modeling the uncertainty of natural language. It is a practical method in a decision-making paradigm whenever ambiguity and vagueness are present. And it can translate linguistic ambiguity into operational numbers for management systems design.

In this paper, a fuzzy logic experience model is proposed, which can represent the full relationship among human resource, tasks and time constraints. Fuzzy logic [7] is applied to the allocation problem. We build fuzzy sets for degree of job satisfaction and acquire individual capability matrixes. Fuzzy appraisal decision-making method provides fuzzy synthesis appraisal matrixes. Task-Arrange or Hungarian Algorithm [8] can obtain the best assignment solution based on the matrixes.

The paper is organized as follows: In section 2, we first present our basic model, and then we describe fuzzy concepts and define fuzzy membership sets in Section 3. We give the optimization target in Section 4. In section 5, our algorithm which achieves the optimization solution is proposed, and a numerical example is also given in section 6 to show how the algorithm works. Finally, we conclude our paper in section 7.

2   The Basic Model

The management of human resource includes three aspects. They are the control of human capital, the control of tasks and the control of time constraints.

2.1   Model on Controlling of Human Capital

1 The follow parameters are aimed at the whole human resource of the firm:

– Matrix $S = [s_{i,j} \mid i,j \in N, 1 \le i \le \mid R \mid, 1 \le j \le \mid T \mid, 0 \le s_{i,j} \le 1]$

In the above definition, S is an experience evaluation matrix for the employee, which is scored by their director depending on the experience of employee i on task j. $\mid R \mid$ represents the number of the employees; $\mid T \mid$ is the number of the tasks.

– group $_r^m$

The parameter r is the group name in a firm, and the parameter m represents the number of employees in the group.

2 For concrete human resource, it needs other parameters:

hr{group, ifFree, freeTimes, tasks, cost}

– group: The name of technique group the employee belongs to.
– ifFree: Indicate whether or not the employee is free currently. The value concludes "1"(occupied) and "0"(free).
– freeTimes: Certain times the employee can keeping free.
– tasks: Which task the employee has been occupied with.
– cost: The premium paid for accomplishing certain job.

2.2   Model on Controlling of Tasks

1 For the whole set of tasks:

- Matrix $T = [t_{i,j} \mid i,j \in N, 1 \le i,j \le \mid T \mid]$

The matrix T represents the time orders among tasks. If $t_{i,j} = 1$, task i is earlier than task j; otherwise, $t_{i,j} = 0$ means task i is not earlier than task j or there is no time order between the two tasks. We can also use PERT for distinguishing distinctly.

2 For concrete task, the parameters are defined here:

task{state, endTime, totalTime, needGroup, penaltyFunc, budget}

- state: Indicate whether or not the task has begun. The value concludes "1"(yes) and "0"(no).
- endTime: The end time of a task.
- totalTime: The total time of a task.
- needGroup: The group and the number of the employees needed by a task. Its data is in the form of g_r^m.
- penaltyFunc: The penalty function if the task is delayed.
- budget: The max cost of the task.

2.3   Model on Controlling Time

time{endTime, totalTime}

- endTime: The end time of a task.
- totalTime: The total time of a task.

3   Fuzzy Logic

Fuzzy Logic and Fuzzy Expert Systems define fuzzy logic as a "superset of conventional (Boolean) logic that has been extended to handle the concept of partial truths - truth values between completely true and completely false". In our paper, we apply the concept to our solution.

Define the Fuzzy Membership Sets We Use as Follows:

1 Fuzzy Sets for Degree of Job Satisfaction

Managers will build the fuzzy sets for degree of job satisfaction. The values of the sets are different aiming at different goals. But they are all based on the experience of the managers. For example, the values may be {A1(excellent), A2(good), A3(adequate)}. The relevant weight sets may be (0.6, 0.3, 0.1). The values of the elements in the weight sets are between "0" and "1". The sum in one set is usually "1". Also these values are commonly based on experience.

2 Experience Evaluation Matrixes of Tasks

Every task has its own experience evaluation matrix. The elements' values indicate the satisfaction probability if the usable employees occupy in certain task.

3 Individual Capability Matrixes

The matrixes are important for managers to make decision. In fact, they are acquired by converting the matrixes above. The elements' values indicate the different satisfaction probability if certain employee occupy in different tasks. For example:

$$R1_r = \begin{bmatrix} 0.5 & 0.6 & 0.7 \\ 0.4 & 0.2 & 0.2 \\ 0.1 & 0.2 & 0.1 \end{bmatrix}$$ provides the prediction on person r doing three tasks.

4 Fuzzy Synthesis Appraisal Matrixes

The matrixes are used in Fuzzy appraisal decision-making method which is an effective appraisal way to solve problems involving multi-factors. The matrixes are derived through the model $M(\cdot, +)$.

4 Optimization Goal of the Model

When directors implement the human resource allocation among tasks, they usually expect the optimized result in the following multi-objectives programming.

1 Minimize the Time in Executing the Entire Tasks

$$\text{Time}_{\text{Task}}(R^*) = T_{\text{Task}}(R^*) + \Delta T_{\text{Task}}(R^*)$$
$$\text{s.t: } 0 \leq \text{Time}_{\text{Task}}(R^*) \leq \text{Task.totalTime} \quad R^* \subset R$$

R is the set of the human resource, R^* is the subset of R. $T_{\text{Task}}(R^*)$ is the calculated time, which is necessary to perform task set. $\Delta T_{\text{Task}}(R^*)$ is the calculated waiting time, which appears only in the case of allocation of so-called busy resources. The time spent by the tasks must be less than the total time regulated beforehand.

2 Minimize the Whole Premium Cost to the Whole Employees

$$\text{Cost}_{\text{Task}}(R^*) = T_{\text{Task}}(R^*) \sum_{r \in R^*} \text{cost}$$
$$\text{s.t: } 0 \leq \text{Cost}_{\text{Task}}(R^*) \leq \text{Task.budget} \quad R^* \subset R$$

R is the set of the human resource, R^* is the subset of R. And the money cost by the employees must be less than the budget regulated beforehand.

To simplify the model, we only consider the human pay with no regard of cost of penaltyFunc of the tasks here.

5 Our Approach

The optimized solution of multi-objective-Function is to make multi-objective-Function be best satisfied. But these functions can not reach the best optimization at the same time. So we can only acquire the fuzzy optimization solution.

1 Definition 1:

If the number of employee certain task i needs is j, then dividing the task i into subtasks. And the number of the subtasks is j. The whole amount of the tasks will be added to j-1. The set of the subtasks replaces its parent task.

$$t_i = t_{i1} + t_{i2} + \ldots \ldots + t_{ij}$$

2 The Algorithm Can Be Described as Follows:

Suppose Task = {task$_1$, task$_2$……task$_n$}, Hr = {hr$_1$, hr$_2$….hr$_m$}

Algorithm:

1. Scheduling the current set of tasks according to matrix T.
   If the value of the element in T is "1", schedule the two tasks. Otherwise, calcu-

late the latest-begin time (task$_i$.endTIme − task$_i$.totalTime) of the other tasks. Scheduling them based on the time.

2. Calculating g$_r^m$ = task$_i$.needGroup. Acquire the sets of subtasks, and replace their parent task. Changing the value of the number of the new tasks set.

3. Finding out the sets of the usable human resource.
   If hr$_i$.ifFree == true then signing the employee.

4. Grouping the sighed employee according to the employee's attributes "group". The result is represented in the form of hr$_i^j$

5. Building the fuzzy sets for degree of job satisfaction A$_i$ and its weight Q$_A$ aiming at factor time.

6. Building the fuzzy sets for degree of job satisfaction B$_i$ and its weight Q$_B$ aiming at factor cost.

7. Building the experience evaluation matrix R$_{task}$ according to A$_i$ and Q$_A$.
   Matrix R$_{task}$ = [s$_{i,j}$ │ i,j ∈ N, 1 ≤ i ≤ │ R │ , 1 ≤ j ≤ │ A │ , 0 ≤ s$_{i,j}$ ≤ 1]
   │ A │ is the number of the elements of the fuzzy set A.

8. Building the experience evaluation matrix R$_{task}$ according to B$_i$ and Q$_B$.
   Matrix R$_{task}$ = [s$_{i,j}$ │ i,j ∈ N, 1 ≤ i ≤ │ R │ , 1 ≤ j ≤ │ B │ , 0 ≤ s$_{i,j}$ ≤ 1]
   │ B │ is the number of the elements of the fuzzy set B.

9. Building the individual capability matrix R$_{hr}$ considering the time factor.
   Matrix R$_{hr}$ = [s$_{i,j}$ │ i,j ∈ N, 1 ≤ i ≤ │ T │ , 1 ≤ j ≤ │ A │ , 0 ≤ s$_{i,j}$ ≤ 1]

10. Building the individual capability matrix R$_{hr}$ considering the cost factor.
    Matrix R$_{hr}$ = [s$_{i,j}$ │ i,j ∈ N, 1 ≤ i ≤ │ T │ , 1 ≤ j ≤ │ B │ , 0 ≤ s$_{i,j}$ ≤ 1]

11. Calculating the fuzzy synthesis appraisal matrix C by Using the model M(·,+) in the fuzzy mathematics.
    Matrix C$_{ij}$ = [c$_{i,j}$ │ i,j ∈ N, 1 ≤ i ≤ │ R │ , 1 ≤ j ≤ │ T │ , 0 ≤ c$_{i,j}$ ≤ 1]

12. Using Task-Arrange or Hungarian Algorithms to obtain the optimization result. Refer to [8] for more detailed description of the algorithms.

6 A Numerical Example

Conditions:

Task = {task1, task2, task3}
task1:{g$_1^2$, g$_2^2$} task2:{g$_1^1$, g$_3^2$} task3:{g$_2^2$, g$_3^2$}

After the fifth step we can get the data:

task = {task$_{11}$, task$_{12}$, task$_{13}$, task$_{14}$, task$_{21}$, task$_{22}$, task$_{23}$, task$_{31}$, task$_{32}$, task$_{33}$, task$_{34}$}
usable hr = {hr$_1^3$, hr$_2^4$, hr$_3^2$}

Then we know:

{hr$_{11}$, hr$_{12}$, hr$_{13}$} are allocated among {task$_{11}$, task$_{12}$, task$_{21}$}
{hr$_{21}$, hr$_{22}$, hr$_{23}$, hr$_{24}$} are allocated among {task$_{11}$, task$_{12}$, task$_{31}$, task$_{32}$}
{hr$_{31}$, hr$_{32}$} are allocated among {task$_{21}$, task$_{22}$, task$_{33}$, task$_{34}$}

First we consider the first allocation.

1. Building the fuzzy sets for degree of job satisfaction due to the goal of time.
   A = {A1(the fastest), A2(faster), A3(slowly)}, weight Q$_A$=(0.5, 0.3, 0.2)

2. Building the fuzzy sets for degree of job satisfaction due to the goal of cost.
   B= {B1(high), B2(medium), B3(low) }, weight Q$_B$=(0.6, 03, 0.1)

3. The weight of the two goals Q = {0.6, 0.4}
4. The experience evaluation matrixes of tasks aiming at employees and A are:

$$R1 = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.6 & 0.2 & 0.2 \end{bmatrix} \qquad R2 = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.5 & 0.4 & 0.1 \end{bmatrix} \qquad R3 = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.6 & 0.3 & 0.1 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$$

5. The individual capability matrixes of $\{hr_{11}, hr_{12}, hr_{13}\}$ about $\{task_{11}, task_{12}, task_{21}\}$ aiming at A are:

$$R1_{r11} = \begin{bmatrix} 0.5 & 0.6 & 0.7 \\ 0.4 & 0.2 & 0.2 \\ 0.1 & 0.2 & 0.1 \end{bmatrix} \quad R1_{r12} = \begin{bmatrix} 0.4 & 0.5 & 0.6 \\ 0.3 & 0.3 & 0.3 \\ 0.3 & 0.2 & 0.1 \end{bmatrix} \quad R1_{r13} = \begin{bmatrix} 0.6 & 0.5 & 0.5 \\ 0.2 & 0.4 & 0.3 \\ 0.2 & 0.1 & 0.2 \end{bmatrix}$$

6. The experience evaluation matrixes of tasks aiming at employees and B are:

$$R1 = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.4 & 0.3 & 0.3 \\ 0.5 & 0.3 & 0.2 \end{bmatrix} \qquad R2 = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.5 & 0.2 & 0.2 \\ 0.4 & 0.4 & 0.2 \end{bmatrix} \qquad R3 = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.6 & 0.3 & 0.1 \end{bmatrix}$$

7. The individual capability matrixes of $\{hr_{11}, hr_{12}, hr_{13}\}$ about $\{task_{11}, task_{12}, task_{21}\}$ aiming at B are:

$$R1_{r11} = \begin{bmatrix} 0.6 & 0.7 & 0.5 \\ 0.2 & 0.2 & 0.4 \\ 0.2 & 0.1 & 0.1 \end{bmatrix} \quad R1_{r12} = \begin{bmatrix} 0.4 & 0.5 & 0.4 \\ 0.3 & 0.2 & 0.3 \\ 0.3 & 0.2 & 0.3 \end{bmatrix} \quad R1_{r13} = \begin{bmatrix} 0.5 & 0.4 & 0.6 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.1 \end{bmatrix}$$

8. Using the model $M(\cdot, +)$ in the fuzzy mathematics, we can get the fuzzy synthesis appraisal matrix.

$$C = \begin{bmatrix} 0.1934 & 0.1526 & 0.1892 \\ 0.214 & 0.176 & 0.3662 \\ 0.205 & 0.8046 & 0.1936 \end{bmatrix}$$

9. Task-Arrange or Hungarian Algorithm gives the best allocation.
   $task_{11}$ is assigned to hr_{13}, $task_{12}$ is assigned to hr_{12}, $task_{21}$ is assigned to hr_{11}.
10. Also we can get the allocation result about other tasks and human resources.

7 Conclusions

In this paper we apply the concept of EM to Human Resource allocation problem. We construct a new model, and use fuzzy logic to get the experience matrix. It is suitable for the multi-objective assignment problem under a dynamic environment. By receiving input through the experience of the decision makers, the fuzzy logic and its solution methods provide better solutions in planning problems than the previous approaches. The new model is flexible, effective and efficient. In the solution, employee's experience is considered mostly. We can get the final human experience evaluation value via fuzzy logic. Fuzzy set theory is proper to be applied to human decision-making where there is a need to do modeling with imprecise reasoning processes or ambiguity. The new model and solution can help the managers make a more reasonable decision on their human resource allocation.

References

1. Grant, E. W, Jr., and Hendon, F. N., Jr.: An application of linear programming in hospital resource allocation. J. Health Care Market. 1: (1987) 69–72
2. Welling, P.: A Goal Programming Model for Human Resource Accounting in a CPA Firm. Accounting, Organizations and Society. 2(4): (1977) 307–316
3. Wikil Kwak.: Human Resource Allocation in a CPA Firm: A Fuzzy Set Approach. Review of Quantitative Finance and Accounting. 20: (2003) 277–290
4. CaiFen Wang.: The Distributed Algorithms of Optimum Distribution for the Limited Resource. Journal of Northwest Normal University. 30(1): (1994) 26–30
5. Keyser Thomas K, Davis Robert.: Distributed computing approaches toward manufacturing scheduling problems. J. IIE Transactions. 30 (4): (1998) 379–390
6. Vigo Daniele, Maniezzo Vittorio1.: A genetic/ tabu thresholding hybrid algorithm for the process allocation problem .J. Journal of Heuristics. 3 (2): (1997) 91–110
7. Omer, K, Andre de Korvin.: Utilizing Fuzzy Logic in Decision-Making: New Frontiers. Application of Fuzzy Sets and Theory of Evidence to Accounting II, Stamford, Connecticut: JAI Press Inc. (1998) 3–14
8. KaiCheng Lu.: Single-objective, multi-objectives and Integer programming. Beijing, Tsinghua University Press. (1999) 171–180
9. Liu, Y. H. and Y. Shi.: A Fuzzy Programming Approach for Solving a Multiple Criteria and Multiple Constraint Level Linear Programming Problem. Fuzzy Sets and Systems. 65: (1994) 117–124,
10. Zebda, A.: Fuzzy Set Theory and Behavioral Models for Decision Making under Ambiguity. Application of Fuzzy Sets and the Theory of Evidence to Accounting II, Stamford, Connecticut: JAI Press Inc. (1998) 15–27

Development of Business Rule Engine and Builder for Manufacture Process Productivity

Hojun Shin[1], Haeng-Kon Kim[2], and Boyeon Shim[1]

[1] CSPI, Inc. #201, Hyunsan B/D, 108-7, Yangjae-Dong,
Seocho-Ku, Seoul, 137-891, Rep. of Korea
{hjshin,byshim}@cspi.co.kr
[2] Dept. Computer Information & Communication Engineering, Catholic University of Daegu,
330 Kumrak 1ri, Hayangup, Kyungsan, Kyungbuk, 712-702, Rep. of Korea
hangkon@cu.ac.kr

Abstract. It is important that we produce a business rule and schedule in the manufacture process. The business rule is related at the dispatching to assign the material(lot) at the resource(equipment). The schedule is to plan a daily job scheduling of a total processes.

In this paper, we develop the builder and editor for business rule flow and the workflow engine for dispatching and scheduling. We propose systemical development process using component and UML(Unified Modeling Language) technology to analysis, design and develop. The development process is an attempt to consider all of the best features of existing domain requirements. Also, we define rule format of XML(eXtended Markup Language) for mapping between a rule engine and builder. The builder generates XML from a rule and parses to save the rule at the repository. Finally we describe how these concepts may assist in increasing the efficiency and productivity in business application and manufacture process.

Keywords: Business Rule, Business Rule Engine & Builder, Component Based Architecture, Component Based Development, UML

1 Introduction

Recently the software lifecycle is getting shorter and web service for paradigm of next generation information technology is more focused on while e-business model has developed very rapidly. Therefore, the development of software which is more functionable, various, stable software, the key of business domain[1].

We propose to organize the description of the business rule engine and builder by using several model, each one addressing one specific set of concerns. The models propose systemical development process using component and UML technology to analysis, design and develop. The development process is an attempt to consider all of the best features of existing domain requirements. The rule builder generates XML format file from a rule and parses to save the rule at the repository. Finally we describe how these concepts may assist in increasing the efficiency and productivity in business application and manufacture process..

2   Related Works

2.1   Business Rule

In November 1995, in its seminal "GUIDE Business Rules Project, Final Report," the Business Rules Group defined a business rule as a "… statement that defines or constrains some aspect of the business … [that is] intended to assert business structure, or to control or influence the behavior of the business. [A business rule] cannot be broken down or decomposed further into more detailed business rules…. If reduced any further, there would be loss of important information about the business."[2]

In general terms, "defines" refers to terms and facts, while "constrains" refers to rules. Together, these three aspects—terms, facts, and rules—represent the core idea of business rules. More precisely, rules in declarative form should represent a primitive component of a business IT architecture - they should be coequal with processes and other components that cannot or should not be derived from anything else. In formal terms, such rules correspond to the if-then or implication connective of predicate calculus; in practical terms, "declarative" means that they should contain no hidden semantics.

Statements given in procedural programming languages therefore do not qualify as an acceptable form of business rule expression because there is hidden meaning—order dependence—in the sequence in which they appear. These languages also typically contain other scaffolding that has nothing to do with actual business logic. Rules expressed in proper declarative form never have any hidden meaning.

The fundamental tenets of the business rule approach, which focuses on building IT architectures, include the following:

- Rules should be written and made explicit.
- Rules should be expressed in plain language.
- Rules can exist independent of procedures and workflows.
- Rules should build on facts, and facts should build on concepts as represented by terms.
- Rules should guide or influence behavior in desired ways.
- Rules should be motivated by identifiable and important business factors.
- Rules should be accessible to authorized parties.
- Rules should be single-sourced.
- Rules should be specified directly by those people who have relevant knowledge.
- Rules must be managed[3].

3   Development Process for Business Rule Engine and Builder

As we suggest CBD reference architecture, component development process based architecture is a set of activities and associated results, which lead to the production of a component as shown in figure 1. These may involve the development of component from specification by using UML model. Here, our main concern is the specification workflow. In addition, we consider systemical development process using UML and model technology to analyze, design, and develop business rule engine and builder. The domain analysis specification, design model, implemented component, which are produced though the process, are stored in the repository.

3.1   Requirements Identification Phase

The requirement of rule engine and builder should be first identified in desired business system. The primary property of business rule is able to analyze after that the description for specific business rule domain and the sorts of essential properties should be understood. At the same time, it is very important to consider weather the requirement, which is already defined, is corresponding to rule type in reference architecture and what business concept is focused on.

For the business domain analysis, UML approach is used. Diagrams used in problem domain analysis are using case diagram. Use case diagram is a diagram that shows a set of use cases and actors and their relationships. It supports the behavior of a system by modeling static aspects of a system. Also, domain analysis is presented on entire domain concept and scenario using activity diagram. Requirement analysis is defined through use case diagram, and use case description.
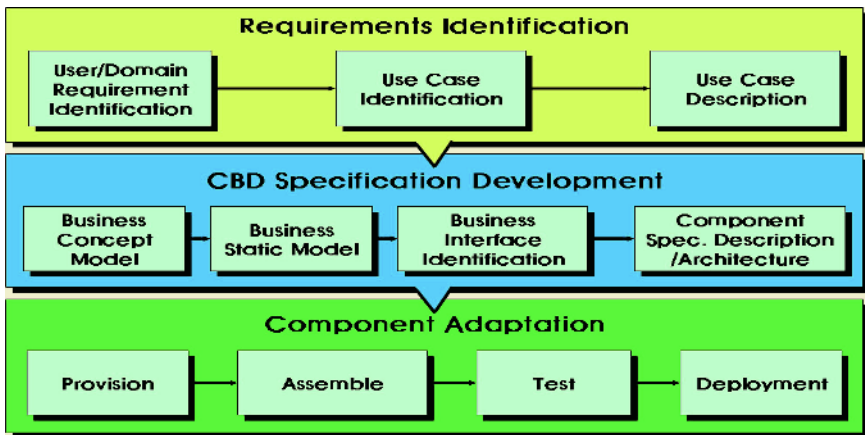
Fig. 1. Development Process for Business Rule Engine and Builder

3.2   CBD Specification Development Phase

We have attempted summarize the process tasks into the four stages: business concept model, business static model, business interface identification and component spec description.

The specification development takes as its input from requirements a use case model. It also uses information about existing software assets, such as legacy systems, packages, and databases, and technical constrains, such as use of particular architectures or tools. It generates a set of component specifications and component architecture. The component specifications include the interface specifications they support or depend on, and the component architecture shows how the components interact with each other.

The identified information based on component users and performance must be provided in specification form for integration. Also, this information can be provided and acquired by producer, consumer and agent in interoperating system. The information of component design and development, and also functional and non-functional

information must be provided by producer, and business rule must provide the commercial information with this. This information is the important standard for choice and the ground for reuse to acquire the component.

3.3 Component Adaptation Phase

These outputs are used in the provisioning phase to determine what components to build or buy, in the assembly phase as an input to test scripts.

The provisioning phase ensures that the necessary components are made available, either by building them form scratch, buying them from a third party, or reusing, integrating, mining, or otherwise modifying an existing component or other software. The provisioning phase also includes unit testing the component prior to assembly.

The assembly phase takes all the components and puts them together with existing software assets and a suitable user interface to form an application that meets the business need. The application is passed to the test phase for system and user acceptance testing.
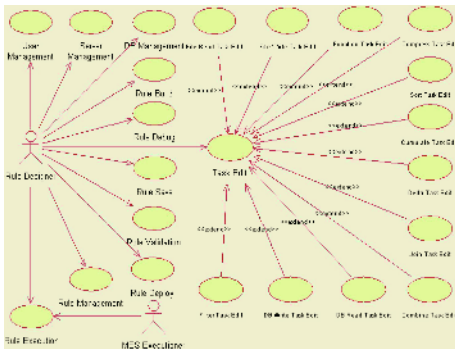
Fig. 2. Use case diagram   **Fig. 3.** Parser Static Model
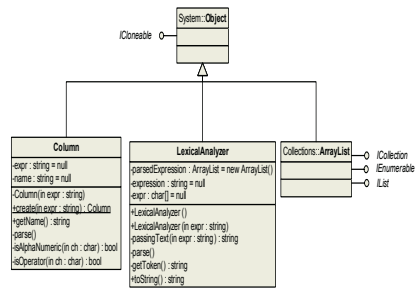
4   Development of Business Rule Engine and Builder

4.1   Use Case Identification

Use Case View uses to reflect user's requirement in business rule domain. The business rule engine and builder design and execute the scheduling rules or dispatching rules that need in a field. The business rule engine and builder design and run scheduling rules or dispatching rules that need in field. Specially, there are very important role for productivity elevation in product line as support system of MES(Manufacturing Execution System). Figure 2 shows the use case diagram for the workflow rule design and execution in rule designer side.

4.2   Business Static Model

Figure 3 show parser static model that lexical analyzer the designed or builded business rule. Figure 4 show source generator static model that generates the business rule source of XML format. Also, it create java class file for business rule execution.

4.3   Component Model

We analyze a component inside action of the system which tries to develop in a component modle. We present a component architected to the foundation the content. Figure 5 presents a component based architected for business rule engine.
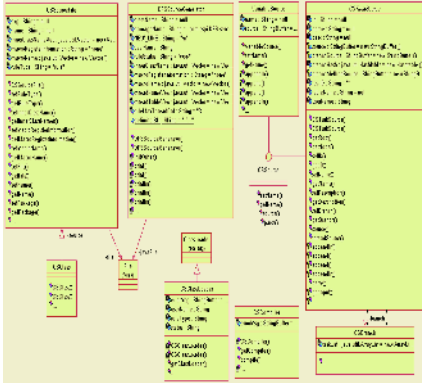
Fig. 4. Source Generation Static Model
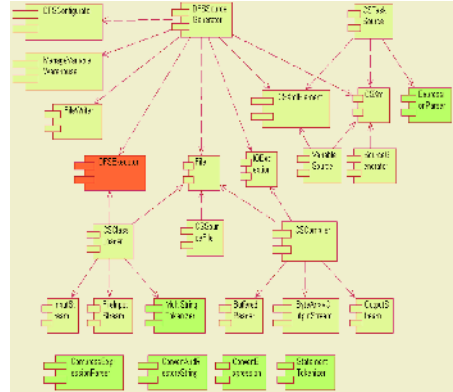
Fig. 5. Component Model

Fig. 6. Architecture Model
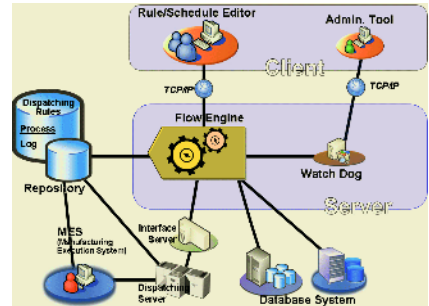
Fig. 7. Application Architecture Model

4.4   Architecture Model

We propose architecture that consists of 5 layers with figure 6. These contents are as following. Client layer takes charge the interaction of user and system. Client layer receives user's request and delivers to the system. As the result, client responses and handles in system to user.

Presentation layer connects with client layer and business layer. It verifies user request in client layer changes to correct form and delivers from business layer. Also, there communicate response that handle in Business layer to client layer.

Business: Business layer handles business according to business rule. Business layer verifies data in presentation layer. It makes data as new result by applying suitable business rule data and existing data, storing or delivering to outside system.

Integration layer connects with integration a system and various outside resource. Integration layer can be connect to database of resources layer as well as that con-

nects with system that know physically with outside resources and do abstract to work that happen in interaction such as data transmission and conversion .

Figure 7 application architecture which expresses in a context diagram. There are data access service and communication service for data access and communication between client and server. A rule engine is located in the server. Figure 8 and 9 present that workflow engine executed and user interface designed for workflow rule.

Fig. 8. Business Rule Engine

Fig. 9. Business Rule Builder

5   Conclusion and Future Works

In this paper, we propose the systemical development process using component and UML technology to analysis, design and develop. The logical view of business rule domain defined 5 hierarchies laying stress on function. The hierarchies are client layer, presentation layer, business layer, integration layer and resource layer. We develop the builder and engine for business rule flow and the workflow engine for dispatching and scheduling. As a result, implementation architecture presented whole system structure and user interface that is developed. In case of increase of new investment or process equipment, huge investment is required. But, can expect productivity improvement introducing administration methodology or software that can raise efficiency of existing production line. Also, bring result that can enjoy big effect into small expense. We are also going to study on the various case study and evaluation for architecture efficiency.

References

1. Barbara von Halle, Business Rules Applied: Building Better Systems Using the Business Rule Approach, Wiley Computer Publishing, 2002.
2. Business Rules Group, "GUIDE Business Rules Project, Final Report," http://www.businessrulesgroup.org, 1995.
3. From R.G. Ross, Principles of the Business Rule Approach, Addison-Wesley, 2003.
4. C.J. Date, What Not How: The Business Rule Approach to Application Development, Addison-Wesley, 2000.
5. George T. Heineman and William T. Councill, "Component-Based Software Engineering", Addison-Wesley, 2001.
6. Claudio Riva, Petri Selonen, Tarja Systä and Jianli Xu, "UML-based Reverse Engineering and Model Analysis Approaches for Software Architecture Maintenance", Proceedings of the 20th IEEE International Conference on Software Maintenance, pp.50-59, 2004.

7. Hugo Vidal Teixeira, Regina M. Braga and Cláudia M. L. Werner, "Model-Based Generation of Business Component Architectures", Proceedings of 30th EUROMICRO Conference, pp. 176-183, 2004.
8. Guoqing Xu, Zongyuan Yang and Haitao Huang, "Article abstracts with full text online: A basic model for components implementation of software architecture", ACM SIGSOFT Software Engineering Notes, Volume 29, pp. 1- 11, Issue 5, 2004.
9. Pearl Brereton and David Budgen, "Component-Based Systems: A Classification of Issues", IEEE Computer, Vol 33, No 11, pp. 54-62, 2000.
10. Jia Yu and Rajkumar Buyya, "A Novel Architecture for Realizing Grid Workflow using Tuple Spaces", Proceedings of Fifth IEEE/ACM International Workshop on Grid Computing, pp. 119-128, 2004.

Automatic Detection of Failure Patterns
Using Data Mining

Youngshin Han, Junghee Kim, and Chilgee Lee

School of Information and Communication Engineering SungKyunKwan University 300,
Chunchun-dong, jangan-gu, Suwon,Kyunggi-do 440-746, S. Korea
yshan@ece.skku.ac.kr

Abstract. In the semiconductor manufacturing, yield enhancement is an important issue. It is ideal to prevent all failures. However, when a failure occurs, it is important to quickly specify the cause stage and take countermeasures. Reviewing wafer level and composite lot level yield patterns has always been an effective way of identifying yield inhibitors and driving process improvement. This process is very time consuming and as such generally occurs only when the overall yield of a device has dropped significantly enough to warrant investigation. The automatic method of failure pattern extraction from fail bit map provides reduced time to analysis and facilitates yield enhancement. The automatic method of failure pattern extraction from fail bit map provides reduced time to analysis and facilitates yield enhancement. This paper describes the techniques to automatically recognize and classifies a failure pattern using a fail bit map, a new simple schema which facilitates the failure analysis.

1 Introduction

Over the years, memory fail bit maps (FBM) have become an important diagnosis tool for testing memory devices. The predominant failure modes of a memory device could quickly be discovered by their characteristic fail patterns on a map. However, as memory devices become denser, it is very difficult to analyze fail bit maps and to estimate the cause [1].

A pattern recognition algorithm suitable for efficient analysis of FBM has been reported [2]. The method enables an engineer to discover the primary failure modes of a memory device over a large population. The statistical analysis method for FBM has been reported [3]. However, these methods do not lead to the cause estimation.

In order to determine problematic wafer processing steps rapidly during mass production, the automatic memory failure analysis using an expert system with a memory tester [4] and the system that utilizes the computers between a conventional memory tester and a yield management database system [5] has been proposed.

The automatic method of failure pattern extraction from fail bit map provides reduced time to analysis and facilitates yield enhancement. This paper describes the techniques to automatically recognize and classifies a failure pattern using a fail bit map, a new simple schema which facilitates the failure analysis.

2 Analyzing Process of Defect Types

Fig.1. shows Electrical Die Sorting (EDS) process. The EDS process aims at improving the whole process of FAB by verifying each die's operation in a wafer state and

analyzing the causes of failure in fail dies. Through the assembly process costs can be saved in the next process by only flowing pass dies that were collected through the assembly process, and improve the quality of those dies through a laser repair process. EDS process is consisted of 3 different processes; PRE-Laser, Laser-repair, POST Laser. DC test, back grinding, Inking can be added. After EDS process, the process goes in to ASS'Y, Package assembly process. Fig.2. shows fail analysis flow.
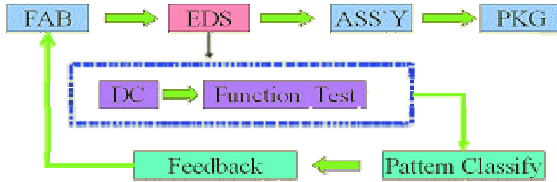
Fig. 1. Electrical Die Sorting process
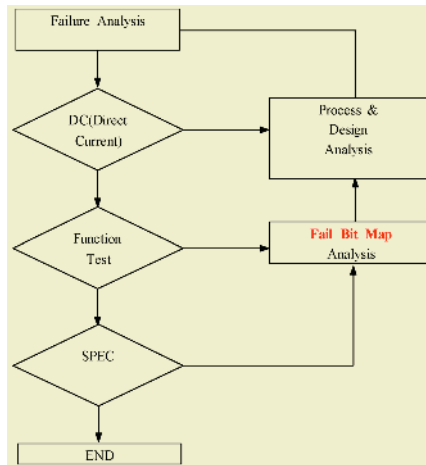
Fig. 2. Fail analysis flow

Types of defect can be classified into two categories: chip level and wafer level. Engineers can define various defects in a chip by classifying them in accordance with specific cells' defects contained in each chip. Chips can be categorized into two groups according to the result of tests: Bin3, ICC failure, and Bin4, function failure. In function test, engineers carry out a wide range of tests on chips. Although engineers can detect failure in chips through the tests, they need another classification according to each failure type as the cause of defect can be changed by the influence of failed block. Defects in wafers can be classified according to the distribution type of defective chips on a wafer, and a method finding the cause of defect can be changed by area character [6][7].

2.1   Definition of a Chip's Defect Types

In type classification, the definition of a chip's defect type should have priority. To categorize the types of defects according to a chip's unit block, this paper uses failed

characters of a device in defining each case of a defect in chips with similar cause. Table 1 shows four types of defects according to a chip's unit block: random, block, area, and special, which consider type and shape of failed test. Special type can be differently defined according to the different character of devices. Random type also can be subdivided into three types according to the distribution density: random, group, and column row type. Finally, block type can be categorized into partial block type, block type, CSL (Column Select Line), and NWE (New Word Line Enable) according to the distribution characters of vertical and horizontal failure.

Table 1. Types of defects according to a chip's unit block

Macro classification	Chip type
Random	Random
	Cluster
	Col/Row
Block	Partial block
	Block
Special	New Word Line Enable
	Column Select Line

2.2   Input Data

Data is being recorded in the process of Failed bit map > chip > unit block.

- Failed bit map: comprised of many chips
- Chip: comprised of X * Y unit block
- Unit block: quantized grade

Chip level information such as wafer ID, total chip number, x-y coordination, Bin value, chip DC, and function test result and grade of each chip's unit block are used as data for wafer information of a failed bit map.

2.3   Data Preprocess

In order to classify the types of chip defects specifically, this paper adjusted the grade of unit block as following: The data in Table 2 are partially adjusted for technical security reasons.

Table 2. Unit block Grade

Grade	Fail bit Count	Fail pattern
0	0	Good
1	1	Pattern A
2	2~10	Pattern B
3	101~300	Pattern C
4	301~500	Pattern D
5	501~1000	Pattern E
6	1001~10000	Pattern F
7	10001~20000	Pattern G
8	20001~	Pattern H

3   Definition Tree Classifier Categorizing Types of Chip Defects

Definition tree classifier has been used in various fields of type understanding. In particular, it is most effective in categorizing and understanding complex classification fields, type classes with a high-degree dimension, and a huge amount of type classes [8]. In Comparison with single stage classifier, definition tree classifier can divide the whole gamut of complex decision-making process into simple narrow decision-making processes through many levels of the tree quickly.

- Classifying Engine: SAS E-Miner decision tree
- Classifying algorithms: C4.5 (Entropy index)
- Input Data: variants generated for character extraction
- Data size for research: samples of typical defective chips were collected from 10% of defective chips found in a daily EDS test samples.

3.1   An Example of Tree Classification Result

As the depth of the tree depth deepens, each node is divided into a group of separate type. Some nodes are mixed with types with high similarity. In this regard, decision rules for each type were generated through adjustment of the tree depth and analysis of some mixed nodes with similarity. Fig.3 shows an example of tree classification result.

Fig. 3. Classification Tree depth = 3

4   Experiment and Results

4.1   Experiment Environment

- Subject Device: 000M DRAM
- Subject Wafer: 26 wf of a daily test quantity 00Lot 000 wf
- Application Method: Automated analysis by program generated algorithms
- Test Environment: OS Window 2000, compiler vc++ 6.0, wafer per 2.5 sec

4.2   Experiment Result

More than 90% of the subjects were specifically classified. In the case of mixed types, error rate was high and improvements in flexibility, extension rate, and false analysis rate were required.

Table 3. Summary of the result of Classification algorithms

Classification algorithm	Sample	Accuracy
Chip type analysis	26*200=5200 chip BIN4 600 chips	91.6%
Wafer type analysis	26 wafers	96%

5   Conclusion

This paper proposed ways to automatically detect and classify types of defective chips used in yield analysis by using failed bit map data that occurs in EDS test of semiconductor fabrication. By applying this method to analyze defects, we automated conventional manual operation and built a system effective in analyzing defects. In addition, we shortened more than 80% of working hours spent in analyzing defects and improved accuracy by detailing data in fail bit unit of a chip compared with conventional failure classifying system using bin information of each chip contained in a wafer. Moreover, this method can be directly applied to conventional products without any increase or loss in test time. By establishing swift and accurate analysis of the cause of failure, we succeeded in improving wafer yield's quality dramatically.

References

1. Koji Nakamae, Atsushi Itoh, Hiromu Fujioka, "Fail pattern classification and analysis system of memory fail bit maps", Modeling and Simulation of Microsystems, 2001
2. M. Faucher, "Pattern recognition of bit fail maps," Proc. International Test Conference, pp.460-463,1983
3. S.P. Cunningham and S. MacKinnon, "Statistical methods for visual defect metrology," IEEE Transaction on Semiconductor Manufacturing, 11,1, pp.48-53, 1998
4. T. Tsujide,et al., "Automatic memory analysis using an expert system with a memory tester/analyszer," Proc. International Test Conference, pp. 184-189, 1993.
5. M. Sugimoto, H. Hamada, and T. Hamada, "High speed processing system for meroty failure analysis," NEC Technical Jaurnal, 50, 6, pp.11-15, 1997.
6. Carlos Ortega, "Human based knowledge for the probe failure pattern classification with the use of a backpropagation neural network. Application on submicron linear technologies", IEEE/SEMI Advanced Semiconductor Manufacturing Conference, 1998
7. Michael W. Cresswell, "A Directed-Graph Classifier of Semiconductor Wafer-Test Patterns", IEEE Transaction on Semiconductor Manufacturing, 1992
8. Seiji Ishikawa, "Fail Bit Analysis System Semiconductor Memory Wafers", Japan International Electronics Manufacturing Technology Symposium, 1993

Logic Frameworks for Components Integration Process

Haeng-Kon Kim[1] and Deok-Soo Han[2]

[1] Department of Computer Information & Communication Engineering
Catholic University of Daegu, Kyungbuk, 712-702, South Korea
hangkon@cu.ackr
[2] Dept. Computer Engineering, Korea Third Military Academy
Kyungbuk,770-849, South Korea
dshan123@hanmail.net

Abstract. This research developed the CBD(Component Based Development) logic frameworks for the Integration of components and a tool which supports this. The Integration of components becomes necessary during the process of reusing or assembling components, and this is because the interface of the component is, in many cases, different than the component the developer wishes to assemble. After the iterations of specification modification and verification in terms of knowledge acquisition activities, Logic frameworks are correctly formed. Occasionally, additional attributes may need to be defined in accordance to new requirements. Consequently, the process for component Integration is crucial for the reuse and assembly of components. In order to support the Integration of components, this research proposes an Integration technique dependent upon binary component Integration techniques and Integration components. In addition, a support tool was developed to support an effective Integration process.

1 Introduction

New expectations are being gathered within the software engineering sector as software development through reuse by Component based Software Development (CBSD) is becoming recognized as a legitimate methodology. When viewing components simply, components are utilized through the application plug-in method, and can be reused (Black-Box Reuse) through the 'As-is" method [1,2,12]. However, as shown through numerous research, "As-is Reuse' is extremely difficult, and is reported to have little practicality within most application development processes. In other words, in order to be a component suitable to the requirements of an application, it is common that a component will require Integrations, whatever Integrations it may be. Especially, in instances when the component was not self-developed, but a developed component was purchased or rented, from the perspective of the component user, it is difficult to guarantee that the particular component will satisfy all requirements necessary within the application system. The problem which occurs most frequently during assembly is that the interface for the purchased component is different from the requirements for the interface of the current application development, or the interface is the same, but the functionalities actually required by the behaviors are different [4,5,6]. The commons methods used to overcome such problems are 1) changing the existing component or system upon integration, 2) modify the purchased component, 3) placing an adaptor between the existing component and system[1,13,15].

Within this research, the component Integration technique which supports black-box reuse of components is defined, and carries an objective of developing a tool which supports this.

2 Necessity for Integration Techniques and Definition of Existing Integration Techniques

This chapter defines the scale of the Integration techniques as well as the requirements which the Integration techniques must satisfy in this research.

2.1 The Necessity of Component Integration

The three problems naturally faced during the reuse process of components are Integrations and interface evolution. Let us first understand the issue with an example which uses components. Suppose there are two applications, A and B, which use components. The components for each application were purchased from different corporations. The class interface for the application is defined as different even class provides same service.

Another problem arises from the fact that the interface is continuously evolving. Evolution is something which occurs because interface corrections of components are necessary during the development process. Changes may occur through newly formed requirements or to correct errors which have been discovered. Additions to the interface accompany additions in new implementations, and if a new implementation is added, this requires for all the components to be re-compiled. These types of changes are unnecessary activities to the codes which have been already compiled, and interface evolution is an operation which also requires additional burden.

The final need for Integrations is in cases when the semantic specification for components changes. For example, one of the Java Components, Date, provides a functionality which expresses time and date. Days are defined with a scale of 1 through 31 and months are defined with a scale of 0 through 11, where 0 through 11 equals months. However, there will be applications in which such method of processing dates is not applicable.

2.2 Requirements of Component Integration Technique

The following is the definition of the requirements which the component Integration technique must satisfy [10,11].

 (1) Transparent: The fact that a component is transparent means that the developer who used the original component before Integration or the developers who used the components after Integration all cannot know about the Integration. Not only this, components not requiring Integration must be able to be accessed without additional effort.
 (2) Black-Box: Although before reusing components, software developers must fully understanding the functionalities of components before reusing components, if possible, it should be possible with only minimal information. These types of requirements presuppose that information regarding how the original component

was implemented is not utilized during the Integration process of components. In other words, Integration of components is limited through only the interface.

(3) Composable: Adapted components must also be able to conform with other components, and adapted components must also be able to be adapted.

(4) Homogeneous: The code which used the previous components must also be able to use the adapted components.

(5) Ignorant: Existing components must have no previous knowledge regarding Integrations.

(6) Identity: Existing components must be independently identifiable irrelevant to Integrations. In other words, they must be able to be used in different Integrations in the future.

(7) Architectural focus: The specification regarding the application architecture which includes the existing components and adapted components must be provided. Here, the specification for the existing components and the adapted components must be different.

3 Component Integration Technique

This research proposes that the technique for the Integration of components which is the Integration of binary codes in which the binary component Integration technique is applied to the EJB components.

3.1 Binary Component Integration Technique

This research has applied a technique utilizing the BCA technique [8,9] for binary component Integrations in consideration of these issues. This is an attempt to make Integrations possible even in instances when the source code cannot be directly accessed. Although the java class loader defined within EJB cannot changed and ultimately wraps the new component with the adapted binary codes, It differs from the existing method in that this is not a wrapping of the contents of the adapted component itself, but rather an addition of only the Integration specification to the existing component, which is wrapped together with the Integration specification processor. If repeated Integrations occur, only the Integration specification becomes added and the problem in which the size of the component grows with continuous wrapping becomes resolved.
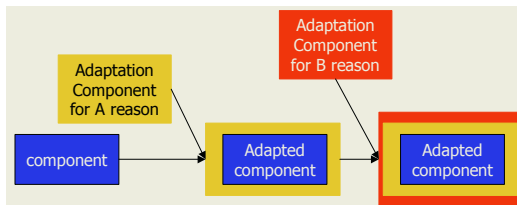
Fig. 1. Integration Technique based on Integration Specification

When examining the example in Fig. 1, in order to adapt existing components, the Integration specification is written, and is wrapped together is the adaptor which

processes this. When requiring additional Integrations in the future, new Integration components can be created by modifying just the existing Integration specification. In addition, an adaptor which has been wrapped already due to an earlier Integration can continuously be reused, and becomes a concept of adapting just the Integration specification when adapting an already adapted component.

Information necessary for Integration and the specification must be written up to execute the proposed binary component Integration. If the component re-user writes up the component contents which need to be changed or added in accordance to the defined specification format, this is converted into a component specification format recognized by the component writer tool, and creates a normal component through the component Integration tool. The component adaptor developed through this research supports all Integrations for adding new attributes, interfaces, and methods to and renaming of existing components. To achieve this, the existing component is decompiled to acquire the source code, and creates a newly adapted component by uniting with the component Integration specification written with relation to the source code. Through this base, the design diagram for a component adaptor has bee defined as <Fig 2>.
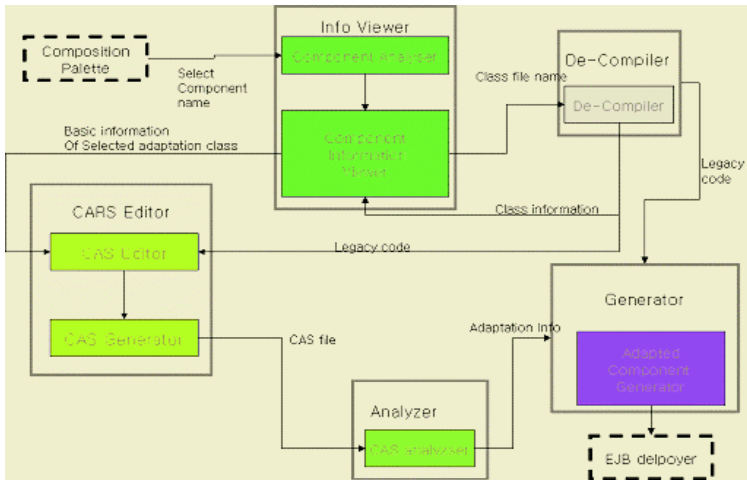
Fig. 2. Component Adaptor System Design

- Component Information Viewer: The list of classes and interfaces included in the package is verified and information regarding each class and interface is provided.
- Decompiler: Various information and source code necessary to understand the classes are acquired through the binary code for the subject for Integration selected in the Component Information Viewer.
- CAS Editor: When requiring component Integrations, The CAS(Component Integration Specification) Editor receives the necessary information by cause for Integration, and creates the CAS file by class or interface which comprise the component.
- Integration Component Generator: New codes are created and packaged by reflecting the contents of the CAS file to the existing component. By analyzing the

CAS Specification File created by the CAS generator through the CAS Analyzer, new codes are created by identifying the information actually adapted and reflecting it to the existing codes created through the decompiler. This is delivered to the deployer which creates an EJB package.

As the specification necessary for Integration, CAS holds specification statements separated by Integration type. The types of Integrations are additions of operations to interfaces, method definition for added operations, renaming of attribute names, renaming of interfaces, and change in previously defined implementation. The scale of Integration can be determined through each of these Integration specification statements. Integration specifications are defined as language types possessing tags.

By analyzing the Integration information from the CAS written in accordance to these types of syntax, the class or interface necessary for Integration from within the original component is adapted by creating a new class and interface by tying them into a new EJB package. By analyzing the written CAS file and extracting the Integration information during the compile process, if this is delivered to the Integration component generator, the Integration component generator then integrates the Integration information with the code created by the Decompiler to create new class codes.

4 Implementation of Component Adaptor

The prototype designs of this research for component Integration is Binary Component Integration(BCA) based Component Integration.

4.1 Introduction to BCA Based Component Integration System

The basic information regarding components is extracted by analyzing the component execution file and as a system which supports the technique of correcting the component execution files itself by analyzing the CAS defined in accordance to the component user, the basic flow is defined as <Fig. 3>.
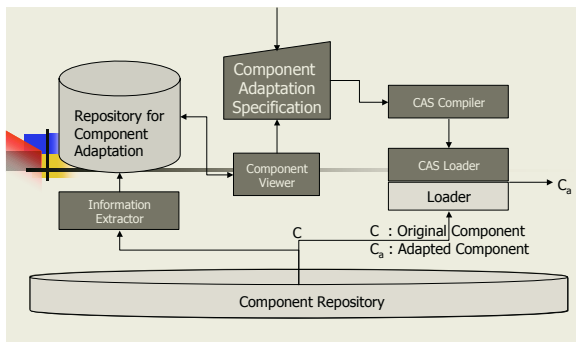
Fig. 3. Introduction to BCA based Component Integration System

The Component Repository operates a functionality which saves and manages components possible for reuse within the common Component Repository. The information analyzer extracts information necessary for component Integration from the

component execution file saved in the repository. The following is the types of extracted information.

1. Interface Information defined in Component
2. Class Information defined in Component
3. Operation Information defined in Class
4. Attribute Information defined in Class
5. Method Information defined in Operation

With this information as a base, the adaptee can understanding the information extracted through the Component Viewer and define the Integration contents through CAS.

(a) CAS & Class Information   (b) Entity Bean Analysis Result Screen
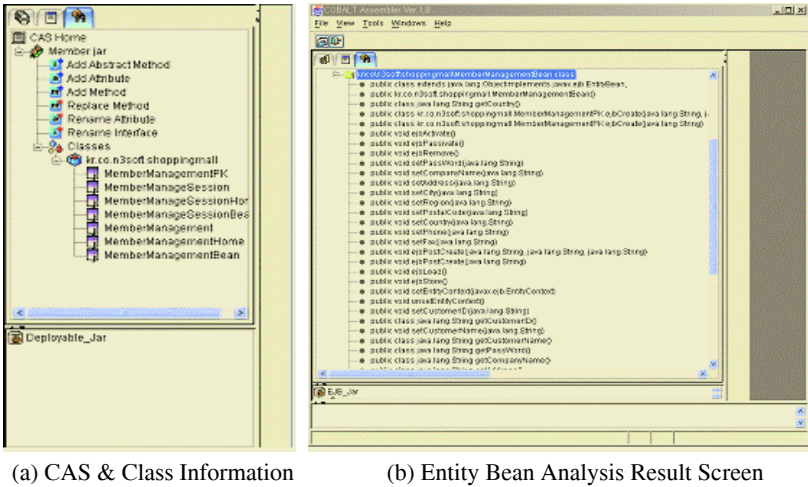
Fig. 4. Component Information acquired through Component Execution File Analyzer

(a) CAS editing for adding Attribute   (b) CAS Editing of adding method
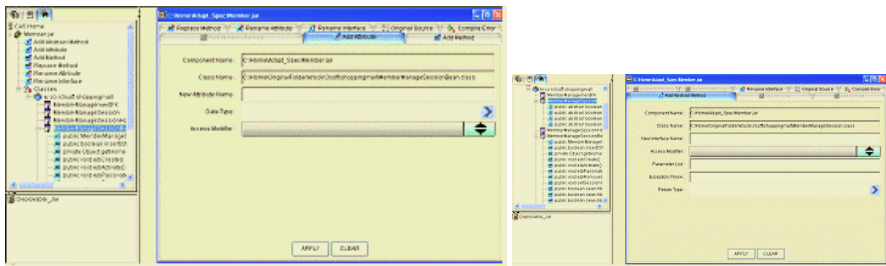
Fig. 5. CAS Editor Screen

5   Conclusion

This research carried the object of designing and developing a component Integration system which supports the Integration process of binary components. For this, the requirements regarding existing component Integration techniques were defined,

possible technical evidence were researched/analyzed and with this foundation, the Integration technologies of components were established and implemented.

This research classified the types of component Integrations largely into 4 types. As the method utilizing the BCA technique is a method which executes the Integration contents defined in the CAS by modifying the execution file itself. As a modification of the existing BCA in order to adapt binary components which do not provide source code, this is an implementation of a method which creates a new component by writing up the Integration specification without source code and integrating this with the existing component. The possibility for EJB application is unclear because instances requiring direct modification of the Java VM (Virtual Machine) may occur when conducting component Integration based on purely BCA methodologies. The disadvantage that platform dependent Integration system must be developed in cases of component Integration based on purely BCA methodologies was resolved. The static information regarding the component structure and the classes which comprise this were acquired by analyzing the existing binary component. After understanding this information, the portions necessary for Integration are identified. The Integration specification is then written, and the existing binary components are modified based upon the developed adaptor.

Through this research, when simple interface modifications are necessary without component re-structuralization, codes for actual components are extracted through the Integration component based Integrations and the class file analyzer, a by-product of BCA technique research. Here, a new component can be created by reflecting the Integration contents defined within CAS. Through this, the integration process for applications can be supported by effective adapting components already constructed or components received from other corporations.

In the future, there remains a need to define the Integration concept itself in a more standardized format by connecting with refactoring. In other words, by standardizing the Integration concept and the processing process classified by cause, errors that may occur during the Integration process can be prevented, and an improvement in the automation of the Integration process can be expected.

References

1. Johannes Sametinger, "Classification of Composition and Interoperation", OOPSLA'96 Poster Presentation
2. Bradford Kain J. "Component: The Basics: Enabling an Application or System to be the Sum of its parts", Object Magazine, Vol 6. No.2, pp. 64-69, April 1996
3. Nierstrasz Oscar, Meijler Theo Dirk, "Research Directions in Software Composition", ACM Computing Surveys, Vol. 27, No. 2, pp. 262-264, June 1995
4. Nierstrasz Oscar, Meijler Theo Dirk, " Component-Oriented SOftware Technology", Object-Oriented Software Composition, Prentice-Hall International, pp. 3-28, December, 1996
5. Jan Bosch. Superimposition: A Component Integration Technique. Information and Software Technology, 41(5):257-273, March 1999.
6. Jim Q. Ning, "Component-Based Software Engineering" IEEE Software, 1997
7. Jim Q. Ning," A Component-Based Software Development Model", in Proceedings of 21th Annual International Computer Software and Application Conference, 1996
8. Ralph Keller, Urs Holzle," Implementing Binary Component Integration for Java", www.cs.ucsb.edu/oocsb

9. Urs Holzle. Integrating Independently-Developed Components in Object-Oriented Languages. Proceedings of ECOOP'93, Springer Verlag LNCS 512, 1993.
10. George T. Heineman, "A Model for Designing Adaptable Software Components," in 22th Annual International Computer Software Annual International Conference, Vienna, Austria, August, 1998
11. George T. Heineman, "An Evaluation of Component Integration Techniques. "Computer Science Department, Worcester Polytechnic Institute, WPI-CS-TR-99-04
12. Erich Gammar,et al, Design Patterns: Elements of Reusable Object-Oriented Software, Addison-Wesley, 1995
13. http://cbs.colognet.org/Integration.php, Component Integration and Assembly
14. Jeong Ah Kim, et al. "Design and Implementation of Component Integration Supporting Tool ," Journal of KIPS(d), 2002
15. P.Oreizy ,et al., "Architecture-based runtime software evolution," International Conference on Software Engineering, 1998
16. Ian Welch and Robert Stroud, "Integration of connectors in software architectures," in 3rd WCOP, 1998
17. Jeong Ah Kim, et al. Component Integration Using Integration Pattern Components, Proceeding of SMC2001, 2001

Automatic Classification Using Decision Tree and Support Vector Machine

Youngshin Han and Chilgee Lee

School of Information and Communication Engineering SungKyunKwan University 300
Chunchun-dong, jangan-gu, Suwon,Kyunggi-do 440-746, S. Korea
yshan@ece.skku.ac.kr

Abstract. The EDS wafer test yield is the most important criteria to evaluate FAB's productivity, so the manufacturing operation's main purpose is to secure new product yield early and maintaining the yield of mass-produced products high. Defining a failed characteristic that's compatible to the device and classifying wafers depending on failure type helps tasks searching for error from FAB become automated. This would be more efficient then existing failed analysis operations and strive to become the basis for improvement in yield and quality. For this method, this research is trying to use a high speed recognition algorithm called SVM (support vector machine) that will define wafer's failed type and automatically classify each one.

1   Introduction

In these days, yield enhancement engineering usually focuses on the investigation of low yield lots, the elimination of defects, process excursions, the correlation between electrical and functional experiment result, and the improvement of baseline product yield [1]. However there is a major problem in conducting verification on types of defect, which is required as a prerequisite for analyzing defects: labor consumption. Not many engineers classify and summarize each case of types of defect manually; after verifying wafer map that occurs everyday, but also many yield related professionals are charged with verifying each case of defect type of every device. Moreover, as defect verification process is done manually, it is difficult to deal with types and causes of defects efficiently. Dramatic increase in memory density fabricated tight design rule and complex cell structure, which in turn gave rise to design margin causing a wide range of failure. However, the failure cannot be used as efficient resource in analyzing yield as characters of each failed device are defined and arranged whenever they are needed. By analyzing the strong points and shortcomings of the research, we propose a new practical method that can be put in to use immediately after using the data obtainable in semiconductor fabrication system and minimizing the error rate through the definition of defects in each type of device.

2   Electrical Die Sorting (EDS)

Fig.1. shows EDS process. The EDS process aims at improving the whole process of FAB by verifying each die's operation in a wafer state and analyzing the causes of

failure in fail dies. Through the assembly process costs can be saved in the next process by only flowing pass dies that were collected through the assembly process, and improve the quality of those dies through a laser repair process.

As shown in Fig.2 EDS process is consisted of 3 different processes; PRE-Laser, Laser-repair, POST Laser. DC test, back grinding, Inking can be added. After EDS process, the process goes in to ASS'Y, Package assembly process.
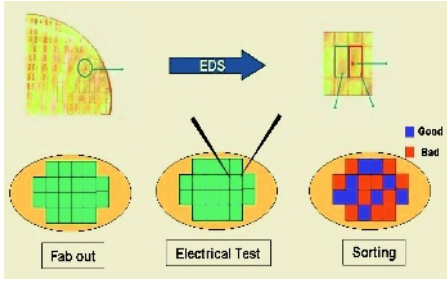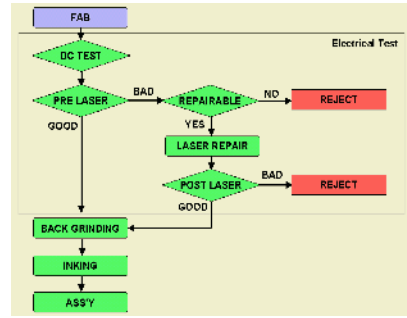
Fig. 1. Electrical Die Sorting process   **Fig. 2.** EDS process flow

3   Analyzing Process of Defect Types

Types of defect can be classified into two categories: chip level and wafer level. Engineers can define various defects in a chip by classifying them in accordance with specific cells' defects contained in each chip. Chips can be categorized into two groups according to the result of tests. In function test, engineers carry out a wide range of tests on chips. Although engineers can detect failure in chips through the tests, they need another classification according to each failure type as the cause of defect can be changed by the influence of failed block. Defects in wafers can be classified according to the distribution type of defective chips on a wafer, and a method finding the cause of defect can be changed by area character. [2][3].

3.1   Definition of a Chip's Defect Types

In type classification, the definition of a chip's defect type should have priority. To categorize the types of defects according to a chip's unit block, this paper uses failed characters of a device in defining each case of a defect in chips with similar cause. There are four types of defects according to a chip's unit block: random, block, area, and special, which consider type and shape of failed test. Special type can be differently defined according to the different character of devices. Random type also can be subdivided into three types according to the distribution density: random, group, and column row type. Finally, block type can be categorized into partial block, block, CSL(Column Select Line), and NWE(New Word Line Enable) according to the distribution characters of vertical and horizontal failure.

4   Chip Level Fail Pattern Classification

4.1   Input Data

Chip level information such as wafer ID, total chip number, x, y-axis coordination and function test result and grade of each chip's unit block are used as data for wafer information of fail bit map

4.2   Decision Tree

Decision Trees [4] are analytical tools for developing hierarchical models of behavior. The tree is built by iteratively evaluating which variable explains the most variability of the response based on the best rule involving the variable. Classes of rules include linear models, binary partitions, and classification groups. Other classes of models could also be considered. Such a rule would partition the data into two groups, one is a group with the true rule and another one is false. The root node of the tree involves a rule that explains the most variability of the response. The child nodes are built by finding other variables that explain the most variability of the data by the first node, etc. The process stops when a point of diminishing returns is reached in a manner similar to automated stepwise regression procedures. Decision trees are useful when the relationships are not known and when you need to make broad categorical classifications or predictions. The following shows a decision rule.

```
IF       93 <=GRADE8      (Character condition)
AND      1 <=Block Count
THEN
NODE 7       (Tree node ID)
N: 58        (Chip number)
Random type              0.0%
Partial block            0.0%
New                      0.0%
CSL                      0.0%
COLROW                   0.0%
Cluster                  0.0%
Block                    100%    (Block type)
```

Example Decision Rule 1 : block type

5   Wafer Level Fail Pattern Classification

Electrical Wafer Sorting (EWS) is the step of the process where all the devices are tested. It consists of a sequence of different electrical measures whereby each device is assigned to a failure class if it fails one of these tests. Electrical failures are the consequence of many different process causes: human mistakes, wrong recipes, equipment or process out of control, equipment malfunctions, too narrow process

windows, scratches due to human and machine causes, and particles. In the wafer level fail pattern classification, failure patterns are searched in a wafer size data where the fail bit map data. We classify wafer level fail bit patterns into 5 categories by considering the characteristics of the manufacturing equipment and their handling of the wafers. Each pattern is ring, central spot, photo spot, continual line, and random type.

5.1   Support Vector Machines

Support vector machines (SVMs) [5] perform pattern classification for two-class problems by determining the separating hyper-plane with maximum distance to the closest points of the training set. These points are called support vectors. If the data is not linearly separable in the input space, a non-linear transformation $\Phi(\cdot)$ can be applied which maps the data points $x \in R^n$ into a high (possibly infinite) dimensional space H that is called feature space. The data in the feature space is then separated by the optimal hyper-plane as described above. The mapping $\Phi(\cdot)$ is represented in the SVM classifier by a kernel function $K(\cdot, \cdot)$ that defines an inner product in H, i.e. $K(x_i, x) = \Phi(x_i) \cdot \Phi(x)$. The decision function of the SVM has the form of:

$$f(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i y_i K(x_i, x),$$

(1)

where $\mathbf{n}$ is the number of data points, and $y_I \in \{-1, 1\}$ is the class label of training point x_j. Co-efficients α_i in Eq. (1) can be found by solving a quadratic programming problem with linear constraints [5]. The support vectors are the nearest points to the separating boundary and are the only ones for which α_i in Eq. (1) can be nonzero. An important family of admissible kernel functions are the Gaussian kernel, $K(x, y) = \exp(-\|x - y\|/2\sigma^2)$, with σ the variance of the Gaussian, and the polynomial kernels, $K(x, y) = (1 + x \cdot y)^d$, with d the degree of the polynomial. Let M be the distance of the support vectors to the hyper-plane. This quantity is called margin and it is related to the co-efficients in Eq. (1),

$$M = (\sum_{i=1}^{n} \alpha_i)^{1/2}$$

(2)

The margin is an indicator of the data separability.

5.2   Experiment and Result

The operating system and programs are consisted of Window 2000, MATLAB 5.2 and VC++6.0. The subject device is a DRAM consisting of 1000 wafers. For the learning process, each class has 30 Fail Bit Map total of 210 as learning materials for the SVM. The size of the input data are 25 by 25 with failed data regarded as 1 and normal data as 0. The input vector of the training samples is also named the characteristic vector. The detailed notations are explained below and represented in Fig. 3.

An experienced engineer will typically classify a failed pattern accurately about 60~80% according to an experiment [6]. Table 1 is the result of classification algo-

rithm. There were 712 simple and 288 complex failed patterns out of a 1000. Among the complex pattern, an excessive complex pattern can result in an error, however only 77 appeared. Extensively classifying failed pattern would be possible for more accurate pattern classification, but would lower the operating speed. The subject research accurately classified over 90%, but the percentage of misclassifying the complex type was high. So, required Rule to improve its flexibility, extensibility and analysis of the probability of misclassification.
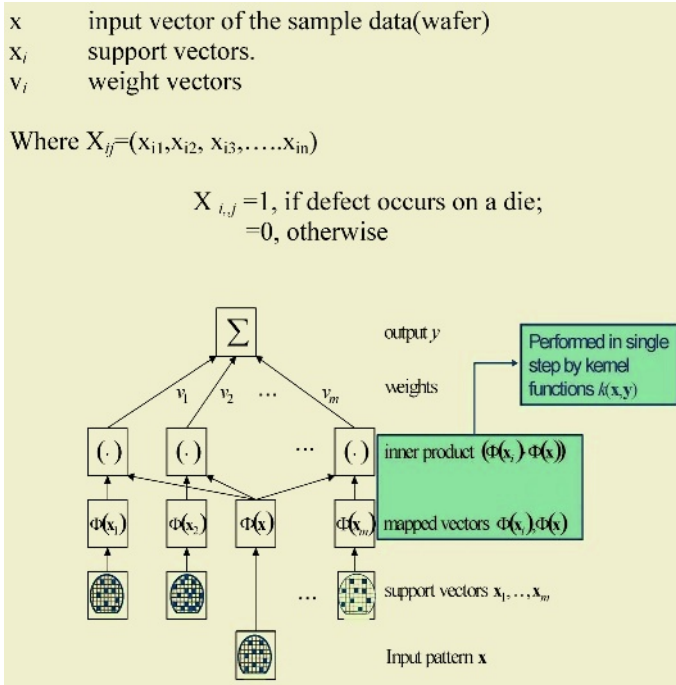
Fig. 3. Fail pattern classification process

Table 1. The result of Classification algorithm

	Sample number	Accuracy
Wafer type analysis	1000 (wafer)	923/1000(92%)

6   Conclusion

During the semiconductor manufacture process, a system that was efficient in fail analysis operation was constructed by presenting a method to automatically classify fail types using the fail bit map data acquired from the EDS tests and apply the method to existing manual operations. Comparing using bin information from each of the wafer's chip to extensively classify into fail bit units from existing methods, resulted in improved accuracy and operation time was reduced by 80%. Moreover, additional operations and possibility of applying without any loss links to quick and accurate fail cause analysis that contribute to improved wafer's yield and quality.

References

1. Chen, F.L., and S.F. Liu, "A neural network approach to recognize defect spatial pattern in semiconductor fabrication", IEEE Transaction on Semiconductors Manufacturing, vol.13, pp. 366-373., august, 2000.
2. W. Shindo, E. H.Wang, R. Akella, and A. J. Strojwas, "Effective excursion detection by defect type grouping in in-line inspection and classification," IEEE Trans. Semiconduct. Manufact., vol. 12, pp. 3–9, Feb. 1999.
3. Fei-Long Chen and Shu-Fan Liu, "A neural-Network approach to recognize defect spatial pattern in Semiconductor fabrication", IEEE *Trans.Semiconduct. Manufact.*, vol.13, no. 3, Agust.2000.
4. C. Westphal and T. Blaxton, "Data Miniing Solutions", John Wiley and Sons, New York(1998)
5. V. N. Vapnik. Statistical Learning Theory. Wiley, New York, 1998.
6. P. B. Chou, A. R. Rao, M. C. Struzenbecker, F. Y. Wu, and V. H. Brecher. Automatic defect classification for semiconductor manufacturing. Machine Vision and Applications, 9(4):201–214, 1997.

Opportunity Tree Framework Design for Quality and Delivery of Software Product

Sun-Myung Hwang[1] and Ki-won Song[2]

[1] Dept. of Computer Engineering, Daejeon University
sunhwang@dju.ac.kr
[2] School of Computer Science and Engineering, Chung-Ang University
from6588@object.cau.ac.kr

Abstract. An organization can optimize a project and strengthen control of it and thus, accomplish its objectives by determining its project capability through the proposed models, by planning the most suitable project to its vision. The approach combines two methods. One looks at an organization's measurement framework in goal-oriented fashion and the other looks at it in vision driven. The goal-oriented method was applied to improve the quality and delivery measurement from the point of view. It identified several new metrics, and also contributed to better understanding the collected data of user requirement. The vision-driven method was used to gain new insights into the existing PCM(Performance Calculation Model) data. This paper gives the case study and its results to qualitatively compare our approach against current ad hoc practices used to improve existing Opportunity Tree frameworks(OTF).

1 Introduction

This paper measures quality and delivery attributes for an organization's external effectiveness. It also measures process capability and project capability through PCM by completing and analyzing a questionnaire based on GQM (Goal Question Metrics) to find a way to improve the process. Based on the analysis of results, PCM can be designed to measure performance score for project development. OTF can be designed to analyze optimized route for improvement through which effective project plan suitable to the organization's vision can be developed

2 Basic Study

2.1 Why Is Earned Value Needed for IT Business?

Promoting resources-managing ability to effectively invest IT resources and maximize their effect is becoming an essential field in the IT industry [2] [3].

   IT emerged as a key area to reengineer and improve business process, along with using computers. Large corporations including IBM, Ford, and GE are enjoying 80% more effect from business process reengineering using IT than from the improvement just using computers.

2.2 The Performance Pyramid

As referred former chapter, SPI is a very important factor in organization's achieving vision. So, an evaluation model is required in order that we evaluate efficiency of SPI for this vision and introduce a optimized strategy[13][15].

But in software process, it is difficult and invisible to understand the extent of achievement for organization's vision due to the immaterial characteristics of a software development process. To solve these problems, Lynch and Cross suggest the methodologies of performance pyramid as <Figure 1>[4][10].

Fig. 1. Lynch and Cross's Performance Pyramid

For example, if an organization's vision is improvement of public recognition on it, it should concentrate on market shares rather than on financial performance, and put its priority on customer satisfaction and flexibility rather than on flexibility and productivity[1][13][14][17]. Quality and delivery is goals to measure external effectiveness while cycle time and waste is to measure internal efficiency [1].

2.3 GQM Approach

GQM approach involves three steps.

First is conceptual step. It consists of elements such as object, purpose, viewpoint and focus. In this step, major goal are set. Second is operational step. In this step, questions are derived from the goals that must be answered in order to determine if the goals are achieved. Third is quantitative step in which proper answers are given to the questions. Through the three steps, metrics system is made. These metrics can be used as a measurement tool [10] [11] [16] [17].

3 PCM (Project Capability Model)

This section suggests PCM which can measure an organization's capability through completion and analysis of questionnaire.

PCM calculates project-performing capability of an organization with GQM approach regarding each of 2 performance attributes in the performance pyramid (Lynch and Cross). Make questions and develop metrics measure accomplishment of the goals using the metrics. Based on the performance pyramid, GQM quantitative questionnaire is made which enables calculation of an organization's capability by using GQM approach. GQM quantitative questionnaire is composed of items with which external effectiveness of an organization can be measured.

3.1   Data Collecting Method

This paper uses data from questionnaires on 20 tasks of corporations which are collected from SPICE assessment[8][9].

3.2   GQM Quantitative Questionnaire from Project Meta Data

This section proposes GQM quantitative questionnaire made from general meta data.

Procedures for making the questionnaire involving three steps: setting goals; giving questions; gaining metrics.

Through the three steps, metrics system is made. 10 measurable metrics were made for 3 questions. GQM results gained though three steps are shown in <Table 1> [5][6][7].

Table 1. Metric Results based on GQM. For Quality and Delivery

Goal	Question	Metric
Quality (To improve quality of product up to level of satisfying customers)	Defect density (In the project, how densely defect are found and properly dealt with.)	Defect rate of products
		Defect rate of technical documents.
		Defect rate of codes
		Defect management rate.
	Impact requirement (How much impact customer's requirement of change has on project?)	Requirement change rate
Delivery (Shorten time needed to deliver product to customers)	Delivery time (Are products delivered to customers on schedule?)	General on-schedule-rate
		On-schedule-rate at planning/analysis stage
		On-schedule-rate at design stage
		On-schedule-rate at implementation stage
		On-schedule-rate at test stage

3.3   Project Capability Measurement Model

This section proposes PCM to calculate project capability in terms of external effectiveness of an organization. Input data of this model is data collected from GQM questionnaire as suggested in 3.2. Factors of PCM to measure a project capability of an organization for 2 goals are shown in <Table2>

Table 2. Factors of PCM to measure a project capability of external effectiveness

Question	Metric	Capability Measure Factor
Defect density (Quality)	Defect rate of products	Total number of defects
	Defect rate of technical documents.	(Requirement specification + design specification) number of defects and total number of pages of outcome
	Defect rate of codes	Number of code defects, total SLOC
	Defect management rate.	Number of complete correcting defects.
Impact requirement (Quality)	Requirement change rate	Number of requirement change, total number of requirement
Delivery time (Delivery)	General on-schedule-rate	Total delivery days, planned delivery days
	On-schedule-rate at planning/analysis stage	Actual delivery days, planned delivery days at planning/analysis stage
	On-schedule-rate at design stage	Actual delivery days, planned delivery days at design stage
	On-schedule-rate at implementation stage	Actual delivery days, planned delivery days at implementation stage
	On-schedule-rate at test stage	Actual delivery days, planned delivery days at test stage

Calculation forms to measure a project capability of an organization in terms of external effectiveness for 2 goals are shown in <Table 3>

Table 3. Calculation forms to measure a project capability of an organization in terms of external effectiveness

External effectiveness PCM(qd) = (PCM(q)+PCM(d)/2	PCM(q): quality effectiveness score	$\sum((100 - each\,defect\,rate) + defect\,management\,rate)/4$
	PCM(d): schedule effectiveness score	$\sum(100 - on\,schedule\,rate\,at\,each\,stage)/5$

Besides, we understand the achievement level of vision in Table3. Because External Effectiveness score of an organization does to base on Performance pyramid, score of Quality and score of Delivery are elements that act great role in customer satisfaction side. Through this, score of each focus on Performance pyramid is able to calculate arithmetically.Example that show score of each focus is shown in <Figure 2>

Fig. 2. Arithmetical capability score through PCM

In case all of goals in each focus of an organization is important equally, <Figure 2> is as result that calculate arithmetically.

4   OTF(Opportunity Tree Framework)Model for Quality and Delivery

In this section, we propose OTF that decision-making of route for effective process improvement that is optimize in organization's vision based on present project performance capability. Decision procedure of OTF is as following. And, the structure is shown in <Figure 3>

Fig. 3. Composition of OTF for Quality and Delivery

4.1   Qualitative Questionnaire

Existent improvement selection method improves the lowest capability score simply. However, the method is difficult to establish optimized improvement strategy that is

set to an organization's vision. Qualitative questionnaire asks vision weight that vision which an organization wants can be reflected. Qualitative questionnaire consist of 3 Phase. Through the Qualitative Questionnaire, In Phase 1, CEOs of an organization can know whether is emphasizing more on some part of Market side. In Phase 2, Middle managers can know whether have vision of some viewpoint among Customer satisfaction, Flexibility. In Phase 3, Project developers can know whether is which side that being very interested part among quality, delivery,. This can grasp all the stakeholder's vision which connect in achievement of an organization.

4.2   Route for Improvement

In this Section, we propose that a Model which look for optimized improvement method. The Model compounds that vision weight of stakeholders that is obtained through the Qualitative Questionnaire and project capability of existent organization that is calculated through PCM. And the Model should be effective all stakeholders.

For this, each focus of performance pyramid like <Table 4> is selected eight kinds of Route for improvement that stakeholder's vision is reflected

Table 4. Route for improvement

Business unit[Bu]	Core business process[Cp]	Department groups[Dg]	Route for improvement [Ri]
[M]arket (External effectiveness)	[C]ustomer Satisfaction	[Q]uality	MCQ = [M]+[C]+[Q]score
		[D]elivery	MCD = [M]+[C]+[D]score
	[F]lexibility		MFD = [M]+[F]+[D]score
			FFD = [F]+[F]+[D]score

4.3   Most Suitable Improvement Route Selection for Vision Achievement

In <Figure 4>, above score can not select route for improvement of most suitable that organization's vision is reflected present project performance capability.

Fig. 4. Calculation of vision achievement

Therefore, we need model to select route for improvement of most suitable. And Vision score means priority that organization wants. The vision score is thing which importance that organization wants is high when point is high.

Therefore, the priority of improvement is combined lower priority of arithmetic score with higher priority of vision score. It means the order that is reflected present performance capability and vision of all stakeholders in organization.

5   Case Study of PCM and OTF, and Verification of Reliability

5.1   Case Study Method of PCM and OTF

This section verifies reliability of PCM and OTF through case study using data on GQM quantitative questionnaires collected from SPICE assessments made during 2003 to 2004 period. This paper uses data of three organizations as project data. <Table 5> provides specifics of each of the three organizations

Table 5. Specifics of each organization

	A company	B company	C company
Nature of task	Development task	Commercialization task	Development task
Motive of project	For commission from other organization.	For commission from other organization.	For internal study of the organization
Project domain	Mobile	Computer	Multimedia

5.2   Verification of Reliability Using PCM and OTF Cases

For case study, this paper used data which were collected from GQM questionnaires of three companies by using Excel as an automation tool

Reliability verification results of 3 companies through case studies are as follows.
<Table 6> and <Table 7> are summarized results of case studies.

Table 6. Analysis results of case studies of PCM

	PCM capability score	Capability score of each goal	
A company	74.36	82.09	66.64
	PCM(QD)	Quality > Delivery	
B company	80.16	88.97	71.34
	PCM(QD)	Quality > Delivery	
C company	86.66	87.42	85.9
	PCM(QD)	Quality > Delivery	

Table 7. Analysis results of case studies of OTF

	Route For improvement	A company Vision:	B company Vision:	C company Vision:
Priority of improvement	MCQ	7	3	8
	MCD	4	2	7
	MFD	2	1	6
	FFD	6	5	3

6   Conclusion and Hereafter Research

In this paper, we could find the optimized process improvement route that can heighten the degree of organization's goal achievement through Case Study.

Our future wok is the following steps:

1. Design PCM metric by GQM approach in Cycle time and waste for efficiency of organization.and Design Route for improvement in OTF.: FFC, MFC, FPW, FPC
2. Build Experience Factory that is accumulated experience and results after software process improvement through OTF model in real projects.
3. Design Integrated WEB-based Opportunity Tree that can provide the most effective improvement strategy or method.

References

1. Richard L.Lynch, Kelvin F. Cross, "Measure up!", 1995, Blackwell.
2. Kyung-whan Lee, "Modeling for High Depending Computing", The fifth Korea Information Science Society's Software Engineering Association,Feb.20. 2003
3. Kyung-whan Lee, "ROI of IT Business" , The federation of Korean Information Industries, 2003. 5
4. Kyung-whan Lee, "Quantitative Analysis for SPI" , Corporation seminar, Feb. 17. 2003.
5. Boehm, C. Abts, A.W. Brown, S. Chulani, B. Clark, E. Horowitz, R. Madachy, D. Riefer, and B. Steece, "Software Cost Estimation with COCOMO II", Prentice Hall, 2000.
6. Steece, B., Chulani, S., and Boehm, B., "Determining Software Quality Using COQUALMO," in Case Studies in Reliability and Maintenance, W. Blischke and D. Murthy, Eds.: Wiley, 2002
7. Mark C., "The Capability Maturity Model Guidelines for Imporving the Software Process", CMU/SEI, 1994
8. ISO/IEC JTC1/SC7 15504: Information Technology-Software Process Assessment, ISO TR, ver.3.3, 1998
9. KSPICE (Korea Association of Software process Assessors), SPICE Assessment Report http://kaspa.org, 2002 2003
10. V. R. Basili, G. Caldiera, H. D. Rombach, "Goal Question Metric Paradigm", Encyclopedia of Software Engineering, John Wiley & Sons, Volume 1, 1994, pp. 528-532.
11. Frank Van Latum, Rini Van Soligen, "Adopting GQM-Based Measurement in an industrial Environment", 1998, IEEE software
12. Tim Kasse, "Action Focused Assessment for software process improvement", Artech House, 2002.
13. Williams A. Florac, Anita D. Carleton, "Measuring the software process", 1999, SEI Series, Addison Wesley.
14. Tom Gilb, "Software Inspection", Addison-Wesley, 2001.
15. Ki-Won Song, "Research about confidence verification of KPA question item through SEI Maturity Questionnaire's calibration and SPICE Level metathesis modeling", SERA03, San Francisco, 2003.06
16. Barry Boehm, "Value-Based Software Engineering: Case Stydt", IEEE Computer, pp33-41, March 2003.
17. Donald J. Reifer, "Making the Software Business Case", Addison-Wesley, 2002.

Breeding Value Classification in Manchego Sheep: A Study of Attribute Selection and Construction

M. Julia Flores and José A. Gámez

Departamento de Informática & SIMD - $i^3\mathcal{A}$
Universidad de Castilla-La Mancha
Campus Universitario s/n. Albacete, 02071
{julia,jgamez}@info-ab.uclm.es

Abstract. Estimating animal's genetic merit (or *breeding value*) plays a major role in the Manchego sheep selection scheme (ESROM), started fifteen years ago with the goal of improving Manchego sheep production figures. In the ESROM scheme the breeding value is estimated each semester by using BLUP animal model. In this paper we study the use of data mining techniques to deal with breeding value classification. The purpose of the paper is not to replace the use of BLUP in the ESROM, on the contrary, we intend to learn in a supervised way from the results produced by BLUP, and to use the learned models to provide preliminary information about the breeding value of an animal. We use standard classification techniques combined with feature subset selection in order to identify good (subsets of) predictors. We also show that the classifiers accuracy can be considerably improved by attribute construction.

1 Introduction

Manchego sheep are the native breed in Castilla-La Mancha (Spain), and the two main products obtained from them (Manchego cheese and Manchego lamb) represent more than 50% of animal production in the region. Because of these economical implications and with the aim of enhancing Manchego sheep production, a selection scheme (called ESROM) based on the animal's genetic merit was started by the authorities fifteen years ago, with the goal of improving milk production of Manchego ewes. One of the major points in the ESROM scheme is the estimation of the animal's genetic merit or breeding value (BV), and its use in flock replacements. In the ESROM scheme the BV is estimated by using BLUP animal model, which is a complex method based on relating different traits by equations and solving them by simultaneously considering all the available information. The estimated BV allows us to place animals in the genealogical ranking and to take decisions about which animals will improve the flock genetic trend, which animals can be entered (or not) in the stud catalogue or market, which ewes are candidates to be used as good mothers by artificial insemination, etc, ... Besides, the ESROM encourages stock breeder to select their flock replacements on the basis of the animal BV.

The BV of an animal is a numerical value that in the ESROM represents the deviation of the animal with respect to the mean BV of the Manchego ewes born in 1990 (the base year). The estimation of the BV [1] by BLUP is a complex process that in the ESROM is carried out every six months in a specialized center. Furthermore, the BV of an animal is a dynamic value, because it can change from one measurement to the next due to changes in the animal's own production data, due to changes in its relatives, due to changes in its flock, etc.

Even though the BV is numerical, many decisions are taken based on the percentile in which the animal is ranked with respect to its/(the whole) flock. Depending on the decision, different thresholds are used: above 50%, above 70%, etc. Because of this, the goal of this paper is to work on the classification of the BV inside the ESROM scheme by using techniques from machine learning [12] and data mining [6], whose application to agriculture has gained interest during the recent years [4]. Obviously, our target is not to replace the use of BLUP methodology, but to study the possibilities of predicting an animal BV by using a data-driven approach, which will use less information (by far) than BLUP, that is simpler and that can be used as soon as the information for an animal is ready, without having to wait for the full breeding six-month evaluation.

To achieve this goal we have structured the paper in five sections apart from this introduction. In Section 2 we describe the datasets used in this work. In Section 3 we describe the initial classification process carried out. Section 4 is devoted to applying different variable selection techniques while in Section 5 attribute construction is studied. Finally, in Section 6 we conclude.

2   Datasets

The datasets used in this work have been obtained from AGRAMA[2] data bases, which contain data from 1989 to 2003. After the data preparation process (described in [7]) and following AGRAMA experts advice we get a dataset with 3087 records and 24 variables (distributed in four groups in Table 1). We seek to get an estimation of ewes BV at early stages, all the records in the dataset correspond to primipara ewes, because it is after the first offspring-birth and its corresponding lactation when ewes are evaluated for the first time, and so, it is of interest to have, as soon as possible, an approximation of their genetic merit in order to make early decisions about them. Besides, only those animals whose (and parents) estimated BV is sufficiently reliable were added to the dataset [7].

From this original dataset we obtained two different ones by discretizing the BV variable in order to transform the regression problem into a classification one. Following the indication of AGRAMA experts two discretizations were carried out, because they can help us to get insight about our problem. Thus BV was discretized in four and five bins of equal frequency giving rise to class variables BV4= $\{f1, f2, f3, f4\}$ and BV5=$\{v1, v2, v3, v4, v5\}$. We denote 4labels and 5labels the two datasets coming from replacing BV with BV4 and BV5 respectively.

---

[1] We should use *Estimated Breeding Value* (EBV), but for the sake of simplicity we maintain the notation of BV, although it is clear that we are dealing with estimations.

[2] National Association of Manchego sheep breeder.

From the study of the variables in the dataset we realize that environmental variables (except TypeOfBirth) are nominal variables with a really large number of possible outcomes (from 64 to 127). This type of variables can introduce a considerably amount of noise in the learning/classification process, so we have preprocessed them by implementing in ELVIRA [3] the method proposed in [1]. This method reduces the number of values for a nominal variable to $|C| + 1$ labels, $|C|$ being the number of classes (and adding an *unknown* extra group).

Two more datasets, 4labels-d and 5labels-d, were obtained by discretizing all the numerical variables by using Fayyad and Irani supervised algorithm [5]. After this process the number of values in the variables of 4labels-d goes from 2 to 24 with a mean of 6.1, and in 5labels-d goes from 2 to 18 with a mean of 5.9.

3   Initial Classification Process

We have used two standard (but competitive) algorithms to carry out the classification process: decision trees (C4.5) [13] and Naive Bayes (NB) [2], concretely their WEKA [14] implementations with the default parameter setting. The accuracy of the classifiers is measured by using stratified 10-fold cross validation.

Instead of directly using all the variables in the dataset, we have run the algorithms starting with a small subset and progressively adding different groups of variables. Thus, we have carried out the following process (see Table 2.a):

1.- Only the BV of both parents, BVp = {BVFather, BVMother}, are considered.
2.- All the BV variables (BVall) are used as predictive attributes. Surprisingly, there is only a slight improvement in one out of the eight cases. Because of this, we maintain these two sets as different starting points for the remaining process.
3.- From these two starting points we considered new cases obtained by adding the remaining groups of variables: environmental (env), lactation data (lact), and mother lactation data (which gives rise to the full set).

Analysis.- As our classification problem has been artificially constructed from a regression one and all the classes are equally distributed, we should not expect high classification rates. In fact, the accuracy obtained by a baseline algorithm such as *OneR* [9] is between 48% (4labels) and 55% (5labels). Having this in mind, we think that the results obtained when all the variables are used (last row of table 2.a) are not bad, specially for C4.5. Furthermore, these results are improved (with significant statistical difference in all the cases but one, C4.5d) when not all the variables are included as predictive attributes. Concretely, in

Table 1. Variables in the data set. **BV** is the target variable.

BV data:	Environmental data:	Mother Lactation data:
BVFather, ReBVF, BVMother, ReBVM,	TypeOfBirth	NLactM
BVMaternalGM, ReBVMGM, BVParentalGM,	StockFarm	AvLactNormM
ReBVPGM, BVMaternalGF, ReBVMGF,	FatherStockFarm	MaxLactNormM
BVParentalGF, ReBVPGF, **BV**	MotherStockFarm	AvLact120M
Lactation data: AvLactNorm, AvLact120		MaxLact120M

most of the cases the best results are obtained when BVp+env+lact are used as predictive attributes, obtaining an accuracy of 76% in the 4labels problem and almost a 70% in the 5labels problem. Besides, it is interesting to point out that C4.5 obtains its best result in the discretized case, while NB does in the non-discretized one. From these observations and from the fact that C4.5 carries out an implicit variable selection process, we can conclude that using all the variables as predictive attributes is not a good idea here.

4   Feature Subset Selection

Feature (or variable, or attribute) Subset Selection (FSS) is the process of identifying the input variables which are relevant to a particular learning (or data mining) problem [8, 10]. Here, the goal of FSS is double: (1) to select the subset of variables yielding the best classification performance, and (2) to identify those variables with major influence in BV prediction. Although the process in Section 3 can be viewed as a *manual* FSS process, now we study the application of automatic FSS techniques. We have considered the following four approaches:

• **Embedded** methods. In this case it is the own learning algorithm which carries out the selection during the learning process. An example of this approach are decision trees. Thus, when C4.5 receives the subset BVp+env+lact as input, only 7 or 5 variables are actually used for the original and discretized datasets respectively. It is worth noting that C4.5 selects the same subset independently of the number of class labels, and that the discretized version (which gets the best results) needs only 5 out of the 7 variables used for the original dataset.

Our approach here consists in using C4.5 as feature selector taking the selected subset as the input for NB (Table 2.b). By doing so, NB improves the best result obtained so far for the discretized versions and also the results obtained when using the full set of variables. However, as the used subsets are biased by C4.5, we can expect to improve these results by using different FSS methods.

• **Filter** methods. They usually work by measuring the relevancy between the predictive attributes and the class. They are fast and independent of the learning algorithm to be used. We have used *symmetrical uncertainty*, $SU(C, X) = \frac{2 \cdot ((H(C) - H(C|X))}{H(C) + H(X)}$, to do the ranking. Basically, SU uses a projection of mutual information onto the $[0, 1]$ interval. In our opinion SU is quite interesting when the number of states in the involved variables is different.

From this ranking we use the first k variables as the selected subset. Due to the knowledge we have gained in Section 3 we have chosen $k = 6$ (at least a lactation variable is included) and $k = 9$ (at least an environmental variable is included). However, only in one case (5labels-d) we improve the best result obtained until now and also only in one case the selection of 9 variables instead of 6 yields a slightly better result (5labels). In contrast, the performance of NB is clearly degraded, which can be due to the selection of redundant variables. In fact, this is one of the main problems of filter methods, two variables can be (independently) relevant for the class, but redundant between them.

- **Wrapper** methods. They use a learning algorithm as part of the selection process, that is, the merit of a given subset is measured by learning (and evaluating) a model using only that subset of variables. Because of this, the wrapper approach (usually) obtains better subsets than filter methods. On the other hand, they are computationally expensive and the obtained subset lacks generality because it is tied to the bias of the classifier used during FSS process. In this work we have used the well known *forward* FSS combined with best first search. Best first search enhances forward selection by allowing backtracking during the search, but it has the disadvantage of (exponentially) increasing the cost with the number of (irrelevant) variables. In our experiments we have set the number of allowed backtracking levels at five.

- **Filter+Wrapper** methods. In this case both approaches are combined in some way. The most common is applied when there is a large number of predictive attributes, and consists in a two-stages approach: apply a fast filter algorithm and get the first k variables, and then use a wrapper one over these k variables. Our scenario is different because we have less than 30 attributes, so we propose the application of a different combination of filter and wrapper approaches.

Our idea (FW algorithm) is to use the filter ranking not to remove variables, but to guide the operation of the wrapper algorithm. Thus, the wrapper stage involves running over that ranking and adding a variable to the subset of selected variables only if such inclusion improves the classifier accuracy. The classical (greedy) stopping criterion for this process is to finish when adding a new variable does not improve the accuracy, which usually gives rise to stop because two correlated variables appear together in the ordering, even if any (uncorrelated) relevant variable appears later in the ranking. Trying to alleviate this problem, we propose to consider a lookahead parameter (k), which allows us to continue the process (discarding the useless variable) if less than k useless variables have been consecutively discarded. Of course, if $k = 0$ we get the greedy behavior and if $k = \infty$ all the predictive variables have a chance. We have used SU and $k = 0, 5$ and ∞.

Analysis.- Now we analyze the behavior of FW and Wrapper and compare them. The greedy approach represented by FW(k=0) only includes the two (three in one case) first variables in the ranking (BVFather and BVMother). Unsurprisingly, with this small subset NB improves filter results in all cases, which remarks the idea of having redundant variables among the (*relevant*) first variables ranked by SU. This is not the case for C4.5, that gets better results with the previous FSS approaches, probably because of its own filter selection process. Things are quite different when using FW(k=5). In this case a new subset, S_3={BVFather, BVMother, AvLact120}, arises as a very good predictor. That is, with a lookahead of 5 FW is able to discard some (possibly relevant) attributes which are redundant with respect to those already included, and to continue the search for new relevant (but not redundant) attributes. S_3 is selected in the four cases in which NB is applied and it is complemented with some other variables when applying C4.5. In all the cases (NB and C4.5) the results obtained when using FW(5) are better than those obtained so far. Finally, we have tried FW(k=∞).

This lookahead value increases the complexity of the process, but in no case the number of evaluated subsets is linear in the number of variables. The results obtained are only slightly better than those obtained with k=5, and we get the same trend with respect to the selected subsets, i.e., S_3 complemented with one or two more variables in most of the cases.

With respect to Wrapper FSS, the experiments (last row of Table 2.b) show that in five out of the eight cases Wrapper gets (slightly) better results than FW. However, there is not statistically significant differences between wrapper and FW in any of the eight cases, which shows us that FW is in fact quite competitive (at least in this domain). In relation to the subsets selected by this approach, again they are extensions of S_3 by adding few new variables, that contain in most cases data about mother lactation (which seems to be quite reasonable) and confidence measures. On the other hand, the wrapper approach has greater complexity than FW(∞) because it evaluates (on average) about 200 subsets, while FW(∞) evaluates exactly 23 subsets over each dataset.

Finally, let us to emphasize that the benefit of using variable selection is reinforced by the fact that in the eight cases there are statistically significant difference between the best result obtained in the initial classification process and the best result obtained in this section.

5   Attribute Construction

Attribute (or variable or feature) construction is the process of deriving new attributes from the original ones. The idea [11] is to apply a set of constructive operators to the existing attributes resulting in the construction of one or more new attributes more appropriate for the description of the target concept. One of the goals of attribute construction [8] is to improve the accuracy when making predictions. In this work we focus on a data-driven approach which uses generic attribute construction methods, thus we create new features by applying simple arithmetic operators ($+$ and $\times$) to pairs of numerical variables. Given that summation and product are commutative, $\frac{n^2-n}{2}$ new features are generated for each one. The choice of these operators is due to the fact that they are simple, and summation can be appropriate when attributes are measured in *similar* domains, while product can be appropriate when they are measured in different domains.

After the attribute construction, $n^2 - n$ new features have been added to the dataset, n being 19. Thus, our new dataset has 366 variables (the class plus 23 original attributes plus 342 constructed attributes). We have used SU as filter to rank the 365 predictive attributes and we get that the first feature is: BVFather+BVMother. This fact is not surprising at all for two reasons:

− These two attributes were ranked as the two of greatest importance with respect to the class variable, and
− The experts frequently use the pedigree index as predictor, which is computed as $(\text{BVFather} + \text{BVMother})/2$.

Therefore, we have identified a good predictor by using the data driven constructive approach, and it agrees with the experts in the domain. The question

now is if more interesting combinations have been identified. By inspecting the ranking produced by SU we observe that the attributes in the top of the ranking are those combining BVFather or BVMother with lactation data. Thus, it seems that we have effectively identified other interesting combinations.

As the datasets have 365 attributes, most of them coming from attribute construction, we can be almost sure that many of them will be irrelevant or redundant with respect to our classification process. Therefore, a selection process is even more necessary than in our previous experiments, and so we have applied FW and wrapper over our new (larger) dataset.

Analysis.- From the results (Table 2.c) we can observe that the accuracy of the classification has been improved considerably with respect to the results obtained over original datasets. Thus, when using Wrapper selection the accuracy has risen 3.6% (on average) for the 4 labels problem and 5% (on average) for the 5 labels problem. On the other hand, the number of attributes selected is quite far from the 365 available features, the best cases being less than ten. In general, NB needs less attributes than C4.5 and Wrapper selects less attributes than FW(∞). With respect to accuracy, wrapper improves FW(5) in all the cases, however it only get statistically significant difference with respect to FW(∞) in one out of the eight cases. On the other hand, Wrapper is (by far) more complex because it needs to evaluate (on average) about 4800 subsets while FW(∞) is linear in the number of attributes, that is, it evaluates exactly 365 subsets. Finally, with regard to the selection, BVFather+BVMother is always included, complemented with different constructed attributes (mainly combinations of these two primary attributes and AvLacNorm with others) that allow the classifiers to be better adapted to some areas of the solution space. However, it seems that the attributes are selected to improve the accuracy, but no semantic interpretation can be (at least easily) obtained from that selection.

To end with this analysis, it is clear from the statistical analysis that the best model in the four labels problem is C4.5(d), while in the five labels case there is no significant difference among all the models but NB.

6   Conclusions

A study of the use of data mining techniques in Manchego sheep breeding value classification has been done in this paper. This task is a key point in the ESROM scheme, used to improve the quality and production figures of Manchego sheep. Two standard algorithms (NB and C4.5) have been used. We have shown that feature selection is a key process with a twofold benefit: (1) identification of small subsets of variables to be used as predictors, and (2) an improvement in the accuracy of the classification. Furthermore, a data-driven attribute construction process has been carried out over the numerical attributes included in the dataset. From this process some interesting new attributes have been identified and also the classification accuracy has been considerably improved. Furthermore, the obtained classifiers can be collapsed to different two-label problems which correspond to different decision problems inside the ESROM [7].

Table 2. Results of the experiments carried out: (a) over (subsets of) the initial dataset, (b) applying feature subset selection and (c) applying attribute construction plus FSS. In boldface we highlight the best result for each dataset and each classification process. The subscripts indicate the cardinality of the selected subset. The symbols used make reference to the statistical analysis (two-tailed paired t-tests) carried out: • means statistically significant difference with respect to the best result in the same column and block (a,b or c); ○ means statistically significant difference with respect to the best result in the column; and, ▷ means statistically significant difference among the best results in the last block (i.e. NB, NB(d), C4.5 and C4.5(d)).

	4 classes				5 classes			
Variables	C4.5	C4.5(d)	NB	NB(d)	C4.5	C4.5(d)	NB	NB(d)
BVp(arents)	• 72,34	• 71,79	○ **72,27**	○ 70,46	• 66,18	• 66,73	○ **66,47**	○ **63,59**
	±2,07	±1,83	±1,65	±1,55	±3,08	±3,11	±2,84	±2,91
BVall	• 70,46	• 71,36	• 65,50	• 62,71	• 64,23	• 66,86	• 58,05	• 55,17
	±2,56	±1,48	±2,10	1,84	±2,36	±2,61	±1,69	±3,05
BVp+env	• 72,47	• 71,92	• 69,39	• 67,15	• 66,11	• 66,21	• 62,78	• 60,61
	±2,52	±1,66	±3,10	±2,83	±2,14	±3,37	±2,19	±2,35
BVall+env	• 70,52	• 71,43	• 64,08	• 63,04	• 64,17	• 66,37	• 58,08	• 54,81
	±2,08	±1,57	±2,39	±2,40	±1,85	±2,63	±1,98	±2,40
BVp+env+lact	○ **75,38**	○ **76,39**	○ 71,33	○ **70,52**	○ **69,42**	○ **69,58**	○ 64,82	○ 63,46
	±2,71	±2,20	±3,61	±3,58	±2,80	±3,49	±2,11	±3,06
BVall+env+lact	○ 74,44	• 75,09	• 70,20	• 66,08	○ 68,74	○ 68,97	• 62,13	• 57,60
	±3,19	±2,23	±2,40	±2,93	±3,38	±3,63	±2,71	±1,48
All variables	• 73,89	• 75,22	• 65,05	• 64,56	• 66,99	○ 69,00	• 56,79	• 57,14
	±2,82	±2,11	±3,35	±2,83	±2,68	±3,45	±2,32	±1,61
				(a) Initial classification process				
FSS (C4.5)	○ 75,41$_7$	○ 76,39$_5$	• 71,65$_7$	• 71,69$_5$	○ 69,45$_7$	○ 69,58$_5$	• 65,24$_7$	• 65,66$_5$
	±2,69	±2,20	±3,77	±3,35	±2,85	±3,49	±2,19	±2,86
FSS (Filter-1)	○ 75,32$_6$	• 75,58$_6$	• 71,75$_6$	• 65,21$_6$	○ 69,90$_6$	• 68,61$_6$	• 66,15$_6$	• 59,54$_6$
	±2,84	±1,77	±2,26	±2,27	±3,61	±3,94	±2,69	±2,64
FSS (Filter-2)	• 75,25$_9$	• 75,38$_9$	• 70,72$_9$	• 63,81$_9$	○ 68,74$_9$	• 68,97$_9$	• 63,46$_9$	• 56,88$_9$
	±2,89	±1,82	±2,55	±3,41	±3,25	±3,44	±2,51	±2,30
FSS (FW(0))	• 72,34$_2$	• 71,79$_2$	• 72,27$_2$	• 70,46$_2$	• 66,31$_3$	• 66,73$_2$	• 66,47$_2$	• 63,59$_2$
	±2,07	±1,83	±1,65	±1,55	±3,17	±3,11	±2,84	±2,91
FSS (FW(5))	○ 76,12$_3$	○ 76,48$_9$	**78,23**$_3$	○ 75,22$_3$	○ 70,07$_5$	• **69,91**$_7$	○ 71,46$_3$	○ 67,22$_3$
	±2,14	±1,87	±3,25	±1,84	±2,86	±3,71	±2,87	±3,70
FSS (FW(∞))	○ 76,54$_7$	○ 76,48$_9$	**78,23**$_3$	○ 75,25$_2$	○ 69,97$_4$	○ **69,91**$_7$	○ 71,36$_5$	○ 67,38$_5$
	±2,53	±1,87	±3,25	±2,16	±3,85	±3,71	±2,71	±3,52
FSS (Wrapper)	○ **76,61**$_6$	○ **76,77**$_5$	**78,23**$_3$	○ **75,31**$_5$	○ **70,52**$_5$	• 69,22$_4$	○ **71,69**$_5$	○ **67,38**$_5$
	±1,81	±1,70	±3,25	±1,96	±3,36	±3,54	±3,09	±3,52
				(b) Feature selection process				
FSS (FW(5))	○ 76,81$_{14}$	○ 73,99$_2$	○ 73,96$_1$	○ 74,22$_2$	○ 68,80$_3$	○ 72,56$_{10}$	○ 69,68$_1$	○ 69,16$_1$
	±2,48	±1,99	±2,22	±1,37	±2,55	±2,71	±2,53	±2,77
FSS (FW(∞))	79,20$_{30}$	○ 80,50$_{18}$	78,56$_{10}$	80,21$_{14}$	73,76$_{26}$	76,25$_{42}$	73,82$_{12}$	**74,99**$_{12}$
	±2,19	±1,66	±2,30	±1,56	±2,85	±2,49	±2,23	±2,66
FSS (Wrapper)	▷ **80,04**$_8$	<u>**81,86**$_{18}$</u>	▷ **79,33**$_4$	▷ **80,24**$_6$	**75,25**$_9$	**76,48**$_8$	▷ **74,76**$_5$	74,83$_8$
	±1,54	±1,84	±1,48	±1,83	±2,78	±3,26	±2,90	±2,05
				(c) Attribute construction + FSS				

For future work we aim at using more sophisticated Bayesian classifiers, because even though the results obtained by NB are competitive with those obtained by C4.5, it is clear that NB independence assumption does not hold in this domain. Thus, we plan to use Bayesian networks classifiers in order to allow dependencies among the predictive attributes. Furthermore, we intend to deal with the problem as it is in nature, that is, as a numerical prediction one.

Acknowledgments

References

1. P. Berka and I. Bruha. Discretization and grouping: Preprocessing steps for data mining. In *Proc. of Principles of Data Mining and Knowledge Discovery PKDD'98*, LNAI 1510, pages 239–245. Springer Verlag, 1998.
2. R. Duda and P. Hart. *Pattern classification and scene analysis*. Wiley&Sons, 1973.
3. Elvira-Consortium. Elvira: An environment for creating and using probabilistic graphical models. *Proc. of 1st European Workshop on Probabilistic Graphical Models*, pages 222–230. 2002.
4. I. Farkas (Ed.). Special issue: Artificial intelligence in agriculture. *Computers and Electronics in Agriculture*, 39(1-3), 2003.
5. U. Fayyad and K.B. Irani. Multi-interval discretization of continuous-valued attributes for classification learning. In *Proc. of Int. Joint Conf. on Artificial Intelligence*, pages 1022–1029, 1993.
6. U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17:37–54, 1996.
7. J.A. Gámez. Mining the esrom: A study of breeding value prediction in manchego sheep by means of classification techniques plus attribute selection and construction. Technical Report DIAB-05-01-3, Universidad de Castilla-La Mancha, 2005.
8. I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.
9. R.C. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11:63–91, 1993.
10. H. Liu and H. Motoda. *Feature Extraction Construction and Selection: a data mining perspective*. Kluwer Academic Publishers, 1998.
11. C. Matheus and L. Rendell. Constructive induction on decision trees. In *Proc. of the International Joint conference on Artificial Intelligence*, pages 645–650. 1989.
12. Tom M. Mitchell. *Machine Learning*. McGraw-Hill, 1997.
13. J.R. Quinlan. Induction of decision trees. *Machine Learning*, 1:81–106, 1986.
14. I.H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, 2000.

Learning Method for Automatic Acquisition of Translation Knowledge

Hiroshi Echizen-ya[1], Kenji Araki[2], and Yoshio Momouchi[1]

[1] Dept. of Electronics and Information, Hokkai-Gakuen University,
S26-Jo, W11-Chome, Chuo-ku, Sapporo, 064-0926 Japan
{echi, momouchi}@eli.hokkai-s-u.ac.jp
Tel: +81-11-841-1161, Fax: +81-11-551-2951

[2] Graduate School of Information Science and Technology, Hokkaido University,
N14-Jo, W9-Chome, Kita-ku, Sapporo, 060-0814 Japan
araki@media.eng.hokudai.ac.jp
Tel: +81-11-706-6534, Fax: +81-11-709-6277

Abstract. This paper presents a new learning method for automatic acquisition of translation knowledge from parallel corpora. We apply this learning method to automatic extraction of bilingual word pairs from parallel corpora. In general, similarity measures are used to extract bilingual word pairs from parallel corpora. However, similarity measures are insufficient because of the sparse data problem. The essence of our learning method is this presumption: in local parts of bilingual sentence pairs, the equivalents of words that adjoin the source language words of bilingual word pairs also adjoin the target language words of bilingual word pairs. Such adjacent information is acquired automatically in our method. We applied our method to systems based on various similarity measures, thereby confirming the effectiveness of our method.

1 Introduction

1.1 Problem in Similarity Measures

Bilingual word pairs - pairs of source language (SL) words and target language (TL) words - are extremely important as translation knowledge in the field of machine translation. However, manual extraction by humans of bilingual word pairs of various languages is costly. For that reason, automatic extraction using a system of bilingual word pairs from parallel corpora is effective. Similarity measures [1] are often used to extract bilingual word pairs from parallel corpora because such measures are language independent. However, similarity measures are insufficient because of the sparse data problem. For example, a system based on cosines [1] would seek to extract (letter; 手紙 [$tegami^1$]) from (I'd like to sent this letter to Japan.; この/手紙/を/日本/に/送り/たい/の/です.[$kono\ tegami\ wo\ nippon\ ni\ okuri\ tai\ no\ desu.$]). The cosine is the effective similarity measure. The cosine is defined as

$$Cosine(W_S, W_T) = \frac{a}{\sqrt{(a+b)(a+c)}} \tag{1}$$

---

[1] Italics means Japanese pronunciation.

In function (1), 'a' is the number of pieces in which both the SL word W_S and TL word W_T are found, 'b' is the number of pieces in which only the W_S is found, and 'c' is the number of pieces in which only the W_T is found.

This system based on the cosine cannot clearly extract (letter; 手紙 [*tegami*]) as a correct bilingual word pair when the respective frequencies of "letter", "手紙 [*tegami*]" and "日本 [*nippon*]" are 1. In that case, the cosine score of "letter" and "手紙 [*tegami*]" becomes $1.0(= \frac{1}{\sqrt{1\times1}})$. The cosine score between "letter" and "日本 [*nippon*]" also becomes $1.0(= \frac{1}{\sqrt{1\times1}})$. Therefore, the system based on the cosine cannot exclusively select (letter; 手紙 [*tegami*]) from among two bilingual word pairs (letter; 手紙 [*tegami*]) and (letter; 日本 [*nippon*]). That is, when several bilingual word pairs have resemblant similarity value candidates, the system based on similarity measures falls into the sparse data problem.

1.2 Motivation

To solve the sparse data problem, we use of this inference: in local parts of bilingual sentence pairs (*e.g.*, phrases, not sentences), the equivalents of words that adjoin the SL words of bilingual word pairs also adjoin the TL words of bilingual word pairs. For example, in (I'd like to send this letter to Japan.; この/手紙/を/日本/に/送り/たい/の/です.[*kono tegami wo nippon ni okuri tai no desu.*]), the system uses the information that "this" which adjoins "letter" corresponds to "この [*kono*]." Moreover, it uses the information that equivalents of words that adjoin the right side of "this" exist on the right side "この [*kono*]" in TL sentences. Consequently, only (letter; 手紙 [*tegami*]) can be extracted as a bilingual word pair. That is, using such adjacent information, the system can limit the search scope for the decision of equivalents in bilingual sentence pairs by extracting only those word pairs that adjoin the adjacent information.

Moreover, the adjacent information is acquired automatically from the perspective of learning [2]. We call this learning method **A**djacent **I**nformation **L**earning (AIL). In this study, we apply AIL to four systems based on the cosine, the Dice coefficient, the **L**og-**L**ikelihood **R**atio (LLR), and Yates' χ^2 to efficiently extract bilingual word pairs as translation knowledge from five kinds of parallel corpora: English – Japanese, French – Japanese, German – Japanese, Shanghai-Chinese – Japanese, and Ainu[4] – Japanese parallel corpora.

2   Outline

Figure 1 shows an outline of the system using AIL. In Fig. 1, AIL corresponds to three processes: the process based on templates, the process based on two bilingual sentence pairs, and the decision process of bilingual word pairs.

First, the user inputs the SL words of bilingual word pairs. In the process based on templates, the system extracts bilingual word pairs using the templates

---

[4] Ainu language is spoken by members of the Ainu ethnic group, who mainly reside in northern Japan and Sakhalin. It is independent, but similar to Japanese or Korean.

Fig. 1. Process flow.

in the dictionary for templates. In this study, templates are designated as the rules for extracting new bilingual word pairs. Similarity values between SL words and TL words in all extracted bilingual word pairs are assigned. The similarity value is defined by function (2) based on the Dice coefficient.

$$sim(W_S, W_T) = \frac{2 \times f_{ST}}{f_S + f_T} \qquad (2)$$

In function (2), f_{ST} is the number of pieces in which both the SL word W_S and the TL word W_T are found, f_S is the number of pieces in which the W_S are found, and f_T is the number of pieces in which the W_T are found. In the process based on two bilingual sentence pairs, the system obtains bilingual word pairs and new templates from two bilingual sentence pairs. Similarity values in all acquired templates are also assigned by function (2). Moreover, during the decision process of bilingual word pairs, the system chooses the most suitable bilingual word pairs using their similarity values when several candidates of bilingual word pairs exist. The system compares the similarity values of chosen bilingual word pairs with a threshold value. Consequently, the system registers the chosen bilingual word pairs to the dictionary for bilingual word pairs when their respective similarity values are greater than the threshold value.

The system extracts bilingual word pairs without AIL in the process based on similarity measures when their similarity values are not greater than the threshold value or when no bilingual word pairs are extracted. Moreover, the extracted bilingual word pairs can be registered into the dictionary efficiently using a morphological analysis system to very minute changes in spellings or words or pronunciation. The system can extract bilingual word pairs even when the scripts of two languages are same because AIL is language independent.

3   Adjacent Information Learning (AIL)

3.1   Process Based on Two Bilingual Sentence Pairs

The system obtains bilingual word pairs and templates using common parts between two bilingual sentence pairs. That is, the bilingual word pairs and the templates can be acquired easily only from a parallel corpus using common parts for which the frequencies are very low, *i.e.* 2. Figure 2 shows examples of extraction of a bilingual word pair and acquisition of a template.

Input word：**the British Museum**
Bilingual sentence pair 1
 (I'm thinking of going *to* **the British Museum**.
                 ；**大/英/博物館**/に/行っ/て/みよ/う/と/思っ/てい/ます.
                 [*dai ei hakubutsukan ni itte miyou to omotte i masu.*])
Bilingual sentence pair 2
 (When should I return it *to* you?；**あなた**/に/いつ/お返し/すれ/ば/いい/です/か？
                 [*anata ni itsu okaeshi sure ba ii desu ka?*])

Noun bilingual word pair：(the British Museum；大/英/博物館 [*dai ei hakubutsukan*])
 Similarity value：1.0

 (I'm thinking of going *to* **@**.；**@**/に/行っ/て/みよ/う/と/思っ/てい/ます.
                 [*@ ni itte miyou to omotte i masu.*])

   Template：(to @；@/に [*@ ni*])   Similarity value：0.50

Fig. 2. An example of the process based on two bilingual sentence pairs.

First, the system selects bilingual sentence pair 1, for which "the British Museum" exists. Furthermore, the system selects bilingual sentence pair 2 for which "to" that adjoins "the British Museum" exists and for which "に [*ni*]" exists as a common part between two TL sentences. The system extracts "大/英/博物館 [*dai ei hakubutsukan*]", which adjoins the common part "に [*ni*]", from the TL sentence of bilingual sentence pair 1. On the other hand, "行っ/て/みよ/う/と/思っ/て/い/ます [*itte miyou to omotte i masu*]", adjoins "に [*ni*]", is not extracted because it is not the word. Consequently, (the British Museum; 大/英/博物館 [*dai ei hakubutsukan*]), is obtained as a correct noun bilingual word pair. Moreover, the system replaces "the British Museum" and "大/英/博物館 [*dai ei hakubutsukan*]" with the variable "@" in bilingual sentence pair 1, and obtains (to @;@/に [*@ ni*]) as template by combining the common parts "to", "に [*ni*]" and variable "@." Similarity values in all obtained bilingual word pairs and templates are assigned by function (2). On the other hand, when several common parts exist, the system extracts the parts between two common parts from TL sentences. As a result, when several candidates of bilingual word pairs are obtained, the system chooses most suitable bilingual word pairs in the decision process of bilingual word pairs as described in section 3.3.

3.2   Process Based on Templates

The system can extract bilingual word pairs efficiently using templates that have adjacent information. Figure 3 shows an example of extraction of a bilingual word pair using templates. In Fig. 3, (eat; 食べ [*tabe*]) was extracted as the verb bilingual word pair using the template (to @;@/に [@ *ni*]). This template (to @;@/に [@ *ni*]) has information that the equivalents of words that adjoin the right side of "to" exist on the left side "に [*ni*]" in TL sentences. Therefore, only (eat; 食べ [*tabe*]) was easily extracted. The use of such templates is effective to solve the sparse data problem because it can limit the search scope for the decision of equivalents in bilingual sentence pairs.

Input word : **eat**

　　　　　　Template　　　　　　　　(*to* **@** ; **@**/*に* [**@** *ni*])

Bilingual sentence pair

(After the test, we all went out for something *to* **eat**. :試験/の/後/で/、/みんな/で/**食べ**/*に*/出かけ/た/ん/です.

　　　　　　　　　　　　　　　　　[*shiken no ato de , minna de **tabe** ni dekake ta n desu.*])

Verb bilingual word pair :　　　(**eat**;**食べ [tabe]**)　　Similarity value : 0.54

Fig. 3. An example of use of templates.

3.3   Decision Process for Bilingual Word Pairs

The most suitable bilingual word pairs are selected according to their similarity values when several bilingual word pairs have been extracted. That is, the extracted bilingual word pairs are sorted so that the bilingual word pairs that have the highest similarity values are ranked at the top. Moreover, when several bilingual word pairs with equal similarity-value candidates exist, the system selects the bilingual word pairs that appear for the first time in a parallel corpus.

4   Process Based on Similarity Measures

The system extracts bilingual word pairs using only the cosine, the Dice coefficient [3], LLR [4], or Yates' χ^2 [5] without AIL when the similarity values are not greater than the threshold value or when no bilingual word pairs are extracted. Moreover, the system chooses the bilingual word pairs that appear in the parallel corpus for the first time when several bilingual word-pair candidates are obtained.

5   Experiments for Performance Evaluation

5.1   Experimental Procedure

Five kinds of parallel corpora were used in this paper as experimental data. These parallel corpora are for English – Japanese, French – Japanese, German –

Japanese, Shanghai-Chinese – Japanese and Ainu – Japanese. They were taken from textbooks containing conversational sentences. The number of bilingual sentence pairs was 1,794. To confirm AIL's effectiveness, we inputted all 1,081 SL words of nouns, verbs, adjectives, adverbs, and conjunctions into the system based on the cosine, the system based on the cosine in which AIL is applied as described in section 2 (herein, we respectively call it the system based on the cosine+AIL), the system based on the Dice coefficient, the system based on Dice+AIL, the system based on LLR, the system based on LLR+AIL, the system based on Yates' χ^2, and the system based on Yates+AIL. The initial conditions of all dictionaries are empty in those respective systems. We repeated experiments for each parallel corpus using respective systems. The system using AIL uses 0.5[2] as its best threshold value. Moreover, we evaluated whether correct bilingual word pairs are obtained or not, and calculated the extraction rates for all 1,081 SL words.

5.2   Experiments and Discussion

Table 1 shows experimental results. The respective extraction rates of systems using AIL were more than 8.0, 8.0, 6.7 and 6.1 percentage points higher than those of the systems based on the cosine, the Dice coefficient, LLR, and Yates' χ^2. These results indicate that our method is effective for various similarity measures. For example, in Fig. 3, the system without AIL must select "食べ [tabe]" as a correct equivalent from many noun and verb candidates "試験 [shiken]", "後 [ato]", "みんな [minna]", "食べ [tabe]", and "出かけ [dekake]". In contrast, the system using AIL can easily select only "食べ [tabe]" using the acquired template (to @;@/に [@ ni]).

Table 1. Results of evaluation experiments.

SL	cosine	cosine +AIL	Dice coefficient	Dice +AIL	LLR	LLR +AIL	Yates' χ^2	Yates +AIL
English	52.1%	61.5%	49.7%	58.0%	52.7%	60.4%	53.8%	59.8%
French	50.8%	58.8%	47.9%	56.7%	54.6%	61.3%	55.4%	60.4%
German	52.3%	57.9%	53.3%	61.0%	54.4%	59.5%	53.3%	58.5%
Shanghai-Chinese	56.8%	63.6%	54.9%	62.9%	57.6%	63.3%	57.6%	62.5%
Ainu	52.6%	62.9%	54.0%	61.5%	53.1%	62.0%	52.1%	62.0%
Total	53.1%	**61.1%**	52.1%	**60.1%**	54.7%	**61.4%**	54.7%	**60.8%**

Moreover, we investigated the extraction rates for which the frequencies are 1. In the systems without AIL, many extracted bilingual word pairs for which the frequencies are 1 were erroneous bilingual word pairs because of data sparseness problems, as described in section 1.1. Therefore, improvement of the extraction rates of such bilingual word pairs indicates that AIL is effective to solve the

---

[2] This value was obtained through preliminary experiments.

sparse data problem. The respective extraction rates of the bilingual word pairs for which the frequencies are 1 improved 11.3, 11.0, 10.9 and 9.7 percentage points using AIL.

Among previous studies, one [6] uses the co-occurrence of words depending on the number of co-occurrence words and their frequency. Such a method is insufficient in terms of efficient extraction of bilingual word pairs. In contrast, the system using AIL requires only a one-word string as the co-occurrence word, *e.g.* only "to" in Fig. 3. Moreover, the system using AIL can extract bilingual word pairs even when the frequencies of the pairs of the co-occurrence words and the bilingual word pairs are only 1, *e.g.*, "to eat" in bilingual sentence pair of Fig. 3. In a study [7] that acquires low-frequency bilingual terms, bilingual dictionary and MT systems are used for measuring similarity. Therefore, it is difficult to deal with various languages because of the use of large-scale translation knowledge.

6 Conclusion

This paper presented **A**djacent **I**nformation **L**earning (AIL) as a new learning method for solving the sparse data problem in similarity measures. Results showed that AIL is effective for various similarity measures. It is also effective as a solution to the sparse data problem. Future studies will apply this method to a multilingual machine translation system.

References

1. Manning, C. D. and Schütze, H. 1999. Foundations of Statistical Natural Language Processing. MIT Press.
2. Echizen-ya, H., K. Araki, Y. Momouchi, and K. Tochinai. 2002. Study of Practical Effectiveness for Machine Translation Using Recursive Chain-link-type Learning. In *Proceedings of COLING '02*, pp.246–252.
3. Smadja, F., K. R. McKeown and V. Hatzivassiloglou. 1996. Translating Collocations for Bilingual Lexicons: A Statistical Approach. *Computational Linguistics*, vol.22, no.1, pp.1–38.
4. Dunning, T. E. 1993. Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, vol.19, no.1, pp.61–74.
5. Hisamitsu, T. and Y. Niwa. 2001. Topic-Word Selection Based on Combinatorial Probability. In *NLPRS'01*, pp.289–296.
6. Kaji, H. and T. Aizono. 1996. Extracting Word Correspondences from Bilingual Corpora Based on Word Co-occurrence Information. In *Proceedings of COLING'96*, pp.23–28.
7. Utsuro, T., K. Hino, and M. Kida. 2004 Integrating Cross-Lingually Relevant News Articles and Monolingual Web Documents in Bilingual Lexicon Acquisition. In *Proceedings of COLING'04*, pp.1036–1042.

Author Index