

Vinu V Das
Gylson Thomas
Ford Lumban Gaol (Eds.)

Communications in Computer and Information Science

147

Information Technology and Mobile Communication

International Conference, AIM 2011
Nagpur, Maharashtra, India, April 2011
Proceedings

Vinu V Das Gylson Thomas
Ford Lumban Gaol (Eds.)

Information Technology and Mobile Communication

International Conference, AIM 2011
Nagpur, Maharashtra, India, April 21-22, 2011
Proceedings

Volume Editors

Vinu V Das
ACEEE, Trivandrum, Kerala, India
E-mail: vinuvdas@theaceee.org

Gylson Thomas
MES College of Engineering, Kuttippuram, Kerala, India
E-mail: gylson_thomas@yahoo.com

Ford Lumban Gaol
Binus University, Jakarta, Indonesia
E-mail: fordlg@gmail.com

ISSN 1865-0929
ISBN 978-3-642-20572-9
DOI 10.1007/978-3-642-20573-6
Springer Heidelberg Dordrecht London New York

e-ISSN 1865-0937
e-ISBN 978-3-642-20573-6

Library of Congress Control Number: 2011925374

CR Subject Classification (1998): C.2, D.2, H.4, H.3, I.2.11, K.4.4

© Springer-Verlag Berlin Heidelberg 2011

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The International Conference on Advances in Information Technology and Mobile Communication (AIM 2011) was sponsored and organized by The Association of Computer Electronics and Electrical Engineers (ACEEE) and held at Nagpur, Maharashtra, India during April 21-22, 2011.

The mission of the AIM International Conference is to bring together innovative academics and industrial experts in the field of computer science, information technology, computational engineering, mobile communication and security to a common forum, where a constructive dialog on theoretical concepts, practical ideas and results of the state of the art can be developed. In addition, the participants of the symposium have a chance to hear from renowned keynote speakers. We would like to thank the Program Chairs, organization staff, and the members of the Program Committees for their hard work this year. We would like to thank all our colleagues who served on different committees and acted as reviewers to identify a set of high-quality research papers for AIM 2011.

The conference received 313 submissions overall. Only 92 papers were accepted and registered for the AIM 2011 proceedings. We also thank Alfred Hofmann, Janahanlal Stephen, and Gylson Thomas for the constant support and guidance. We would like to express our gratitude to the Springer LNCS-CCIS editorial team, especially Ms. Leonie Kunz, for producing such a wonderful quality proceedings book.

February 2011

Vinu V. Das

AIM 2011 – Organization

Honorary Chairs

Shuvra Das
Jiguo Yu

University of Detroit Mercy, USA
Qufu Normal University, China

Technical Chairs

Sumeet Dua
Vijayakumar
Amit Banerjee

Louisiana Tech University, USA
MG University, India
The Pennsylvania State University, USA

Technical Co-chairs

Natarajan Meghanathan
Gylson Thomas
Hicham Elzabadani
Shahrokh Valaee

Jackson State University, USA
MES College of Engineering, India
American University in Dubai
University of Toronto, Canada

General Chair

Janahanlal Stephen

ILAHIA College of Engineering, India

Organizing Chair

Vinu V. Das

The IDES

Organizing Co-chairs

T.S.B. Sudarshan
Ford Lumban Gaol

BITS Pilani, India
University of Indonesia

Publicity Chairs

Amlan Chakrabarti
Prafulla Kumar Behera,

University of Calcutta, India
Utkal University, India

Publication Chairs

Vijayakumar	NSS Engineering College, India
T.S.B. Sudarshan	BITS Pilani, India
K.P. Soman	Amritha University, India
N. Jaisankar	VIT University, India

Program Committee Chairs

Harry E. Ruda	University of Toronto, Canada
Deepak Laxmi Narasimha	University of Malaya, Malaysia
N. Nagarajan	Anna University, Coimbatore, India
Akash Rajak	Krishna Institute of Engineering and Technology, UP, India
M. Ayoub Khan	CDAC, NOIDA, India

Table of Contents

Full Paper

Efficient Object Motion Prediction Using Adaptive Fuzzy Navigational Environment	1
<i>Vijay S. Rajpurohit and M.M. Manohara Pai</i>	
An Efficient Protocol Using Smart Interval for Coordinated Checkpointing	6
<i>Jagdish Makhijani, Manoj Kumar Niranjani, Mahesh Motwani, A.K. Sachan, and Anil Rajput</i>	
Face Recognition System Using Discrete Wavelet Transform and Fast PCA	13
<i>K. Ramesha and K.B. Raja</i>	
Mining Indirect Association between Itemsets	19
<i>B. Ramasubbareddy, A. Govardhan, and A. Ramamohanreddy</i>	
Reaction Attacks in the Matrix Scheme of NTRU Cryptosystem	27
<i>Rakesh Nayak, Jayaram Pradhan, and C.V. Sastry</i>	
Process Corner Analysis for Folding and Interpolating ADC	33
<i>Shruti Oza and N.M. Devashrayee</i>	
Fast Near-Lossless Image Compression with Tree Coding Having Predictable Output Compression Size	39
<i>Soumik Banerjee and Debashish Chakroborty</i>	
Over Load Detection and Admission Control Policy in DRTDBS	45
<i>Nuparam and Udai Shanker</i>	
Wavelet Transform Based Image Registration and Image Fusion	55
<i>Manjusha Deshmukh and Sonal Gahankari</i>	
60 GHz Radio Channel Characteristics in an Indoor Environment for Home Entertainment Networks	61
<i>T. Rama Rao, S. Ramesh, and D. Murugesan</i>	
Improved Back Propagation Algorithm to Avoid Local Minima in Multiplicative Neuron Model	67
<i>Kavita Burse, Manish Manoria, and Vishnu Pratap Singh Kirar</i>	
Technical White Paper on “Time and Frequency Synchronization in OFDM”	74
<i>Anagha Rathkanthiwar and Mridula Korde</i>	

Cell-ID Based Vehicle Locator and Real-Time Deactivator Using GSM Network	82
<i>Nilesh Dubey, Vandana Dubey, and Shivangi Bande</i>	
A Novel Design of Reconfigurable Architecture for Multistandard Communication System	87
<i>T. Suresh and K.L. Shunmuganathan</i>	
Two Novel Long-Tail Pair Based Second Generation Current Conveyors (CCII)	95
<i>Amisha Naik and N.M. Devashrayee</i>	
Generating Testcases for Concurrent Systems Using UML State Chart Diagram	100
<i>Debashree Patnaik, Arup Abhinna Acharya, and Durga P. Mohapatra</i>	
Intelligent Agent Based Resource Sharing in Grid Computing	106
<i>V.V. Srinivas and V.V. Varadhan</i>	
Wideband Miniaturized Patch Antenna Design and Comparative Analysis	111
<i>Sanket Patel, Yogeshwar Kosta, Himanshu Soni, and Shobhit Patel</i>	
Predicting Number of Zombies in a DDoS Attack Using ANN Based Scheme	117
<i>B.B. Gupta, R.C. Joshi, M. Misra, A. Jain, S. Juyal, R. Prabhakar, and A.K. Singh</i>	
A Novel Biometric Watermarking Approach Using LWT- SVD	123
<i>Meenakshi Arya and Rajesh Siddavatam</i>	
Detection and Prevention of Phishing Attack Using Dynamic Watermarking	132
<i>Akhilendra Pratap Singh, Vimal Kumar, Sandeep Singh Sengar, and Manoj Wairiya</i>	
A Search Tool Using Genetic Algorithm	138
<i>M.K. Thanuja and C. Mala</i>	
Heterogeneous Data Mining Environment Based on DAM for Mobile Computing Environments	144
<i>Ashutosh K. Dubey, Ganesh Raj Kushwaha, and Nishant Shrivastava</i>	
Selection of Views for Materialization Using Size and Query Frequency	150
<i>T.V. Vijay Kumar and Mohammad Haider</i>	
Key Validation Using Weighted-Edge Web of Trust Model	156
<i>Sumit Kumar, Nahar Singh, and Ashok Singh Sairam</i>	

A Novel Reconfigurable Architecture for Enhancing Color Image Based on Adaptive Saturation Feedback	162
<i>M.C. Hanumantharaju, M. Ravishankar, D.R. Rameshbabu, and S. Ramachandran</i>	
Signal Processing Approach for Prediction of Kink in Transmembrane α -Helices	170
<i>Jayakishan K. Meher, Nibedita Mishra, Pranab Kishor Mohapatra, Mukesh Kumar Raval, Pramod Kumar Meher, and Gananath Dash</i>	
Cascaded H-Bridge Multilevel Boost Inverter without Inductors for Electric/Hybrid Electric Vehicle Applications	178
<i>S. Dhayanandh, A.P. Ramya Sri, S. Rajkumar, and N. Lavanya</i>	
Design of Microstrip Meandered Patch Antenna for Mobile Communication	184
<i>Shobhit Patel, Jaymin Bhalani, Yogesh Kosta, and Sanket Patel</i>	
Building Gaussian Mixture Shadow Model for Removing Shadows in Surveillance Videos	190
<i>Archana Chougule and Pratap Halkarnikar</i>	
FAutoREDWithRED: To Increase the Overall Performance of Internet Routers	196
<i>K. Chitra and G. Padmavathi</i>	
Short Paper	
Scan Based Sequential Circuit Testing Using DFT Advisor	203
<i>P. Reshma</i>	
Rate Adaptive Distributed Source-Channel Coding Using IRA Codes for Wireless Sensor Networks	207
<i>Saikat Majumder and Shrish Verma</i>	
Web Cam Motion Detection Surveillance System Using Temporal Difference and Optical Flow Detection with Multi Alerts	214
<i>V.D. Ambeth Kumar and M. Ramakrishan</i>	
Rts-Mirror: Real Time Synchronized Automated Rear Vision Mirror System	222
<i>Kuldeep Verma, Ankita Agarkar, and Apoorv Joshi</i>	
Fuzzy Based PSO for Software Effort Estimation	227
<i>P.V.G.D. Prasad Reddy and CH.V.M.K. Hari</i>	
SLV: Sweep Line Voronoi Ad Hoc Routing Algorithm	233
<i>E. Rama Krishna, A. Venkat Reddy, N. Rambabu, and G. Rajesh Kumar</i>	

Hybrid Routing for Ad Hoc Wireless Networks	240
<i>Ravilla Dilli, R.S. Murali Nath, and P. Chandra Shekar Reddy</i>	
Implementation of ARINC 429 16 Channel Transmitter Controller on FPGA	245
<i>Debasis Mukherjee, Niti Kumar, Kalyan Singh, Hemanta Mondal, and B.V.R. Reddy</i>	
Segmentation of Image Using Watershed and Fast Level Set Methods . . .	248
<i>Minal M. Puranik and Shobha Krishnan</i>	
Tree Structured, Multi-hop Time Synchronization Approach in Wireless Sensor Networks	255
<i>Surendra Rahamatkar, Ajay Agarwal, Praveen Sen, and Arun Yadav</i>	
A New Markov Chain Based Cost Evaluation Metric for Routing in MANETs	259
<i>Abhinav Tiwari, Nisha Wadhawan, and Neeraj Kumar</i>	
Texture Image Classification Using Gray Level Weight Matrix (GLWM)	263
<i>R.S. Sabeenian and P.M. Dinesh</i>	
Formal Verification of IEEE802.11i WPA-GPG Authentication Protocol	267
<i>K.V. Krishnam Raju and V. Valli Kumari</i>	
Adaptive Steganography Based on Covariance and Dct	273
<i>N. Sathisha, Swetha Sreedharan, R. Ujwal, Kiran D'sa, Aneeshwar R. Danda, K. Suresh Babu, K.B. Raja, K.R. Venugopal, and L.M. Patnaik</i>	
Image Segmentation Using Grey Scale Weighted Average Method and Type-2 Fuzzy Logic Systems	277
<i>Saikat Maity and Jaya Sil</i>	
Cluster Analysis and Pso for Software Cost Estimation	281
<i>Teggyot Singh Sethi, CH.V.M.K. Hari, B.S.S. Kaushal, and Abhishek Sharma</i>	
Controlling Crossover Probability in Case of a Genetic Algorithm	287
<i>Parama Bagchi and Shantanu Pal</i>	
A Qualitative Survey on Unicast Routing Algorithms in Delay Tolerant Networks	291
<i>Sushovan Patra, Anerudh Balaji, Sujoy Saha, Amartya Mukherjee, and Subrata Nandi</i>	

Designing and Modeling of CMOS Low Noise Amplifier Using a Composite MOSFET Model Working at Millimeter-Wave Band	297
<i>Adhira Raj, Karthigha Balamurugan, and M. Jayakumar</i>	
Designing Dependable Business Intelligence Solutions Using Agile Web Services Mining Architectures	301
<i>A. V. Krishna Prasad, S. Ramakrishna, B. Padmaja Rani, M. Upendra Kumar, and D. Shravani</i>	
A Modified Continuous Particle Swarm Optimization Algorithm for Uncapacitated Facility Location Problem	305
<i>Sujay Saha, Arnab Kole, and Kashinath Dey</i>	
Design of Hybrid Genetic Algorithm with Preferential Local Search for Multiobjective Optimization Problems	312
<i>J. Bhuvana and C. Aravindan</i>	
Synergy of Multi-agent Coordination Technique and Optimization Techniques for Patient Scheduling	317
<i>E. Grace Mary Kanaga and M.L. Valarmathi</i>	
Naive Bayes Approach for Website Classification	323
<i>R. Rajalakshmi and C. Aravindan</i>	
Method to Improve the Efficiency of the Software by the Effective Selection of the Test Cases from Test Suite Using Data Mining Techniques	327
<i>Lilly Raamesh and G.V. Uma</i>	
A Hybrid Intelligent Path Planning Approach to Cooperative Robots	332
<i>K. Prasad, Vinodh P. Vijayan, and Biju Paul</i>	
Towards Evaluating Resilience of SIP Server under Low Rate DoS Attack	336
<i>Abhishek Kumar, P. Shanthy Thilagam, Alwyn R. Pais, Vishwas Sharma, and Kunal M. Sadalkar</i>	
Poster Paper	
Performance Analysis of a Multi Window Stereo Algorithm on Small Scale Distributed Systems: A Message Passing Environment	340
<i>Vijay S. Rajpurohit and M.M. Manohara Pai</i>	
Ant Colony Algorithm in MANET-Review and Alternate Approach for Further Modification	344
<i>Jyoti Jain, Roopam Gupta, and T.K. Bandhopadhyay</i>	

RDCLRP-Route Discovery by Cross Layer Routing Protocol for Manet Using Fuzzy Logic	348
<i>Mehajabeen Fatima, Roopam Gupta, and T.K. Bandhopadhyay</i>	
A New Classification Algorithm with GLCCM for the Altered Fingerprints	352
<i>R. Josphineleela and M. Ramakrishnan</i>	
Footprint Based Recognition System	358
<i>V.D. Ambeth Kumar and M. Ramakrishnan</i>	
An Efficient Approach for Data Replication in Distributed Database Systems	368
<i>Arun Kumar Yadav, Ajay Agarwal, and S. Rahmatkar</i>	
Unicast Quality of Service Routing in Mobile Ad Hoc Networks Based on Neuro-fuzzy Agents	375
<i>V.R. Budyal, S.S. Manvi, and S.G. Hiremath</i>	
Performance Comparison of Routing Protocols in Wireless Sensor Networks	379
<i>Geetika Ganda, Prachi, and Shaily Mittal</i>	
Security and Trust Management in MANET	384
<i>Akash Singh, Manish Maheshwari, Nikhil, and Neeraj Kumar</i>	
Analytical Parametric Evaluation of Dynamic Load Balancing Algorithms in Distributed Systems	388
<i>Mayuri A. Mehta and Devesh C. Jinwala</i>	
Wavelet Based Electrocardiogram Compression at Different Quantization Levels	392
<i>A. Kumar and Ranjeet</i>	
Content Based Image Retrieval by Using an Integrated Matching Technique Based on Most Similar Highest Priority Principle on the Color and Texture Features of the Image Sub-blocks	399
<i>Ch. Kavitha, M. Babu Rao, B. Prabhakara Rao, and A. Govardhan</i>	
Understanding the Impact of Cache Performance on Multi-core Architectures	403
<i>N. Ramasubramaniam, V.V. Srinivas, and P. Pavan Kumar</i>	
An Overview of Solution Approaches for Assignment Problem in Wireless Telecommunication Network	407
<i>K. Rajalakshmi and M. Hima Bindu</i>	
Design of Domain Specific Language for Web Services QoS Constraints Definition	411
<i>Monika Sikri</i>	

Modified Auxiliary Channel Diffie Hellman Encrypted Key Exchange Authentication Protocol	417
<i>Nitya Ramachandran and P. Yogesh</i>	
Bilateral Partitioning Based Character Recognition for Vehicle License Plate	422
<i>Siddhartha Choubey, G.R. Sinha, and Abha Choubey</i>	
Strategies for Parallelizing KMeans Data Clustering Algorithm	427
<i>S. Mohanavalli, S.M. Jaisakthi, and C. Aravindan</i>	
A Comparative Study of Different Queuing Models Used in Network Routers for Congestion Avoidance.....	431
<i>Narendran Rajagopalan and C. Mala</i>	
SAR Image Classification Using PCA and Texture Analysis	435
<i>Mandeep Singh and Gunjit Kaur</i>	
Performance of WiMAX/ IEEE 802.16 with Different Modulation and Coding	440
<i>Shubhangi R. Chaudhary</i>	
A Novel Stair-Case Replication (SCR) Based Fault Tolerance for MPI Applications.....	445
<i>Sanjay Bansal, Sanjeev Sharma, and Ishita Trivedi</i>	
Evaluating Cloud Platforms- An Application Perspective.....	449
<i>Pankaj Deep Kaur and Inderveer Chana</i>	
An Intelligent Agent Based Temporal Action Status Access Control Model for XML Information Management	454
<i>N. Jaisankar and A. Kannan</i>	
An Economic Auction-Based Mechanism for Multi-service Overlay Multicast Networks	461
<i>Mohammad Hossein Rezvani and Morteza Analoui</i>	
Server Virtualization: To Optimizing Messaging Services by Configuring Front-End and Back-End Topology Using Exchange Server in Virtual Environments	468
<i>R. Anand and T. Deenadayalan</i>	
Blind Source Separation for Convolutional Audio Mixing	473
<i>V. Jerine Rini Rosebell, D. Sugumar, Shindu, and Sherin</i>	
ICA Based Informed Source Separation for Digitally Watermarked Audio Signals	477
<i>R. Sharanya, D. Sugumar, T.L. Sujithra, Susan Mary Bose, and Divya Mary Koshy</i>	

Evaluation of Retrieval System Using Textural Features Based on Wavelet Transform	481
<i>Lidiya Xavier and I. Thusnavis Bella Mary</i>	
Behavioural Level Watermarking Techniques for IP Identification Based on Testing in SOC Design.....	485
<i>Newton david Raj, Josprakash, AntopremKumar, Daniel, and Joshua Thomas</i>	
Rough Set Approach for Distributed Decision Tree and Attribute Reduction in the Disseminated Environment	489
<i>E. Chandra and P. Ajitha</i>	
MIMO and Smart Antenna Technologies for 3G and 4G	493
<i>Vanitha Rani Rentapalli and Zafer Jawed Khan</i>	
A GA-Artificial Neural Network Hybrid System for Financial Time Series Forecasting	499
<i>Binoy B. Nair, S. Gnana Sai, A.N. Naveen, A. Lakshmi, G.S. Venkatesh, and V.P. Mohandas</i>	
A Preemptive View Change for Fault Tolerant Agreement Using Single Message Propagation	507
<i>Poonam Saini and Awadhesh Kumar Singh</i>	
A Model for Detection, Classification and Identification of Spam Mails Using Decision Tree Algorithm	513
<i>Hemant Pandey, Bhasker Pant, and Kumud Pant</i>	
Author Index	517

Efficient Object Motion Prediction Using Adaptive Fuzzy Navigational Environment

Vijay S. Rajpurohit¹ and M.M. Manohara Pai²

¹ Gogte Institute of Institute of Technology, Belgaum, India
vijaysr2k@yahoo.com

² Manipal Institute of Institute of Technology, Manipal, India
mm.pai@rediffmail.com

Abstract. This paper proposes an adaptive Fuzzy rule based motion prediction algorithm for predicting the next instance position of a moving object. The prediction algorithm is tested for real-life bench-marked data sets and compared with existing motion prediction techniques. Results of the study indicate that the performance of the predictor is comparable to the existing prediction methods.

Keywords: Short Term Motion prediction, Rule base Optimization, Fuzzy Predictor algorithm, Adaptive navigational Environment.

1 Introduction

Short term Object motion prediction in a dynamic Robot navigation environment refers to, the prediction of next instance position of a moving object based on the previous history of its motion. Research literature has addressed solutions to the short term object motion predictions with different methods such as, Curve fitting or Regression methods, Neural network based approaches, Hidden Markov stochastic models, Bayesian Occupancy Filters, Extended Kalman Filter and Stochastic prediction model[1][2][4][6][7]. The design of a navigational model in an automated mobile Robot system is influenced by its specific applications, the environment in which it operates and the sensory system. Many navigational model representations have been proposed, tested and implemented[3].

Based on the literature survey it is observed that i) The existing models lack flexibility in handling the uncertainties of the real life situations. ii) Probabilistic models sometimes fail to model the real-life uncertainties. iii) The existing prediction techniques show poor response time due to their complex algorithmic structure. iv) Most of the approaches validate the results with simulated data.

The present work provides a novel solution for short term motion prediction using adaptive Fuzzy prediction technique. History of moving object motion positions are captured in the form of Fuzzy rule base and the next instance object position is predicted using fuzzy inference process. Because of the multi valued nature of fuzzy logic this approach enjoys high robustness in dealing with noisy and uncertain data.

However, direct implementation of the rule base is not suitable for real-life navigation systems due to the formation of huge number of rules. To overcome this drawback rule-base is optimized by adaptive navigational environment.

2 Fuzzy Rule Based Object Motion Prediction

The navigational environment is modeled as Fuzzy World model [3] which can be observed in most of the applications. The Fuzzy representation of the environment is shown in Figure 1 with numerical notation for each region.

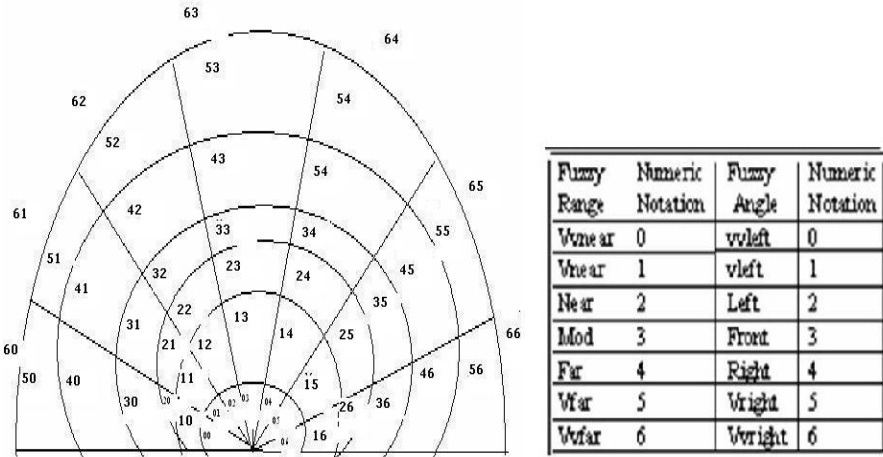


Fig. 1. Division of Navigation Space into Fuzzy subsets of Range and Direction

In the rule-base formation phase, rules are defined and added to the rulebase using real-life data, expert knowledge base and a simulator. At time t_1 , the position (Angle and Range) of the moving object from the Robot is read. Using Fuzzification the observed data is converted to Fuzzy value. At time t_2 ($t_2 > t_1$ and $t_2 - t_1 > \delta$, where δ is threshold time difference greater than or equal to 1 sec), the sensor reads the position of the same object. The read value is converted to Fuzzy value. The same process is followed at time t_3 ($t_3 > t_2$ and $t_3 - t_2 = t_2 - t_1$) to get the Fuzzy value of the location of the same object under observation. A Fuzzy rule with the positions of the moving object at time t_1 and t_2 as the antecedent and the position of the object at time t_3 as the consequent is formed and added to the rule-base. Each rule in the rule-base is represented as

$$\text{IF } (R_1, \theta_1) \text{ and } (R_2, \theta_2) \text{ THEN } (R_3, \theta_3)$$

where R_1 and θ_1 represent the Range and the Angle respectively of the object at time t_1 , R_2 and θ_2 represent the Range and the Angle respectively of the object at time t_2 , and R_3 and θ_3 represent the Range and the Angle respectively of the object at time t_3 .

Similar rules are added to the rule-base for different objects observed at various positions in the navigation environment. In the implementation phase of the predictor, the Robot observes the moving object at time t_1 and t_2 and sends the data to the Fuzzy predictor algorithm. With the application of Fuzzy inference process, prediction of the next instance position of the moving object is carried out. The complete process of short term motion prediction is represented in Figure 2.

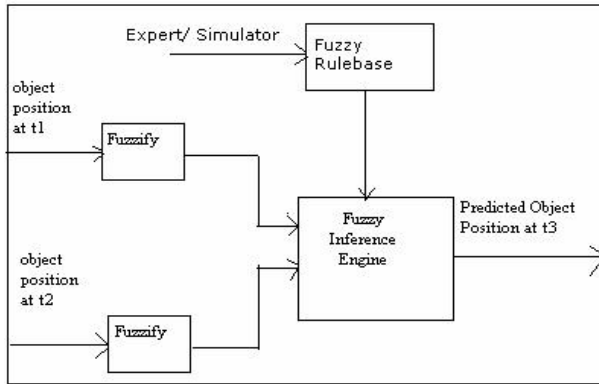
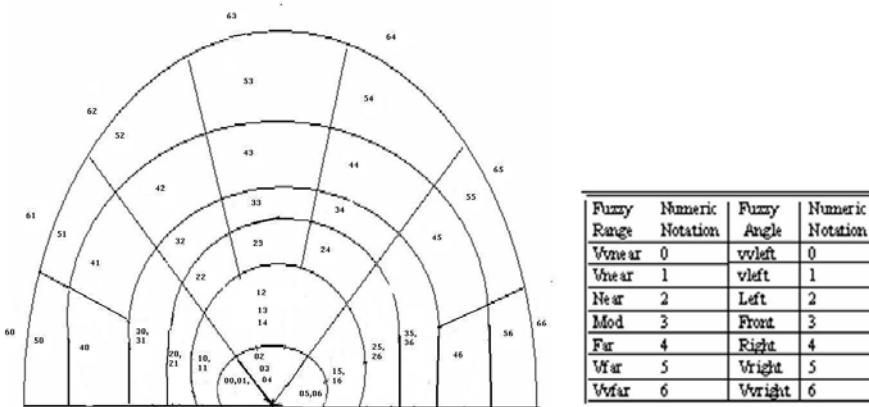


Fig. 2. Short term motion prediction

3 Rulebase Optimization Using Adaptive Navigational Environment

To enhance the performance of the predictor algorithm, the basic navigation environment is altered such that at nearer distance only three, at moderate distance five



a. Division of space in Fuzzy subsets of Range and Direction

b. Numeric notation for Fuzzy range and Angle Values

Fig. 3. Adaptive division of navigation space into Fuzzy regions

and at the far distance seven Fuzzy membership functions are defined for angular subset by merging adjacent members in the angular subset(Fig. 3). By defining Adaptive navigational environment, the number of Fuzzy rules can be decreased as well as the accuracy of the results can be further improved.

4 Experimental Results

The Fuzzy predictor algorithm is developed in C++ language. The algorithm is tested on 1.66 GHz machine in VC++ environment. The tests are carried out for real-life benchmarked datasets [5]. Figure 4 represents the movement of the objects from left to right direction and the corresponding short term motion prediction path. P_i and A_i represent the predicted and the actual path traversed by the moving object. $P_i(G)$ and $A_i(G)$ represent the predicted goal and the actual goal of the object. $A1$ is the actual path observed and $A1(G)$ is the actual goal reached by the object $A1$.

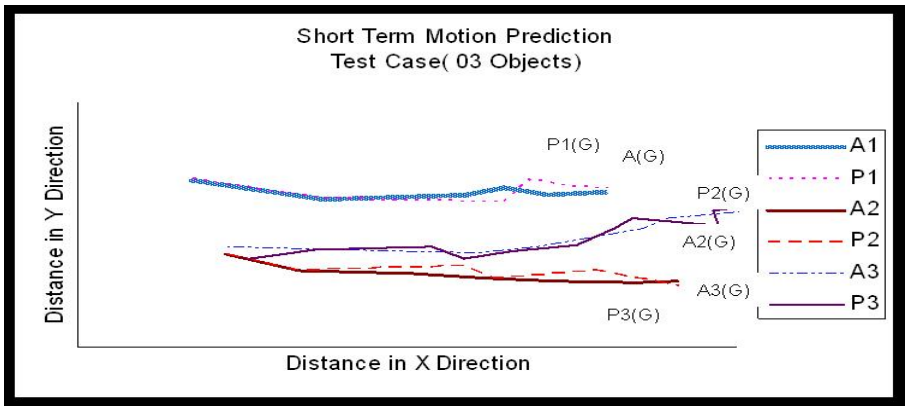


Fig. 4. Prediction graphs showing the few of the path prediction solutions for Short term motion prediction

Table 1. Comparison of Short term predictors

Short Term Predictor	Relative Error	Response time in seconds
Neural Network predictor	6-17%	560×10^{-3} sec
Bayesian Occupancy Filters	1-10%	100×10^{-3} sec
Extended Kalman Filter	1-20%	0.1 sec
Proposed Fuzzy Predictor Algorithm	1-10%	02×10^{-3} sec to 05×10^{-3} sec

Table 1 compares a few of the well known prediction techniques which are re-implemented and compared with the developed Fuzzy predictor in respect of response time and relative error. From the table it can be observed that the performance of the predictor is comparable with regard to relative error but better than the other prediction methods as far as response time is concerned.

5 Conclusion

In a dynamic navigation system the Robot has to avoid stationary and moving objects to reach the final destination. Short Term motion prediction for moving objects in such an environment is a challenging problem. This paper proposes a simplified approach for predicting the future position of a moving object using fuzzy inference rules derived from expert knowledge. Fuzzy based prediction is more flexible, can have more real life parameters, comparable to the existing approaches and suited for real life situations. The results of the study indicate that, the Fuzzy predictor algorithm gives comparable accuracy with quick response time when compared to existing techniques.

Acknowledgments. The authors are thankful to the benchmark dataset provided by EC Funded CAVIAR project, CMU Graphics lab and Motion capture web group.

References

1. Foka, A., Trahanias, P.E.: Predictive Autonomous Robot navigation. In: Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems, EFPL, Lausanne, Switzerland, pp. 490–494 (October 2002)
2. Fayad, C., Web, P.: Optimized Fuzzy logic based algorithm for a mobile robot collision avoidance in an unknown environment. In: 7th European Congress on Intelligent Techniques & Soft Computing, Aachen, Germany, September 13-16 (1999)
3. Angelopoulou, E., Hong, T.-H., Wu, A.Y.: World Model Representation for Mobile Robots. In: Proceedings of the Intelligent Vehicles 1992 Symposium, pp. 293–297 (1992)
4. Madhavan, R., Schlenoff, C.: Moving Object Prediction for Off-road Autonomous Navigation. In: Proceedings of the SPIE Aerosense Conference, April 21-25 (2003)
5. Fisher, R., Santos-Victor, J., Crowley, J.: CAVIAR Video Sequence Ground Truth (2001), <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>
6. Zhuang, H.-Z., Du, S.-X., Wu, T. -j.: On-line real-time path planning of mobile Robots in dynamic uncertain environment. Journal of Zhejiang University Science A, 516–524 (2006)
7. Zhu, Q.: Hidden Markov Model for Dynamic Object Avoidance of Mobile Robot Navigation. IEEE Transactions on Robotics and Automation, 390–396 (1991)

An Efficient Protocol Using Smart Interval for Coordinated Checkpointing

Jagdish Makhijani¹, Manoj Kumar Niranjana², Mahesh Motwani³, A.K. Sachan⁴,
and Anil Rajput⁵

¹ Rustamji Institute of Technology, BSF Academy, Tekanpur
j_makhijani@yahoo.com

² Rustamji Institute of Technology, BSF Academy, Tekanpur
manoj_niranjana2000@yahoo.co.in

³ Rajiv Gandhi Technical University, Bhopal
mahesh.7@sify.com

⁴ Radharaman Institute of Technology & Science, Bhopal
sachanak_12@yahoo.com

⁵ Bhabha Engineering Research Institute, Bhopal
dranilrajput@hotmail.com

Abstract. Checkpointing using message logging is a very popular technique for fault tolerance in distributed systems. The proposed protocol controls the lost messages, orphan messages and also simplifies garbage collection which is not available in most of existing protocols. In the protocol, all processes take checkpoints at the end of their respective smart interval to form a global consistent checkpoint. Since the checkpointing is allowed only within smart interval, the protocol will minimize various overheads like checkpointing overhead, message logging overhead etc.

Keywords: Distributed Systems, Checkpointing, Fault Tolerance and Message Logging, Smart Interval.

1 Introduction

A distributed system consists of multiple autonomous computers that communicate through a computer network in order to achieve a common goal. The distributed computing systems generally tolerate undesired changes in their internal structure or external environment in regular working which can be referred to as faults. A Fault may be a design or operational fault and may occur at once or many times. To make a system fault tolerable, fault tolerance techniques such as Checkpointing may be used. Checkpointing is the method of periodically recording the state of the system in stable storage. Any such periodically saved state is called the checkpoint of the process [1]. A global state [2] of a distributed system is a set of individual process states, on per process [1]. Checkpointing may be one of two types, i.e., independent and coordinated checkpointing. In independent checkpointing, each process takes checkpoint independently without requiring any synchronization when a checkpoint is taken [3].

In coordinated checkpointing, the processes coordinate their checkpointing action in such a way that the set of local checkpoints taken is consistent [4,5,6].

The present work suggests a new coordinated checkpointing algorithm which effectively manages the lost and orphan messages. In this algorithm, each process takes turn to act as checkpoint initiator. The checkpoint initiator sends messages to other processes to be prepared for checkpoint and then to take checkpoint. However, a process has to maintain a log of received, sent, and unacknowledged messages of the current checkpointing interval. The initiator issues commit message to all other processes after receiving checkpoint from all processes. The set of these checkpoints made permanent after issue of commit. If initiator does not receive all the local checkpoints, it issues abort message for not making the tentative checkpoint permanent.

2 Existing Works

In the existing work, the initiator communicates with other processes to create a checkpoint. In these old checkpointing protocols, if message communication can take place after checkpoint request of initiator, the global checkpoint may be inconsistent. This is shown in fig. 1 in which message m is sent by P_0 after receiving a checkpoint request from the initiator. If m reaches P_1 before the checkpoint request, the checkpoint will become inconsistent because checkpoint $c_{1,x}$ confirms that message m is received from P_0 , while checkpoint $c_{0,x}$ says that it is not sent from P_0 . [11]

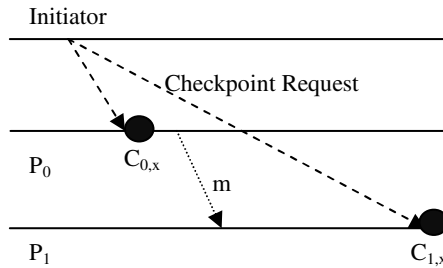


Fig. 1. Message communication between P_0 and P_1 causing inconsistent checkpoint

In another protocol, the message communication is allowed within a fixed time interval only. This concept reduces message communications [7] which is beneficial in decreasing the communication overhead. The main drawback of this protocol is the fixation of a particular process as initiator process. Since a fixed process will act as initiator in entire system execution, thus the probability of failure will be high.

In another checkpointing protocol, the process initiator is not fixed which reduces the probability of failure of initiator. The drawback of this protocol is that the message communication could be accomplished at any time i.e., there is no concept of fixed time interval for message communication. Hence it increases communication overhead and output commit latency [8].

The proposed protocol overcomes to these shortfalls. The proposed protocol uses a fixed time interval for message communications which controls the message

communication. This fixed time interval is called smart interval. This concept reduces the communication overhead. The protocol also gives chance to every process to act as initiator process which reduces the probability of failure of initiator.

3 System Model

Let us consider a system of 'n' processes, P_0, P_1, \dots, P_{n-1} . The no. of processes 'n' is fixed for the duration of execution. Let the checkpoints be denoted as CP_k^i , i.e., initial checkpoint CP_k^0 ($i=0$), first checkpoint CP_k^1 ($i=1$), second checkpoint CP_k^2 ($i=2$) and so on (here k is the process no.). The initial checkpoint is taken when the system is being initialized. Each process maintains its own independent data structures, states and computations. Processes have no shared memory and no global clock. All communications among processes are through message passing only. We are assuming followings:

- 1) The underlying network guarantees reliable FIFO (First In First Out) delivery of messages between any pair of processes. The assumption of FIFO delivery assures the message synchronization.
- 2) Each process takes turn to initiate checkpointing at regular interval. The initial checkpoint (CP_k^0) is taken at the time of system initialization and initiated by P_0 . The next checkpoint, i.e. first checkpoint (CP_k^1) will be initiated by P_1 and so on.
- 3) The initiator process cannot be fail. If the initiator process fails, the global checkpoint (which is always stored at initiator process) will be lost and the entire system process will be collapsed.

The message communication will took place only in *smart interval*. The *smart interval* is a specified time interval which is elapsed between the control messages for prepare checkpoint and take checkpoint. If any process sends a message within smart interval, it has to be logged and the process execution is continued. This enables handling of lost messages. [10] The initiator process sends the control messages for prepare checkpoint and take checkpoint to other processes.

4 Protocol Description

The checkpoint initiator process sends *checkpoint-prepare-request-message* to other processes to start checkpointing. The other processes send their responses to the initiator process. If initiator process received replies from all processes within smart-interval then it sends *take-checkpoint-request-message* and if initiator process does not receive replies from any process within smart-interval then it will send *abort-checkpoint-request-message*. The set of checkpoint of all processes received by initiator process is called global checkpoint. A local checkpoint is denoted by CP_k^i where k is the process id and i is the checkpoint number. The i^{th} global checkpoint is the set $CP^i = \{CP_0^i, CP_1^i, \dots, CP_{n-1}^i\}$ in a system of n processes. CP^i is said to be consistent if and only if $\forall j, k \in [0, n-1]; j \neq k \Rightarrow (CP_j^i \rightarrow CP_k^i)$ where \rightarrow denotes the happened-before relation described by Lamport in [9].

The maximum transmission delay to reach a message to destination is t. The T is the checkpointing interval. Here $T > 3t$, since checkpoint interval (T) is obviously

greater than smart-interval and the length of smart-interval is bound to be at least $3t$ to survive the transmission delay of control messages (*checkpoint-prepare-request-message*, *response of checkpoint-prepare-request-message* and *take-checkpoint-request-message* and each transmission will take at least t) and to enable logging of computational messages.

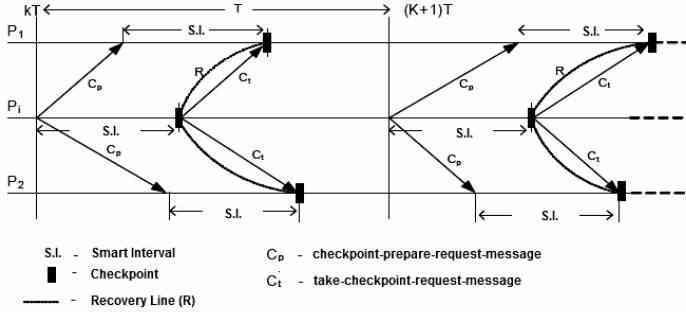


Fig. 2. Diagram showing message communication during smart interval

Now, let us define the following terms:

t^{prep} =instant at which initiator process starts sending prepare request

t^{rec} =instant at which a process receives prepare request

t^s =instant at which a computation message is sent

t^r =instant at which a computation message is received

T^{trms} =maximum transmission time for message including allowable delay (which is t)

Tag1=normal acknowledgement

Tag2=acknowledgement indicating that sender must keep this message logged

save_state (P_i)=procedure that saves the current state of process P_i

ack[]=keeps record of whether acknowledgement has come for the corresponding message (ack[a] is set to 0 when a^{th} message is sent and set to 1 when acknowledgement with tag1 for a^{th} message comes back, otherwise it is set to 0 if acknowledgement with tag2 comes back for a^{th} message.)

send(), receive()=functions for sending and receiving messages respectively.

s_id=sender id

5 Checkpointing Process

The checkpoint process starts at the time of system initialization. After T time interval (which is decided by the programmer) of previous checkpoint, the next initiator process starts the process of next checkpoint. The initiator process P_i sends *checkpoint-prepare-request-message* to all other processes at t^{prep} . On receiving *checkpoint-prepare-request-message*, each process write *tentative checkpoint* after sending response to the initiator.

- 1) Now, if initiator receives response from all processes, within $(t^{\text{prep}}+2*T^{\text{trns}})$, the initiator process sends *take-checkpoint-request-message* to all processes. When receiver receives *take-checkpoint-request-message* from initiator process, the *tentative checkpoint* is made *permanent*. This will save the states of all processes which are responsible for preparing a global checkpoint. Now, suppose if one or more process fails after responding to *checkpoint-prepare-request-message*, then the *tentative checkpoint* (which is prepared in response to *checkpoint-prepare-request-message*) is used to recover the failed process.
- 2) Now suppose if one or more process fails to respond to *checkpoint-prepare-request-message*, the initiator process sends *abort-checkpoint-request-message* to all processes. On receiving this, the *tentative checkpoint* is deleted. The copy of unacknowledged message keeps in a log in this case.

6 Algorithm

Step-I:

This step is executed at initiator process P_i

- i. Send *checkpoint-prepare-request-message* to remaining processes at t^{prep} for $(k+1)^{\text{th}}$ checkpoint
- ii. Remove $(k-1)^{\text{th}}$ checkpoint, if exist.
- iii. Receive response from other processes within $(t^{\text{prep}}+2*T^{\text{trns}})$
- iv. If all processes respond positively then
 - Send *take-checkpoint-request-message* to all processes
 - Else (if even a single process does not respond positively or response does not arrive to initiator process)
 - a. Send *abort-checkpoint-request-message* to all processes
 - b. Retain copies of unacknowledged messages in a log

Step-II:

This step is executed at other process P_{oth}

- i. Receive *checkpoint-prepare-request-message* from initiator at t^{rec}
- ii. Send own response to initiator
- iii. If response is positive then Call `save_state(P_{oth})` to write *tentative-checkpoint* asynchronously
- iv. Wait for decision of P_i till $(t^{\text{rec}}+T^{\text{trns}}+T^{\text{trns}})$
- v. If received decision is *take-checkpoint-request-message* then Change status of *tentative-checkpoint* to permanent
 - Else
 - Delete *tentative-checkpoint*
- vi. Delete messages whose acknowledgements have received. Log unacknowledged messages.

Step-III:

This step is executed at any process P_{any} for receiving message

- i. If $((\text{checkpoint number in message})=(\text{checkpoint number in } P_{\text{any}}))$

- a. Send (tag1,s_id)
- b. Receive(message)
- ii. else if ((checkpoint number in message)>(checkpoint number in P_{any}))
 - a. save_state(P_{any})
 - b. send(tag1,s_id)
 - c. receive(message)
- iii. else if ((checkpoint number in message)<(checkpoint number in P_{any}))
 - a. send (tag2,s_id)
 - b. receive(message)

Step-IV:

This steps is executed at any process P_{any} for writing unacknowledged messages

- i. for all k
 - if (ack[k]=0) then write k^{th} message in buffer

7 Performance Results

The proposed algorithm is simulated in Windows Environment using Windows XP Operating System and Visual Studio. It is assumed that coordinator process and network will never fail. The result shows that a distributed system with proposed algorithm will not fail, whether the total execution time may increase. This increase in execution time will depend on no. of fault. Table-1 shows the summarized result:

Table 1. Performance Result of proposed algorithm

Total Time to Complete	No. of Process	Time to Check point	Checkpoint Preparation Time	Time to Complete without Error (with algorithm)	No. of Errors occurred	Execution Time Increase % (with errors)	Execution Time Increase % check pointing (with errors)
100	10	1	0.1	110	10	21.00	19.09
150	10	1	0.1	165	10	17.33	10.51
200	10	1	0.1	220	10	15.50	7.05
100	20	1	0.1	110	10	21.00	19.09
150	20	1	0.1	165	10	17.33	10.51
200	20	1	0.1	220	10	15.50	7.05

8 Conclusion

Consistent checkpointing is formed in a distributed manner by using proposed checkpointing protocol. The checkpointing protocol of this paper also manages the unacknowledged messages to decrease the communication overhead. A global checkpoint includes each and every checkpoint taken by a process. Hence the last global checkpoint has to be retained. This decrease a lot of overhead since it required only $3*(n-1)$ messages during checkpointing in ideal case for n number of processes. The proposed protocol uses the concept of smart interval which allows message communication in a fixed time interval. This provides faster message communication since the bandwidth will be used for less no. of messages. The protocol minimizes output commit latency

and simplifies garbage collection. The message communication may result high output commit latency due to communication required and saving the message log to stable storage. The proposed protocol allows message communication only in smart interval. Since no. of messages is reduced, output commit latency is minimized. Garbage collection is a technique of managing storage memory efficiently by removing unwanted checkpoints from stable storage. In the proposed protocol, whenever initiator process P_i sends *checkpoint-prepare-request-message* for $(k+1)^{\text{th}}$ checkpoint, the protocol will automatically delete the $(k-1)^{\text{th}}$ global checkpoint which results simplified garbage collection. The protocol is useful in tolerating all types of software faults occurred on non-initiator processes.

There is no provision for failure of initiator process which may be considered as a drawback of our protocol.

References

1. Manivannan, D., Netzer, R.H.B., Singhal, M.: Finding Consistent Global Checkpoints in a Distributed Computation. *IEEE Trans. On Parallel & Distributed Systems* 8(6), 623–627 (1997)
2. Tsai, J., Kuo, S.: Theoretical Analysis for Communication-Induced Checkpointing Protocols with Rollback-Dependency Trackability. *IEEE Trans. on Parallel & Distributed Systems* 9(10), 963–971 (1998)
3. Bhargava, B., Lian, S.R.: Independent Checkpointing and Concurrent Rollback for Recovery in Distributed Systems-An Optimistic Approach. In: *Proceeding of IEEE Symposium on Reliable Distributed Systems*, pp. 3–12 (1988)
4. Cao, G., Singhal, M.: On Coordinated Checkpointing in Distributed Systems. *IEEE Transactions on Parallel And Distributed Systems* 9(12), 1213–1222 (1998)
5. Sharma, D.D., Pradhan, D.K.: An Efficient Coordinated Checkpointing Scheme for Multi-computers. In: *Proc. IEEE Workshop on Fault-Tolerant Parallel and Distributed Systems*, pp. 36–42 (June 1994)
6. Elnozahy, E.N., Johnson, D.B., Zwaenepoel, W.: The Performance of Consistent Checkpointing. In: *Proc. 11th Symp. Reliable Distributed Systems*, pp. 39–47 (October 1992)
7. Subba Rao, C.D.V., Naidu, M.M.: A New, Efficient Coordinated Checkpointing Protocol Combined with Selective Sender-Based Message Logging. In: *IEEE/ACS International Conference on Computer Systems and Applications, AICCSA 2008*, pp. 444–447 (2008)
8. Neogy, S., Sinha, A., Das, P.K.: CCUML: A Checkpointing Protocol for Distributed System Processes. In: *IEEE Transactions on TENCON 2004, IEEE Region 10 Conference, November 21-24, vol. B*, pp. 553–556 (2004)
9. Chandy, K.M., Lamport, L.: Distributed Snapshots: Determining Global States of Distributed Systems. *ACM Trans. on Computer Systems* 3, 63–75 (1985)
10. Subba Rao, C.D.V., Naidu, M.M.: A Survey of Error Recovery Techniques in Distributed Systems. In: *Proc. 28th Annual Convention and Exhibition of IEEE India Council*, pp. 284–289 (December 2002)
11. Mootaz Elnozahy, E.N., Alvisi, L., Wang, Y.-M., Johnson, D.B.: A Survey of Rollback-Recovery Protocols in Message-Passing Systems. *ACM Computing Surveys (CSUR)* 34(3), 375–408 (2002)

Face Recognition System Using Discrete Wavelet Transform and Fast PCA

K. Ramesha¹ and K.B. Raja²

¹ Department of Telecommunication Engineering, Vemana Institute of Technology,
Koramangala, Bangalore-560034
rameshk13@yahoo.co.uk

² Department of Electronics and Communication Engineering, University Visvesvaraya
College of Engineering, Bangalore
University, K.R. Circle, Bangalore-560001

Abstract. The face recognition system is used to create a national database for the purpose of identity cards, voting in an electoral systems, bank transaction, food distribution system, control over secured areas etc. In this paper we propose the Face Recognition System using Discrete Wavelet Transform and Fast PCA (FRDF). The Discrete Wavelet Transform is applied on face images of Li-labor Spacek database and only LL subband is considered. Fast Principal Component Analysis using Gram-Schmidt orthogonalization process is applied to generate coefficient vectors. The Euclidean Distance between test and database face image coefficient vectors are computed for face recognition based on the threshold value. It is observed that the face recognition rate is 100% and the proposed algorithm for the computation of eigenvalues and eigenvectors improves the computational efficiency as compared to Principal Component Analysis (PCA) with same Mean Square Error (MSE).

Keywords: Face Recognition, Fast Principal Component Analysis, Discrete Wavelet Transform, Eigenvalue, Eigenvector, Error Vector.

1 Introduction

Biometric authentication of a person has many advantages compared to the existing traditional methods such as personal identity number, a key, a card etc. Biometric mainly classified based on the characteristics of a person into two groups viz., physiological characteristics i.e., Face, Fingerprint, Hand geometry, Hand vein, Iris, Retina, DNA, Facial thermo gram and behavioral characteristics such as Signature, Keystroke dynamics, Speech. The physiological biometrics has an advantage over behavioral biometrics as the characteristics of a particular person do not change over a period of years. The face recognition system attains lot of interest in these days due to its applications such as proof of identity card for access control to physical facilities, fight against terrorism and crime, access control to services, law enforcement system and content based video processing system. Fully automatic robust face recognition is required due to the widespread use of photo identity for personal security and

identification. Automatic face recognition system is more reliable, effective and easy to implement compared to other biometric systems such as fingerprint, hand geometry, hand vein, iris, retina, facial thermo gram, Signature, Keystroke dynamics, DNA and Speech and it does not require special knowledge of a person or cooperation.

The process of face recognition normally consists of two steps. The first step is face detection and localization, in which faces have to be found in the input image and separated from the background. The second step is face feature extraction and recognition. Challenges in face recognition research are face illumination variations, face rotation, different hair style of a person wearing different kinds of spectacles, ageing effects and facial expressions. Eigenface based analysis is a method about the features of the whole face appearance and even to the whole sample set. Face recognition under Pattern recognition, the most popular and necessary problem is dimensionality reduction. Over the past few years, several face recognition systems based on PCA have been proposed [1]. Instead of using N intensity values for an N pixel image, it is feasible to specify an image by a set of M features, where $M \ll N$. The selected features must be able to uniquely represent the right class for their corresponding facial images.

Kishore S Kinare and Bhirud [2] proposed Two Dimensional Principal Component Analysis (2DPCA) on wavelet subband. Haar, Daubechies, Coiflet, Symlet, Biorthogonal and Reverse Biorthogonal wavelet transforms are used to extract image features of facial images by decomposing face image in subbands of 1 to 8. By using 2DPCA and Euclidean Distance (ED) measures the features are analyzed. Daw-Tung Lin [3] developed a method utilizing PCA to perform facial expression recognition using Hierarchical Radial Basis Function Network for the facial expression classifications based on local feature extraction by PCA from lips and eyes images. Sheifali Gupta et al., [4] provided a method for face recognition using PCA, an eigenface approach, in which a small set of characteristic pictures are used to describe the variation between face images. ED is used to match facial images. Vinod Pathangay and Sukhendu Das [5] proposed the use of selective subbands for PCA based frontal face recognition with variations in illumination and expression. Subband face representation is evaluated using PCA. Kai Chen and Le Jun Zhao [6] presented real time face recognition and tracking system to detect face for recognition and tracking. Hybrid algorithm wavelet, PCA and Support Vector Machine is used for face recognition. Face tracking is done by using meانشift and Kalman filter.

2 Proposed Face Recognition System Model

The face recognition approach based on DWT and Fast PCA is as shown in Figure 1.

2.1 Face Image Database

The face images are collected from Libor Spacek database for eleven persons from F1 to F11. The data set used for training and testing purposes contains male, female and old person images of $180 * 200$ pixels size. In this data set, twenty images of each person without background with very minor variation in head turn, tilt and slant are considered. The data set has images of small changes in face position, because images

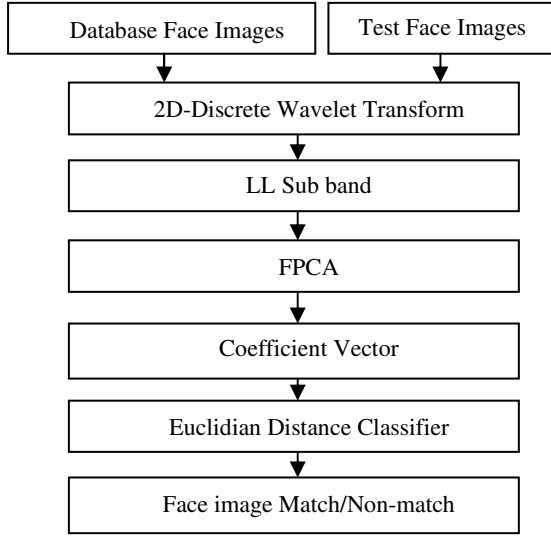


Fig. 1. FRDF Face Recognition System

have been acquired in speech mode with no variation in hair style and lighting. The face image considered for testing may be from database or from out of database.

2.2 Two Dimensional DWT

The decomposition is applied at different levels repeatedly on low frequency channel (LL) to obtain next level decomposition. The image is decomposed into four subbands LL, LH, HL, and HH subbands by applying 2D DWT on face image. The LL subband corresponds to low frequency components of an image and HL, LH and HH are high frequency components of an image corresponds to vertical, horizontal and diagonal subbands respectively. The LL subband we obtain is half the original image. Figure 2 shows the image decomposition based on wavelet scales. 2D DWT gives dimensional reduction for less computational complexity, insensitive feature extraction, and multiresolution data approximation. The transform decomposes an image and hence different facial expressions are attenuated by removing high frequency components. Wavelet coefficients are obtained by convolving a target function with wavelet kernels and mathematically DWT can be given as in Equation (1)

$$DWT_{x(n)} = \begin{cases} d_{p,q} = \sum x(n) h_p^*(n-2^p q) \\ a_{p,q} = \sum x(n) g_p^*(n-2^p q) \end{cases} \quad (1)$$

The coefficients $d_{p,q}$ gives the component details to the wavelet function, where as $a_{p,q}$ gives approximation components of the image. The $h(n)$ and $g(n)$ in the Equation (1)

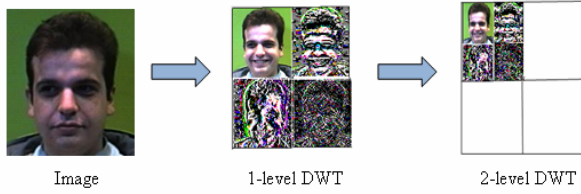


Fig. 2. Image decomposition based on wavelet scales

are functions, gives the coefficients of high pass and low pass filters respectively. The parameters p and q refers to wavelet scale and translation factors.

2.3 FPCA

PCA based system is affected by non-convergence state of the algorithm and high MSE. The major difference is for facial features such as eigenvalues and eigenvectors extraction using Gram–Schmidt Orthogonalization method instead of eigenvalue decomposition method in PCA. Fast PCA using Gram-Schmidt orthogonalization process to find leading eigenvectors converges in little iteration without any initial setting. Face recognition using FPCA decreases the decision time of a system, especially when high resolution images are used, hence FPCA is computationally more efficient, easy to implement and generates same MSE as that of PCA.

2.4 Identify the Known/Unknown Face Image

The Euclidean distance between database set and test faces becomes error vector and the average error vector becomes the threshold value for face recognition. The minimum ED between database and test image are recorded which leads to Difference Error Vector (DEV). If the value of DEV is less than the threshold value, then the face image is concluded as match otherwise non-match.

3 Performance Analysis and Results

For testing and analysis purpose, 400 images of 20 persons are considered. First 10 images of each person are used to create a face database. The second set of 10 images of each person is used as test images to determine recognition rate. For Non-matching, the test images are from different persons other than the persons used to create database. It is observed from the Table 1 that the average recognition rate is 100% in the case of proposed algorithm compared to the average recognition rate of 80% in the case of existing algorithm.

The CPU time increases exponentially as the the database size increases in the case of PCA method [7] of face recognition, whereas in the case of proposed FRDF the CPU time variation is very low for larger database. The FRDF algorithm using Harr

Table 1. FPCA and FRDF Face Recognition Rates

Database	Recognition Rate	
	FPCA [8]	FRDF
F1	90%	100%
F2	100%	100%
F3	100%	100%
F4	50%	100%
F5	90%	90%
F6	90%	100%
F7	100%	100%
F8	100%	100%
F9	90%	100%
F10	0%	100%
F11	90%	100%

wavelet, scale-1 requires only few leading principal components to achieve 100% recognition accuracy with less computational time.

4 Conclusion

The face recognition is normally used to identify a person for natural security. The FRDF algorithm is proposed in the paper. The face images from Libor Spacek database is considered for training and testing. The DWT is applied on face and considered only LL subband by leaving other subbands. The FPCA using Gram–Schmidt orthogonalization process is applied on LL subband to generate leading eigenvalues to compute face features. The Euclidean distance is used to compare the face features of database and test face images to obtain face match/non-match. It is observed that in the proposed algorithm CPU time and recognition rate is improved compared to the existing algorithm. In future different transformations may be used on large face database for robust identification with minimum time consumption.

References

1. Nicholl, P., Amira, A.: DWT/PCA Face Recognition using Automatic Coefficient Selection. In: 4th IEEE International Symposium on Electronic Design, Test and Applications, Hong Kong, pp. 390–393 (2008)
2. Kinare, K.S., Bhirud, S.G.: Face Recognition based on Two-Dimensional PCA on Wavelet Subband. *Int. J. Recent Trends in Engineering* 2(2), 51–54 (2009)
3. Lin, D.-T.: Facial Expression Classification using PCA and Hierarchical Radial Basis Function Network. *J. Information Science and Engineering*, 1033–1046 (2006)
4. Gupta, S., Sahu, O.P., Gupta, R., Goel, A.: A Bespoke Approach for Face Recognition using PCA. *Int. J. Computer Science and Engineering* 2(2), 155–158 (2010)

5. Pathangay, V., Das, S.: Exploring the use of Selective Wavelet Subbands for PCA based Face Recognition. In: National Conference on Image Processing, pp. 182–185. IISc, Bangalore (2005)
6. Chen, K., Zhao, L.J.: Robust Real Time Face Recognition and Tracking System. *J. Computer Science and Technology* 9(2), 82–88 (2009)
7. Sharma, A., Paliwal, K.K.: Fast Principal Component Analysis using Fixed-Point Algorithm. *ELSEVIER I. J. Pattern Recognition Letters* 28, 1151–1155 (2007)
8. Sajid, I., Ahmed, M.M., Taj, I.: Design and Implementation of a Face Recognition System using Fast PCA. In: International Symposium on Computer Science and its Applications, Hobort, ACT, pp. 126–130 (2008)

Mining Indirect Association between Itemsets

B. Ramasubbareddy¹, A. Govardhan², and A. Ramamohanreddy³

¹ Associate Professor, Jyothishmathi Institute of Technology and Science, Karimnagar, India
rsreddyphd@gmail.com

² Professor & Principal, JNTUH college of Engineering, Karimnagar, India
govardhan_cse@yahoo.co.in

³ Professor, S.V.U. College of Engineering, S.V. University, Tirupati, India
ramamohansvu@yahoo.com

Abstract. Discovering association rules is one of the important tasks in data mining. While most of the existing algorithms are developed for efficient mining of frequent patterns, it has been noted recently that some of the infrequent patterns, such as negative associations and indirect associations, provide useful insight into the data. Existing indirect association mining algorithms mine indirect associations between items and require two join operations. But in this paper, we propose an algorithm for mining the complete set of indirect associations between pair of items and itemsets which require only one join operation.

Keywords: Data mining, positive and negative association rules, indirect association.

1 Introduction

Association rule mining is a data mining task that discovers associations among items in a transactional database. Association rules have been extensively studied in the literature for their usefulness in many application domains such as recommender systems, diagnosis decisions support, telecommunication, intrusion detection, etc. Efficient discovery of such rules has been a major focus in the data mining research. From the celebrated *Apriori* algorithm [1] there have been a remarkable number of variants and improvements of association rule mining algorithms [2]. A typical example of association rule mining application is the market basket analysis. In this example, the behavior of the customers is studied with reference to buying different products in a shopping store. The discovery of interesting patterns in this collection of data can lead to important marketing and management strategic decisions. For instance, if a customer buys bread, what are chances that customer buys milk as well? Depending on some measure to represent the said chances of such an association, marketing personnel can develop better planning of the shelf space in the store or can base their discount strategies on such associations/correlations found in the data. All the traditional association rule mining algorithms were developed to find positive associations between items.

In [9], a new class of patterns called indirect associations has been proposed and its utilities have been examined in various application domains. Consider a pair of items

X and Y that are rarely present together in the same transaction. If both items are highly dependent on the presence of another itemset M, then the pair (X, Y) is said to be indirectly associated via M. There are many advantages in mining indirect associations in large data sets. For example, an indirect association between a pair of words in text documents can be used to classify query results into categories [9]. For instance, the words *coal* and *data* can be indirectly associated via *mining*. If only the word *mining* is used in a query, documents in both *mining* domains are returned. Discovery of the indirect association between *coal* and *data* enables us to classify the retrieved documents into *coal mining* and *data mining*. There are also potential applications of indirect associations in many other real-world domains, such as competitive product analysis and stock market analysis [9].

This paper is structured as follows: the next section contains preliminaries about Indirect Association Rules, In Section3, existing strategies for mining indirect association rules are reviewed. The proposed algorithm is presented in Section 4 for finding all valid indirect association rules for pairs of multiple itemsets. Section 5 contains conclusions and future work.

2 Basic Concepts and Terminology

Let $I = \{i_1, i_2, \dots, i_m\}$ be a set of m items. A subset $X \subseteq I$ is called an itemset. A k -itemset is an itemset that contains k items. Let $D = \{T_1, T_2, \dots, T_n\}$ be a set of n transactions, called a transaction database, where each transaction $T_j, j = 1, 2, \dots, n$, is a set of items such that $T_j \subseteq I$. Each transaction is associated with a unique identifier, called its TID. A transaction T contains an itemset X if and only if $X \subseteq T$. The support of an itemset X is the percentage of transactions in D containing X . An itemset X in a transaction database D is called “frequent itemset” if its support is at least a user-specified minimum support threshold viz., *minsup*. Accordingly, an infrequent itemset is an itemset that is not a frequent itemset.

2.1 Negative Association Rules

An association rule is an implication of the form $X \Rightarrow Y$, where $X \subset I, Y \subset I$, and $X \cap Y = \emptyset$. Here, X is called the antecedent and Y is called the consequent of the rule. The confidence of an association rule $X \Rightarrow Y$ is the conditional probability that a transaction contains Y , given that it contains X . The support of rule $X \Rightarrow Y$ is defined as: $\text{sup}(X \Rightarrow Y) = \text{sup}(X \cup Y)$. Negative association was first pointed out by Brin et al. in [5]. Since then, many techniques for mining negative associations have been developed [8, 9, 10]. In the case of negative associations we are interested in finding itemsets that have a very low probability of occurring together. That is, a negative association between two itemsets X and Y , denoted as $X \Rightarrow \neg Y$ or $Y \Rightarrow \neg X$, means that X and Y appear very rarely in the same transaction. Mining negative association rules is computational intractable with a naive approach because billions of negative associations may be found in a large database while almost all of them are extremely uninteresting. This problem was addressed in [8] by combining previously discovered positive associations with domain knowledge to constrain the search space such that fewer but more interesting negative rules are mined. A general framework for mining both positive and

negative association rules of interest was presented in [10], in which no domain knowledge was required, and the negative association rules were given in more concrete expressions to indicate actual relationships between different itemsets. However, although the sets of the positive and negative itemsets of interest in the database were minimized in this framework, the search space for negative itemsets of interest was still huge. Another problem was that it tended to produce too many negative association rules, thus the practical application of this framework remained uncertain.

2.2 Indirect Association

Indirect association is closely related to negative association, they are both dealing with itemsets that do not have sufficiently high support. Indirect associations provide an effective way to detect interesting negative associations by discovering only “infrequent itempairs that are highly expected to be frequent” without using negative items or domain knowledge.

Definition (Indirect Association). A pair of itemsets X and Y is indirectly associated via a mediator M , if the following conditions hold:

1. $\text{sup}(X, Y) < t_s$ (Itempair Support Condition)
2. There exists a non-empty set M such that
 - (a) $\text{sup}(X \cup M) \geq t_f$, $\text{sup}(Y \cup M) \geq t_f$; (Mediator Support Condition)
 - (b) $\text{dep}(X, M) \geq t_d$, $\text{dep}(Y, M) \geq t_d$, (Mediator Dependence Condition)

where $\text{dep}(P, Q)$ is a measure of the dependence between itemsets P and Q .

The thresholds above are called itemset pair support threshold (t_s), mediator support threshold (t_f), and mediator dependence threshold (t_d), respectively. In practice, it is reasonably to set $t_f \geq t_s$.

Condition 1 is needed because an indirect relationship between two items is significant only if both items rarely occur together in the same transaction. Otherwise, it makes more sense to characterize the pair in terms of their direct association.

Condition 2(a) can be used to guarantee that the statistical significance of the mediator set. In particular, for market basket data, the support of an itemset affects the amount of revenue generated and justifies the feasibility of a marketing decision. Moreover, support has a nice downward closure property which allows us to prune the combinatorial search space of the problem. Condition 2 (b) ensures that only items that are highly dependent on the presence of X and Y will be used to form the mediator set.

Over the years, many measures have been proposed to quantify the degree of dependence between attributes of a dataset. From statistics, the Chi-Square test is often used for this purpose. However, the drawback of this approach is that it does not measure the strength of dependencies between items [19]. Furthermore, the Chi-Square statistic depends on the number of transactions in the database. As a result, other statistical measures of association are often used, including Pearson's Φ coefficient, Goodman and Kruskal's λ , Yule's Q and Y coefficients, etc [16]. Interest factor is another measure that has been used quite extensively to quantify the strength of dependency among items [11, 12, 13].

Definition. Given a pair of itemsets, say X and Y , its' IS measure can be computed using the following equation:

$$IS(X, Y) = \frac{P(X, Y)}{P(X)P(Y)} \quad (1)$$

Where P denotes the probability that the given itemset appears in a transaction.

3 Related Work in Indirect Association Rule Mining

It is observed that automated document translation systems tend to produce lexicon translation tables that are full of indirectly-associated words [15]. A lexicon translation table encodes the probability that two words from different languages being semantically equivalent to another. The presence of indirect association can pollute the resulting tables, thereby reducing the overall precision of the system. An iterative strategy was proposed in [15] to clean up existing translation tables by finding only the most probable translations for a given word.

The notion of internal and external measures of similarity between attributes of a database relation was introduced in [14]. Internal similarity between two attributes X and Y is a measure whose value depends only on the values of X and Y columns. Conversely, external measure takes into account data from other columns (called the probe attributes). Their notion of probe attributes is similar to mediators for indirect association in [14]. However, their sole purpose of using probe attributes is to perform attribute clustering.

Indirect association is closely related to the notion of negative association rules [17]. In both cases, we are dealing with itemsets that do not have sufficiently high support. A negative association rule discovers what are the items a customer will not likely buy given that he/she buys a certain set of other items. Typically, the number of negative association rules can be prohibitively large and the majority of them are not interesting to a data analyst. The use of domain knowledge, in the form of item taxonomy, was proposed in [17] to decide what constitutes an interesting negative association rule. The intuition here is that items belonging to the same parent node in taxonomy are expected to have similar types of associations with other items. If the observed support is significantly smaller than its expected value, then there is a negative association exists between the items. Again, unlike indirect association, these types of regularities do not specifically look for mediating elements.

Another related area is the study of functional dependencies in relational databases. Functional dependencies are relationships that exist between attributes of a relation. However, the emphasis of functional dependencies is to find dependent and independent attributes for applications such as semantic query optimization [18] and reverse engineering [18].

In [20], authors proposed an efficient algorithm, called HI-mine, based on a new data structure, called HI-Struct, for mining the complete set of indirect associations between items. Experimental results show that HI-mine's performance is significantly better than that of the previously developed algorithm for mining indirect associations on both synthetic and real world data sets over practical ranges of support specifications.

In [21], IAM algorithm proceeds in four phases: an initialization phase, a pruning phase, a bridge itemset calculation phase, and a ranking phase. The purpose of the initialization phase is to allocate the memory needed. The second phase is a process of pruning for the purpose of minimizing the search space of problem. The threshold value of pruning is $\min\text{-sup}(s)$. The third phase, the Bridge Itemset Calculation Phase, is the

most important for this algorithm. The last phase, a ranking phase, is mainly to finish the ranking operation according to 'the closeness value in the linked vector C for the purpose of providing decision makers the most useful indirect association rules.

4 Algorithm

The existing work done on generating indirect associations between pair of items only. In this paper, we propose a new method which generates indirect associations between pair of itemsets. This method contains two algorithms. Algorithm1 finds set of all frequent itemsets and set of all Valid Candidates (VC). An itemset V is said to be Valid candidate if $\text{sup}(V) \leq t_s$ and all subsets of V are frequent. Algorithm 2 finds set of all indirect association rules between pairs of itemsets.

Algorithm1. Finding Frequent(F) and ValidCandidates (VC)

Input: TDB- Transactional Database, m_s , t_s

Output: F- Frequent itemsets, VC- ValidCandidates

Method:

1. Find F_1 , the set all frequent 1-itemsets
2. for($K=2; F_{K-1} \neq \Phi ; K++$)
3. $\{ C_K = F_{K-1} \bowtie F_{K-1}$
- // Pruning infrequent itemsets
4. for each $c \in C_K$ {
5. if any sub-set of c is not a member of F_{K-1} then
6. $C_K = C_K - \{ c \}$ }
- // find support count for each itemset in C_K
7. for each c in C_K
8. if $\text{support}(c) \geq m_s$ then $F_k = F_k \cup \{ c \}$
9. for each c in C_K
10. if $\text{support}(c) \leq t_s$ then $VC = VC \cup \{ c \}$
11. $F = F \cup F_K$ }

Algorithm2. Mining Indirect Association Rules

Input: F, VC, t_f , t_d

Output: Indirect Association Rules

Method:

1. for each $I (= X \cup Y) \in VC$ {
2. for each $i \in F$ {
3. If ($\text{support}(X \cup i) \geq t_f$ && $\text{support}(Y \cup i) \geq t_f$)
4. If ($\text{dependency}(X \cup i) \geq t_d$ && $\text{dependency}(Y \cup i) \geq t_d$)
5. $IAR = IAR \cup (X, Y/i)$ }

- Line1, each I is infrequent but all subsets of I are frequent because all items of VC passed Apriori property.
- Line2, i is a frequent itemset
- Line3, since X, Y and i are frequent itemsets then $X \cup i$ and $Y \cup i$ may be frequent
- Line4, IS measure is used to find the dependency between two itemsets.

5 Experimental Results and Performance Evaluation

To evaluate the performance of proposed algorithm experiments are performed on two synthetic transactional databases containing 5400 and 12000 transactions each and implemented on java platform. We concentrate on mediator support which is a support of item-set and mediator and mediator dependency which is estimated by “Eq. (1)”.

Data set consisting of 5400 transactions with mediator support as 0.2,0.25,0.3,0.35; mediator dependence as 0.4,0.45,0.5,0.55 and the total number of rules generated as 96,15,36 and 20 respectively. Figure 1 shows the graph showing the mediator support and mediator dependency vs. total number of rules.

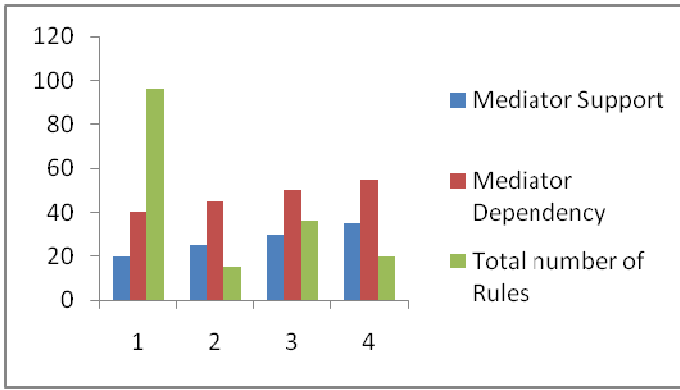


Fig. 1. Graph showing the mediator support and mediator dependency vs. total number of rules for 5400 transactions

Figure 2 is generated by considering 12000 transactions with mediator support as 0.2,0.25,0.3,0.35,0.4 mediator dependence as 0.4,0.45,0.5,0.55,0.6 and the total number of rules generated 31,27,7,7,and 6 respectively.

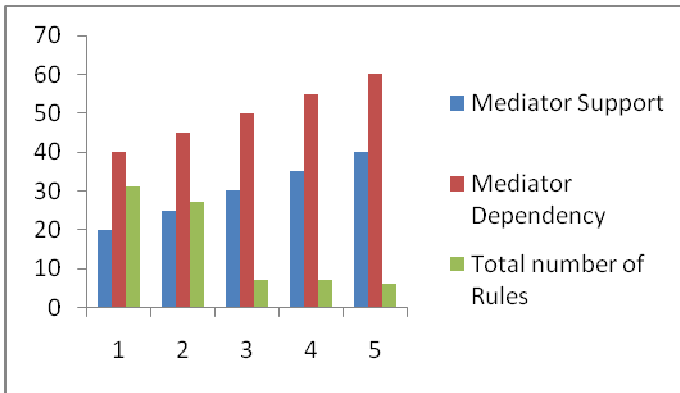


Fig. 2. Graph showing the mediator support and mediator dependency vs. total number of rules for 12000 transactions

6 Conclusion and Future Work

Existing indirect association mining algorithms mine indirect associations between items whilst indirect association rules are called negative indirect association rules if its mediator set contains presence and absence of items. In this paper, we propose an efficient algorithm to discover all indirect associations between itemsets. In indirect association mining for itempairs, algorithms require two join operations. To overcome this disadvantage we have proposed a new algorithm to mine indirect associations between itempairs and itemsets. This algorithm features, performing only one join operation and generating indirect associations for pair of items and itemsets. In future we propose to elaborate this work by conducting experiments on large databases to test the scalability and to generate negative indirect association rules. Threshold selection is another issue that needs further investigation.

References

1. Agarwal, R., Aggarwal, C., Prasad, V.V.V.: Depth first generation of long patterns. In: Proceedings of ACM-SIGKDD International Conference on Knowledge Discovery and Data Mining (2000)
2. Agarwal, R., Aggarwal, C., Prasad, V.V.V.: A tree projection algorithm for generation of frequent itemsets. Journal of Parallel and Distributed Computing, Special Issue on High Performance Data (2000)
3. Ramasubbareddy, B., Govardhan, A., Ramamohanreddy, A.: Mining Positive and Negative Association Rules. In: IEEE ICSE 2010, Hefei, China (August 2010)
4. Ramasubbareddy, B., Govardhan, A., Ramamohanreddy, A.: An Approach for Mining Positive and Negative Association Rules. In: Second International Joint Journal Conference in Computer, Electronics and Electrical, CEE (2010)
5. Brin, S., Motwani, R., Ullman, J., Tsur, S.: Dynamic itemset counting and implication rules for market basket data. In: Proceedings of the International ACM SIGMOD Conference, Tucson, Arizona, USA, pp. 255–264 (May 1997)
6. Tan, P., Kumar, V.: Interestingness measures for association patterns: A perspective. In: KDD 2000 Workshop on Postprocessing in Machine Learning and Data Mining, Boston, MA (August 2000)
7. Tan, P., Kumar, V.: Mining indirect associations in web data. In: Kohavi, R., Masand, B., Spiliopoulou, M., Srivastava, J. (eds.) WebKDD 2001. LNCS (LNAI), vol. 2356, p. 145. Springer, Heidelberg (2002)
8. Savasere, A., Omiecinski, E., Navathe, S.: Mining for strong negative associations in a large database of customer transactions. In: Proceedings of the 14th International Conference on Data Engineering, Orlando, Florida, pp. 494–502 (February 1998)
9. Tan, P., Kumar, V., Srivastava, J.: Indirect association: mining higher order dependencies in data. In: Proceedings of the 4th European Conference on Principles and Practice of Knowledge Discovery in Databases, Lyon, France, pp. 632–637 (2000)
10. Wu, X., Zhang, C., Zhang, S.: Mining both positive and negative association rules. In: Proceedings of the 19th International Conference on Machine Learning (ICML 2002), Sydney, Australia, pp. 658–665 (July 2002)

11. Brin, S., Motwani, R., Silverstein, C.: Beyond market baskets: Generalizing association rules to correlations. In: Proc. ACM SIGMOD Intl. Conf. Management of Data, Tuscon, AZ, pp. 265–276 (1997)
12. Brijs, T., Swinnen, G., Vanhoof, K., Wets, G.: Using association rules for product assortment decisions: A case study. In: Proc. of the Fifth ACM SIGKDD Conf. on Knowledge Discovery and Data Mining, San Diego, Calif., pp. 254–260 (August 1999)
13. Cooley, R., Clifton, C.: TopCat: Data mining for topic identification in a text corpus. In: Żytkow, J.M., Rauch, J. (eds.) PKDD 1999. LNCS (LNAI), vol. 1704, pp. 174–183. Springer, Heidelberg (1999)
14. Das, G., Mannila, H., Ronkainen, P.: Similarity of attributes by external probes. In: Proc. of the Fourth ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, New York, NY, pp. 23–29 (1998)
15. Melamed, D.: Automatic construction of clean broad-coverage translation lexicons. In: 2nd Conference of the Association for Machine Translation in the Americas, ATMA 1996 (1996)
16. Reynolds, H.T.: The Analysis of Cross-Classifications. Macmillan Publishing Co., New York (1997)
17. Savasere, A., Omiecinski, E., Navathe, S.: Mining for strong negative associations in a large database of customer transactions. In: Proceedings of the 14th International Conference on Data Engineering, Orlando, Florida, pp. 494–502 (February 1998)
18. Tari, Z., Bukhres, O., Stokes, J., Hammoudi, S.: The reengineering of relational databases based on key and data correlations. In: Spaccapietra, S., Maryanski, F. (eds.) Searching for Semantics: Data Mining, Reverse Engineering, etc. Chapman and Hall, Boca Raton (1993)
19. Winkler, R., Hays, W.: Statistics: Probability, Inference and Decision, 2nd edn. Holt, Rinehart & Winston, New York (1975)
20. Wan, Q., An, A.: An Efficient Approach to Mining Indirect Associations, pp. 1–26. Kluwer Academic Publishers, Boston
21. Li, L., Xu, F., Wang, H., She, C., Zhihua Fan, I.A.M.: An Algorithm of Indirect Association Mining. In: Proceedings of the 2004 International Conference on Intelligent Mechatronics and Automation Chengdu, China (August 2004)

Reaction Attacks in the Matrix Scheme of NTRU Cryptosystem

Rakesh Nayak¹, Jayaram Pradhan², and C.V. Sastry³

¹ Associate Professor in Department of IT,
Sri Vasavi Engineering College, Tadepalligudem,
Andhra Pradesh, India
nayakrakesh8@gmail.com

² Professor, Computer Sciences, Behrampur University,
Behrampur, Odisha, India
jayarampradhan@hotmail.com

³ Professor, School of Computer Science and informatics,
Sreenidhi Institute of Science and Technology,
Hyderabad, Andhra Pradesh, India
cvsastry40@yahoo.co.in

Abstract. An attacker produces a sequence of encrypted messages E , each of which has a distinct probability, however small, of decrypting into a valid message and also a probability of decrypting into an invalid message. The smallest modification that an attacker can make to the cipher-text and can still decrypt it correctly, gives information about the private key used to encrypt the message. In this paper we assume that the attacker knows or gets hold of intermediate text which arises during the process of decryption.

Keywords: Encryption, Decryption, Cipher Text, Private Key, Public Key.

1 Introduction

In a recent preprint, Hall, Goldberg, and Schneier [1] have proposed an attack against several public key cryptosystems based on lattice problems. They call their attack a Reaction Attack. In their paper they describe how to mount a Reaction Attack on a number of cryptosystems, including those suggested by McEliece, Atjai-Dwork, and Goldreich-Goldwasser-Halevi. Since the NTRU public key cryptosystem is also based on an underlying lattice problem, it is natural that a Reaction Attack should exist for NTRU. In this paper we explain how an NTRU Reaction Attack would work and we describe a number of ways in which a user of the NTRU public key cryptosystem can thwart such attacks. Our main point will be that before such an attack can begin to yield useful information, it must deform the encrypted message sufficiently that the attack can be detected by the NTRU user. We also note that such an attack can only hope to succeed if a single private key is used for the decryption of a large number of messages. This is a situation that will rarely arise in an NTRU based system, since it is easy to generate a key, meaning that keys can be changed frequently.

2 Mathematical Preliminaries

2.1 Modular Arithmetic

Let A be an n x n matrix defined as

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

And let p be an integer. Then we define A (Mod p) as

$$A(\text{Mod } p) = \begin{bmatrix} a_{11}(\text{Mod } p) & \cdots & a_{1n}(\text{Mod } p) \\ \vdots & \ddots & \vdots \\ a_{n1}(\text{Mod } p) & \cdots & a_{nn}(\text{Mod } p) \end{bmatrix}$$

If p is a positive integer, and A, B are two matrices then, A is said to be congruent to B modulo m, if A – B is divisible by p denoted by $a \equiv b$ modulo m or simply $a \equiv b \pmod{m}$.

A matrix and congruence with the same modulus may be added, subtracted, and multiplied just as is done with matrix operations.

The properties of modular arithmetic on matrix [13,14] are

(i) $[A \pmod{p} + B \pmod{p}] \pmod{p} = (A + B) \pmod{p}$

(ii) $[A \pmod{p} * B \pmod{p}] \pmod{p} = (A*B) \pmod{p}$

(iii) $A^{-1} \pmod{p} = (A \pmod{p})^{-1} \pmod{p}$

(iv) If $A \equiv B \pmod{p}$ and $C \equiv D \pmod{p}$ then $Ax + Cy \pmod{p} = Bx + Dy \pmod{p}$

for all integers x and y.

(v) If $A \equiv B \pmod{p}$ then $An \equiv Bn \pmod{p}$ for any positive integer n

And the property $[A*(A^{-1} \pmod{p})] \pmod{p} = I$ is already shown in [5].

2.2 NTRU Encryption on Matrix

The NTRU Crypto-system [2] is based on three parameters p, q and N where p is a small prime number and q and p are relatively-prime and N is the degree of the polynomial in the ring of polynomials. Recently the NTRU cryptosystem using a ring of polynomials has been extended [5] for a more compact matrix formalism using modular arithmetic.

Bob chooses two matrices X and Y, where matrix X is an invertible matrix (modulo p). He keeps the matrices X and Y private and generates a public key H as follows:

$$H = p Xq * Y \pmod{q}$$

Where Xq is X^{-1} modulo q or $X * Xq = I \pmod{q}$ When Alice wants to send a message to Bob, she converts the message to the form of binary matrix M (which is of the same order as X and Y). She uses Bob's public key and generates the cipher text E as follows:

$E = H * R + M$ modulo q , where R is a compatible matrix and serves to obscure the original message M . Bob after receiving the encrypted message uses the following procedure to decrypt the message:

$$\begin{aligned} A &= X * E \text{ (modulo } q) \\ &= X * (H * R + M) \text{ (modulo } q) \\ &= X * (p * Xq * Y * R + M \text{ (modulo } q)) \\ &= p * Y * R + X * M \\ \text{Let } B &= A \text{ (modulo } p) = X * M \text{ modulo } p \\ \text{Now } C &= Xp * X * M \text{ modulo } p = M, \text{ the original message.} \end{aligned}$$

3 Motivation

In the Reaction Attacks described in [2], one starts with a valid encrypted message e and creates small modifications $e' = e + \epsilon$. The attacker makes the modification ϵ larger and larger until the modified message e' causes a decryption error. By comparing the ϵ 's that cause decryption errors to those that do not, the attacker gains information about either the plaintext message m underlying e or about the private key used to encrypt m .

An encrypted NTRU message has the form $e \equiv \phi h + m \pmod{q}$. The smallest modification that an attacker can make to e and still have it sometimes decrypt correctly is, say, obtained by adding $-npX_i$ to e i.e., $e' = e + npX_i$ for some $0 \leq i < N$ and some $n \geq 1$. This will cause a decryption failure (a so-called wrap or gap failure) if some coefficient of the intermediate decryption polynomial $a = p\phi g + mf$ is within np of $q/2$ and if $X_i f$ has a corresponding $+1$ coefficient. The important point to observe is that for the correct choice of n , the i 's which cause decryption failure for $e + npX_i$ will reflect (with some shifting and possible duplication) the i 's for which the private key f has a term of the form $+X_i$. Thus the attacker potentially gains information about the $+1$ bits in the private key f . Similarly, using negative values for n may give information about the -1 bits of f .

The Reaction Attack does not compromise the hard mathematical problem underlying the NTRU PKCS, which is the problem of finding the shortest vector in a lattice of high dimension. None-the-less, it is a potentially serious attack for implementations of the NTRU PKCS in hardware or software.

4 Proposed Method

An attacker would try to figure out the private key X , from the encrypted text. The smallest modification that an attacker can make to the encrypted text and still decrypt correctly gives some information about the private key. Let $E' = E + np * m(i, j)$ for some $1 \leq i \leq N$, $1 \leq j \leq N$ and for some n , where $m(i, j)$ is $N \times N$ matrix with 1 at (i, j) th position and the rest zero. In this paper we assume that the attacker knows or the descriptor reveals the value of A which is equal to $X * E$.

Let E be the valid encrypted message that decrypts correctly. So we have all the elements of the matrix A satisfy $a_{\mu} < q/2$ and $a_{\nu} > -q/2$, where a_{μ} is the largest and a_{ν} is the smallest element in the encrypted message.

$$\text{We have } E' = E + np * m[i, j] = \begin{cases} E, & \text{if } m[i, j] = 0 \\ E + np, & \text{if } m[i, j] = 1 \end{cases}$$

$m[i, j]$ is a matrix where only one entry is 1 and the rest of the elements are 0's. We know that $E[i, j]$ lies in $(-q/2, q/2)$.

$$\text{Now calculate } A' \text{ as } A' = X * E' \pmod{q} = X * [E + p * m(i, j)] \pmod{q}$$

$$\begin{aligned} &= [X * E + X * p * m(i, j)] \pmod{q} \\ &= [X * E \pmod{q} + X * np * m(i, j) \pmod{q}] \pmod{q} \\ &= [A + X * p * m(i, j) \pmod{q}] \pmod{q} \text{ -----} \end{aligned} \tag{i}$$

We know that X is a matrix with entries $(-1, 0, 1)$ and $m[i, j]$ is a matrix with only one entry with 1 and rest of the elements are 0's. p is the NTRU parameter and n is the smallest integer which causes wrap failure. q being very large compared to all other parameters, it is safe to assume that $X * np * m(i, j) \pmod{q} = X * np * m(i, j)$.

$$\text{So we can write (i) as } A' = [A + X * np * m(i, j)] \pmod{q}$$

$$\begin{aligned} &= A \pmod{q} + X * np * m(i, j) \pmod{q}. \\ \text{Now find } A' - A &= X * np * m(i, j) \pmod{q}. \text{ ----} \end{aligned} \tag{ii}$$

If we absorb np in $m(i, j)$, we can write the above equation as

$$A' - A = X * Y(i, j) \quad \text{where } Y(i, j) = np * m(i, j)$$

The above procedure can be repeated for different values of i and j in $m(i, j)$ such that $\sum(A' - A) = X * \sum Y(i, j)$ with $\text{Det}(\sum Y(i, j)) = 1$.

Now multiplying Inverse of $\sum Y(i, j)$ to $\sum(A' - A)$ we get the required private key.

4.1 Example

Let $X = \{\{1, 0, 0\}, \{1, -1, 1\}, \{0, 0, -1\}\}$, $M = \{\{0, 0, -1\}, \{0, -1, 1\}, \{-1, 1, -1\}\}$, and $A = \{\{0, 3, 1\}, \{1, 1, 6\}, \{1, -1, 3\}\}$. We know that $p=3$, take $n=3$. Now send different

Table 1. Calculation of A' and $(A-A')$ for a given $m_i(i, j)$

$m_i(i, j) =$	$A' =$	$A' - A =$
$\{\{1, 0, 0\}, \{0, 0, 0\}, \{0, 0, 0\}\}$	$\{\{-3, 3, 1\}, \{-2, 1, 6\}, \{4, -1, 3\}\}$	$\{\{-3, 0, 0\}, \{-3, 0, 0\}, \{3, 0, 0\}\}$
$\{\{0, 1, 0\}, \{0, 0, 0\}, \{0, 0, 0\}\}$	$\{\{0, 0, 1\}, \{1, -2, 6\}, \{1, 2, 3\}\}$	$\{\{0, -3, 0\}, \{0, -3, 0\}, \{0, 3, 0\}\}$
$\{\{0, 0, 1\}, \{0, 0, 0\}, \{0, 0, 0\}\}$	$\{\{0, 3, -2\}, \{1, 1, 3\}, \{1, -1, 6\}\}$	$\{\{0, 0, -3\}, \{0, 0, -3\}, \{0, 0, 3\}\}$
$\{\{0, 0, 0\}, \{1, 0, 0\}, \{0, 0, 0\}\}$	$\{\{0, 3, 1\}, \{4, 1, 6\}, \{1, -1, 3\}\}$	$\{\{0, 0, 0\}, \{3, 0, 0\}, \{0, 0, 0\}\}$
$\{\{0, 0, 0\}, \{0, 1, 0\}, \{0, 0, 0\}\}$	$\{\{0, 3, 1\}, \{1, 4, 6\}, \{1, -1, 3\}\}$	$\{\{0, 0, 0\}, \{0, 3, 0\}, \{0, 0, 0\}\}$
$\{\{0, 0, 0\}, \{0, 0, 1\}, \{0, 0, 0\}\}$	$\{\{0, 3, 1\}, \{1, 1, 9\}, \{1, -1, 3\}\}$	$\{\{0, 0, 0\}, \{0, 0, 3\}, \{0, 0, 0\}\}$
$\{\{0, 0, 0\}, \{0, 0, 0\}, \{1, 0, 0\}\}$	$\{\{0, 3, 1\}, \{-2, 1, 6\}, \{-2, -1, 3\}\}$	$\{\{0, 0, 0\}, \{-3, 0, 0\}, \{-3, 0, 0\}\}$
$\{\{0, 0, 0\}, \{0, 0, 0\}, \{0, 1, 0\}\}$	$\{\{0, 3, 1\}, \{1, -2, 6\}, \{1, -4, 3\}\}$	$\{\{0, 0, 0\}, \{0, -3, 0\}, \{0, -3, 0\}\}$
$\{\{0, 0, 0\}, \{0, 0, 0\}, \{0, 0, 1\}\}$	$\{\{0, 3, 1\}, \{1, 1, 3\}, \{1, -1, 0\}\}$	$\{\{0, 0, 0\}, \{0, 0, -3\}, \{0, 0, -3\}\}$

Table 2. Calculation of $(A' - A)/np$ for a given $m_i(i, j)$

$(A' - A)/np =$
$\{-1, 0, 0\}, \{-1, 0, 0\}, \{1, 0, 0\}$
$\{0, -1, 0\}, \{0, -1, 0\}, \{0, 1, 0\}$
$\{0, 0, -1\}, \{0, 0, -1\}, \{0, 0, 1\}$
$\{0, 0, 0\}, \{1, 0, 0\}, \{0, 0, 0\}$
$\{0, 0, 0\}, \{0, 1, 0\}, \{0, 0, 0\}$
$\{0, 0, 0\}, \{0, 0, 1\}, \{0, 0, 0\}$
$\{0, 0, 0\}, \{-1, 0, 0\}, \{-1, 0, 0\}$
$\{0, 0, 0\}, \{0, -1, 0\}, \{0, -1, 0\}$
$\{0, 0, 0\}, \{0, 0, -1\}, \{0, 0, -1\}$

$m_i(i, j)$ and assume that somehow the attacker knows the value of corresponding A' which is shown in table 1.

In a 3 X 3 binary matrix, there are 84 different matrixes we can get whose determinant is 1. Any matrix with determinant 1 ensures that in its inverse there is no fractional part. This small procedure finds the number of different matrix whose determinant is 1 and its corresponding inverse.

$i=1;$

For[$y=1, y \leq 512, y++$,

$s=$ IntegerDigits[$y, 2, 9$]; // IntegerDigits[n, b, len] gives a list of the base- b digits in
// the integer n and pads the list on the left with zeros to give
// a list of length len .

$m=$ Partition[$s, 3$]; // Partition[list, n] partitions list into non-overlapping
// sublists of length n .

If[Det[m]==1, Print[$i++$, " ", m , " ", Inverse[m]];] // Det[f] gives the determinant.

Let us take another set of m_i 's $\{1, 0, 0\}, \{0, 0, 0\}, \{0, 0, 0\} + \{0, 0, 1\}, \{0, 0, 0\}, \{0, 0, 0\} + \{0, 0, 0\}, \{1, 0, 0\}, \{0, 0, 0\} + \{0, 0, 0\}, \{0, 1, 0\}, \{0, 0, 0\} + \{0, 0, 0\}, \{0, 0, 0\}, \{0, 0, 1\} = \{1, 0, 1\}, \{1, 1, 0\}, \{0, 0, 1\}$. Whose inverse is $\{1, 0, -1\}, \{-1, 1, 1\}, \{0, 0, 1\}$

Add the corresponding $(A' - A)/np$. We get $\{-1, 0, 0\}, \{-1, 0, 0\}, \{1, 0, 0\} + \{0, 0, -1\}, \{0, 0, -1\}, \{0, 0, 1\}$
+ $\{0, 0, 0\}, \{1, 0, 0\}, \{0, 0, 0\} + \{0, 0, 0\}, \{0, 1, 0\}, \{0, 0, 0\} + \{0, 0, 0\}, \{0, 0, -1\}, \{0, 0, -1\} = \{-1, 0, -1\}, \{0, 1, -2\}, \{1, 0, 0\}$

Now $\sum [(A - A')/np] \cdot (\text{Inverse } m_i \text{'s}) = \{-1, 0, -1\}, \{0, 1, -2\}, \{1, 0, 0\} \cdot \{1, 0, -1\}, \{-1, 1, 1\}, \{0, 0, 1\}$
 $= \{-1, 0, 0\}, \{-1, 1, -1\}, \{1, 0, -1\}$. This is the required X .

5 Conclusion

The method described in this paper to obtain the private key is possible if in advertently the attacker gets hold of the intermediate text before final decryption.

References

- [1] Hall, C., Goldberg, I., Schneier, B.: Reaction attacks against several public-key cryptosystems (April 1999) (preprint), <http://www.counterpane.com>
- [2] NTRU Cryptosystem, Technical Reports, the free encyclopedia. NTRU Cryptosystems Inc. (2002), <http://www.ntru.com> Wikipedia
- [3] Silverman, J.H.: Wraps, Gaps, and Lattice Constants NTRU Cryptosystems Technical Report (March 15, 2001)
- [4] Hoffstein, J., Silverman, J.H.: Reaction Attacks Against the NTRU Public Key Cryptosystem, NTRU Cryptosystems Technical Report (June 2000)
- [5] Nayak, R., Sastry, C.V., Pradhan, J.: A matrix formulation for NTRU cryptosystem. In: Proceedings 16th IEEE International Conference on Networks (ICON 2008), New Delhi, December 12-14 (2008)
- [6] Hoffstein, J., Pipher, J., Silverman, J.H.: NTRU: A High Speed Public Key Cryptosystem. Presented At He Hump Session of Euro. Crypt. 1996 (1996) (preprint)
- [7] Hoffstein, J., Lieman, D., Silverman, J.: Polynomial Rings and Efficient Public Key Authentication. In: Blum, M., Lee, C.H. (eds.) Proceeding of the International Workshop on Cryptographic Techniques and E-Commerce (CrypTEC 1999). City University of Hong Kong Press (1999)
- [8] El-Gamal, T.: A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE Transactions on Information Theory 31, 469–472 (1985)
- [9] Hoffstein, J., Pipher, J., Silverman, J.: NTRU: A Ring Based Public Key Cryptosystem. In: Buhler, J.P. (ed.) ANTS 1998. LNCS, vol. 1423, pp. 267–288. Springer, Heidelberg (1998)
- [10] Wells Jr., A.L.: A polynomial form for logarithms modulo a prime. IEEE Transactions on Information Theory 30, 845–846 (1984)
- [11] Diffie, W., Hellman, M.E.: New directions in cryptography. IEEE Information theory June 23- 25 (1975); IEEE International Symposium on Information Theory, Sweden, June 21-24 (1976)
- [12] Brassard, G., Bratley, P.: Fundamentals of Algorithm, PHI (1996)
- [13] Horowitz, E., Sahani, S., Rajasekharan, S.: Fundamental of Computer Algorithm, Galgotia (1998)

Process Corner Analysis for Folding and Interpolating ADC

Shruti Oza¹ and N.M. Devashrayee²

¹ EC Department, Kalol Institute of Technology & Research Centre, Nr. Highway,
Kalol-382721, Gujarat, India
swajan_2004@yahoo.com

² EC Department, IT-Nirma University of Science & Technology, S.G. Road,
Ahmedabad Gujarat, India
nalin_deepika@yahoo.com

Abstract. Folding and Interpolating ADC have been shown to be an effective means of digitization of high bandwidth signals at intermediate resolution. The paper designs Folding and Interpolating ADC using cascaded folding amplifier to observe the effect of process variations. The primary circuit effects, resulted from process variations that are liable to degrade the performance of ADC are transistor mismatch, resistor mismatch and amplifier-comparator offsets. The device matching in reference generation, folding amplifier, interpolation and comparator offsets specify overall performance of ADC. Since the mismatches are random, Monte Carlo Analysis is used to estimate the linearity performance. In this paper the design is simulated using 0.35 μm , 3.3V to study the effect of process corners.

Keywords: Cascaded Folding Amplifier, Comparator, Encoder, Folding and Interpolating ADC, Interpolation, Process Corners.

1 Introduction

ADC is one of the most important building blocks to transform analog signals to digital signal process systems. For high-speed application, Flash ADC is widely used. However, N-bit Flash ADC needs $2^N - 1$ comparators, which consume large power and occupy area. Folding architecture is an alternative approach to reduce the complexity of Flash ADC. Before the outputs of the preamps are fed into comparators, folding amplifiers are inserted. Folding amplifier combines the outputs of several preamplifiers and generates folding waveforms, which contains information of those preamplifiers. After folding processing, one comparator deals with more quantization levels. Hence, the number of the comparators is reduced. The number of comparators required for a Folding ADC decreases as the folding order increases. The architecture still keep high conversion rate [1-4].

Another attractive feature of Folding and Interpolating ADC is that high-speed sample and hold amplifier is optional due to parallel operation of fine and coarse converter. However, in most of the papers coarse converter is similar to Flash ADC.

Liu et al proposed design of coarse converter using folding circuit similar to fine converter [4]. Wenzek suggested cascaded dummy differential amplifiers with similar structure like those in the folding stage of the fine converter to overcome the different latencies of coarse and fine converter [5].

Systematic and random variations in process are posing a major challenge to the future high performance VLSI design. Variation in the process parameters such as impurity concentration densities, oxide thickness and diffusion depths caused by non-uniform conditions during deposition and/or during diffusions of the impurities. Variations in the dimensions of the devices are due to limited resolution of the photolithographic process [6-8].

The main focus over here is to observe effect of process parameter variations at device level on the performance of Folding and Interpolation ADC. Section-2 describes design of Folding and Interpolating design with details of implemented folding amplifier, interpolation technique, comparator section, coarse converter and encoder. Section-3 discusses simulation results through Monte Carlo Analysis with wide range of randomly chosen device parameter.

2 Design of Folding and Interpolation ADC

Process variations are due to manufacturing phenomena and static in nature. Process variations are deviations from intended or designed values for the structural or electrical parameters of concern. The process variations can be lot-to-lot, wafer-to-wafer (inter process), die-to-die or within a die (intra process). The paper focuses on Front-End of Line (FEOL) variations, which refers to the variations at device level. The major sources of FOEL variations consist of transistor gate length and gate width variations, gate oxide thickness variations, doping-related variations, etc. The threshold voltage can vary due to changes in oxide thickness, substrate, polysilicon and implant impurity levels and surface charge [6-8]. These process parameter variations result in variation of power and delay and non-linearity in ADC. Therefore it is very important to observe effect of these parameters on the performance of ADC. The effects of process variation can be observed through typical, slow and fast transistors.

The concept of Folding ADC was first introduced by Arbel and Kurz in 1975. The main motivation was the dramatic reduction of the number of comparators required in the design. In Folding and Interpolating ADC, number of zero-crossing points Z is determined by the following equation:

$$Z = N_F \cdot F_F \cdot I \quad (1)$$

Where N_F is the number of primary folding waveform, F_F is the folding factor and I is the interpolating rate. The choice of N_F , F_F and I play very important role in design of Folding ADC.

The aim of the fine converter inside the Folding and Interpolating ADC is to process the input signal, either a sampled or an unsampled signal, and to resolve it into the lower significant bits of the word length of the whole analog-to-digital converter. The output signals of the fine converter are then synchronized with output signals of the coarse converter and combined to yield the output of the complete converter. The fine

converter is composed of a folding stage, an interpolation stage, a comparator stage, and an encoder as shown in Fig. 1(a).

Figure 1(a) shows block diagram of implemented 6-bit Folding and Interpolating converter. Total 64 zero crossing points are obtained by selecting folding factor $F_F=8$, Interpolating Factor $I=2$ and Number of Folding block $N_F=4$. The Fine converter uses cascaded folding amplifier with folding factor $F_F=8$. The encoder generates 4 fine bits while 2 coarse bits are generated using pre-processing folding amplifier.

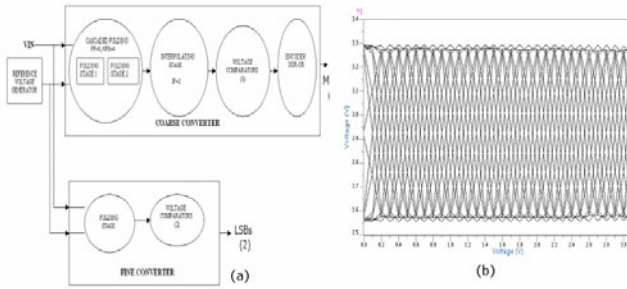


Fig. 1. (a) Folding and Interpolating ADC (b) zero crossing points

The function of a folder is to separate the input signal range into several intervals. By gradually developing higher folding factor using cascaded folding amplifier helps in achieving a high folding degree while avoiding the high gain and/or bandwidth requirements of each folder. The first stage is implemented with folding factor=4. The goal of the folding amplifier is to overcome all the limitations of existing design, utilization of all transistor and minimizing power and settling time. Figure 2 (a) shows implanted folding amplifier with folding factor=4 for the first stage. The second stage is shown in Figure 2 (b). The differential output of first stage is two signals FA_P and FA_N, applied as input for the second stage. Such two folding amplifiers' output are applied to the second stage to generate FF=8 as shown in figure. The total power required by cascaded folding amplifier is only 180uW.

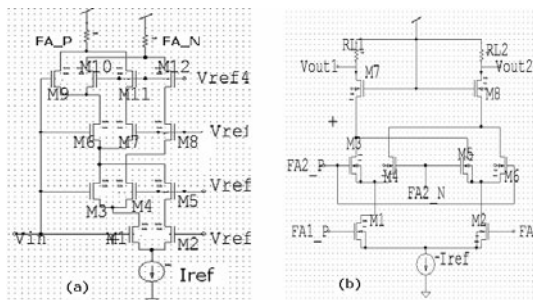


Fig. 2. (a) First stage of Cascaded Folding Amplifier with FF=4 (b) Second Stage

Interpolation is often employed to generate extra folding waveforms without increasing number of folding amplifiers. There are basically two methods to interpolate the folded signals, namely voltage-mode (resistive) interpolation and current-mode interpolation. The voltage-mode interpolation can be implemented using a resistance ladder. The advantage of a voltage-mode interpolation over current mode is its design simplicity and low power operation. Figure 1 (b) shows 64 zero crossing points generated by folding and interpolation stages, using $N_F = 4$, $F_F = 8$ (first stage $F_F = 4$, second stage $F_F = 2$) and interpolating factor=2.

In order to achieve low power, high-speed operation of the design, comparator is another important block. When a comparator must drive a significant amount of output capacitance in very short times, it is advisable to follow the latch by circuits that can quickly generate large amount of current. A high-speed comparator following these principles is designed in Fig. 3(a). The first stage is a low gain, high bandwidth preamplifier that drives a latch (Decision Circuit). The latch outputs are used to drive an inverter (Output Stage). Figure 3(b) shows output of fine comparator (cyclic thermometer) with process variations.

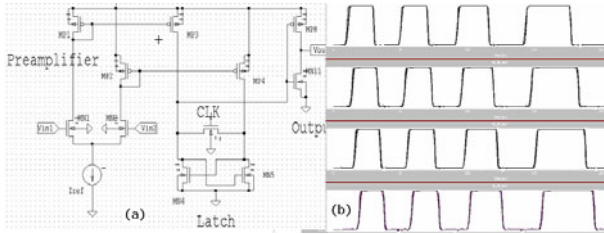


Fig. 3. (a) Voltage Comparator (b) Cyclic Code (output) of Fine Converter

The coarse converter divides input into several intervals and provides MSBs. Instead of using traditional Flash architecture for implementing coarse converter, folding circuit is used as pre-processing. This helps in reducing number of comparators and encoder logic.

The cyclic output (Fig. 3(b)) can be easily converted into binary code. To convert cyclic code into binary, logic shown in equations (2) can be used, comparing binary and cyclic code. The encoder based on XOR-OR logic can be used to convert the cyclic code into binary. The simulation results of the encoder-digital output with process variations are shown in Fig 5(a).

Encoder logic:

$$\begin{aligned}
 B3 &= C7 \\
 B2 &= C7 \oplus C3 \\
 B1 &= C7 \oplus C5 + C3 \oplus C1 \\
 B0 &= C7 \oplus C6 + C5 \oplus C4 + C3 \oplus C2 + C1 \oplus C0 \dots
 \end{aligned} \tag{2}$$

Where B0 to B3 are 4-bit fine converter output and C0 to C7 are 8-bit fine comparator output.

3 Monte Carlo Analysis

The folding architecture imposes tight constraints on device matching as mismatch results in shifting zero crossing points from the ideal values. Mismatch in first stage of folding creates large performance degradation. For accurate zero crossings, folding and interpolating stages must be well matched, which depends on differential pair mismatch, current source mismatch and resistance mismatch.

The mismatch within a folding amplifier includes mismatch in the threshold voltage and beta, mismatch between various tail current sources and mismatch of slope of two folding signals. The input offset voltage depends on load resistor mismatch, transistor dimensions mismatch and threshold voltage mismatch. In this paper, V_{TH} is varied randomly with 6% variation and μ_0 is varied randomly with 5% on each side from their standard value. During simulations, Monte Carlo runs are performed 100 times to get accurate results. Figure 4(a) shows results for output of folding amplifier. Among all the corners, the worst-case error = 0.22%.

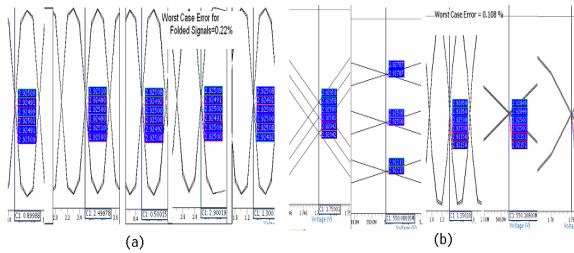


Fig. 4. Process Corner Analysis (a) for Folding Amplifier (b) for Interpolation Stage

Fig. 4(b) shows the simulated results for interpolated signal due to variations in process parameters on each side from their standard values. For 100 Monte Carlo runs, worst-case error is 0.108% interpolation, which is half than worst-case error for folded signal. This indicates that interpolating helps in reducing non-linearity through averaging.

Figure 5 (b) shows switching power of converter at 100MHz clock frequency, 1MHz input. The range of average power variation is 10.05-10.6mW and peak power is 19.88-24.9mW. The dc power required is in range of 10.05mW-17.9mW.

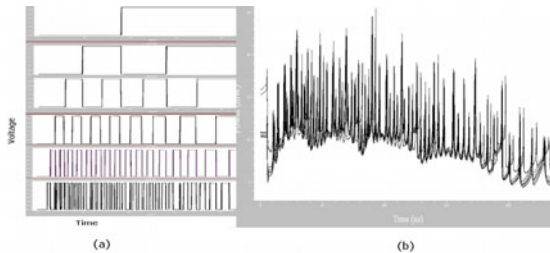


Fig. 5. (a) Digital Output of encoder (b) Switching Power due to Process Variations at Clock Frequency of 100MHz, Input Frequency of 1MHz

4 Conclusion

In this paper, 6-bit low power Folding and Interpolating ADC is designed with minimum comparators and hardware. The architecture uses novel cascaded folding amplifier to achieve high folding factor. The design uses folding pre-processing circuit for both coarse and fine converter. To achieve low power operation, folding block, comparator, encoder, and coarse converters are optimized. The design is simulated using 0.35 μ m technology at 3.3V. The effect of process variations is observed using 100 Monte Carlo runs. The worst-case error in zero crossings after folding stage is 0.22%. The error is reduced due to resistive interpolation and is 0.108%. The worst-case error at final stage (MSBs and LSBs) is 2% in LSB. No error (error in codes, spikes or pulse duration/delay variations) is observed for remaining bits of converter. Due to cyclic output of folding block comparison, number of comparators required is only 10 in case of 6-bit converter. The design also helps in reducing latency difference of coarse and fine converter.

References

1. Nauta, B., Venes, A.: A 70MS/s/110-mW 8-b CMOS Folding and Interpolating A/D Converter. *IEEE Journal Of Solid-State Circuits* 30(12), 1302–1308 (1995)
2. Thirugnanam, R., Ha, D.S., Choi, S.S.: Design of a 4-bit 1.4 GSamples/s Low Power Folding ADC for DS-CDMA UWB Transceivers. In: *IEEE International Conference on Ultra-Wide Band, ICU 2005*, pp. 536–541 (2005)
3. Kim, K.M., Yoon, K.S.: An 8-Bit CMOS Current-Mode Folding And Interpolation A/D Converter With Three-Level Folding Amplifiers. In: *IEEE 39th Midwest Symposium on Circuits and Systems*, vol. 1, pp. 201–204 (1996)
4. Liu, Z., Wang, Y., Jia, S., Ji, L., Zhang, X.: Low-Power CMOS Folding and Interpolating ADC with a Fully-folding Technique. In: *7th International Conference on ASIC, ASICON 2007*, pp. 265–268 (2007)
5. Lin, I.K.-L., Kemna, A., Hosticka, B.J.: *Modular Low-Power, High-Speed CMOS Analog-To-Digital Converter of Embedded Systems*. Kluwer Academic Publishers, Dordrecht (2003)
6. Bowman, K., et al.: Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration. *IEEE Journal of Solid-State Circuits* 37, 183–190 (2002)
7. Borkar, S.: *Parameter Variations and Impact on Circuits & Microarchitecture*. C2S2 MARCO review (2003)
8. <http://www.britannica.com/bps/..18/..Chapter-2-Process-Variations>

Fast Near-Lossless Image Compression with Tree Coding Having Predictable Output Compression Size

Soumik Banerjee and Debashish Chakroborty

Department of Computer Sc. & Engineering.
St. Thomas' College of Engg. & Tech., Kolkata, India
{soumik.stcet, sunnydeba}@gmail.com

Abstract. Image compression, in the present context of heavy network traffic, is going through major research and development. Traditional entropy coding techniques, for their high computational cost is becoming inappropriate. In this paper a novel near-lossless image compression algorithm had been proposed which follows simple tree encoding and prediction method for image encoding. The prediction technique uses a simple summation process to retrieve image data from residual samples. The algorithm had been tested on several gray-scale standard test images, both continuous and discrete tone, and had produced compression comparable to other state-of-the-art compression algorithms. The output compressed file sizes had shown that they are independent of image data, and depends only on the resolution of the image, an unique property that can be exploited for networking bandwidth utilization.

Keywords: Modelling – coding architecture, median filtering, residual samples, tree encoding, progressive transmission.

1 Introduction

Image compression researches are conducted to optimize storage space requirements for efficiency in storage and data-transfer over networks [1]. The lossless concept of image compression [2] although produces errorless output, cannot provide excellent compression. The lossy concept exploits the fact that human cognition is unable to detect very small intensity variation in small areas of an image. Although this concept provides excellent compression but sometimes loss of image information becomes unacceptable, for instance in medical image processing or image archival where machine based image processing is involved.

Recently, the research interest in this field had been diverted to development of near-lossless compression methods. Algorithms like LOCO-1[5], JPEG-LS[3], JBIG[4] and FELICS[7] have been developed on this concept. All these algorithms follow the *modeling-coding* architecture.

Firstly, image data was modeled using prediction algorithm like DCT[8] or DWT[9] and the deviation of predicted value from actual, called residual sample, was stored. Then the residual samples were encoded using entropy encoding techniques like Adaptive Huffman, Hierarchical Interpolation or Tree encoding[6]. These

approaches involved intensive computation due to the statistical data manipulation and encoding techniques employed by them. Each of the standards provides a faster version, which results in a trade-off in compression quality.

In this paper a new compression technique for gray-scale images had been proposed. The technique involves three basic steps: 1) *image smoothing* 2) *tree encoding* of the image information and 3) *representing intensity of tree using ASCII character* set for optimized storage requirement. The computational cost was low since only linear search and subtraction operation was involved in image modeling part. The simulated results on standard test images had given comparable or better compression ratio than the standard algorithms.

The remaining paper had been organized as follows: section 2 discusses the detailed algorithm, section 3 presents the simulation results and in section 4, the paper is concluded.

2 Proposed Algorithm

2.1 Image Compression

The proposed algorithm complied with the *modeling-coding* architecture. The image had been smoothed to emphasize intensity correlation (section 2.1.1). The image modeling part had been done using residual calculation from maximum intensity in blocks of the image (section 2.1.2). The encoding of thus produced residual samples had been done using tree encoding (section 2.1.2). Consequently, data representation completed the process (section 2.1.3).

2.1.1 Image Smoothing

The image had been smoothed using median filtering on 2X2 masks throughout the image. This step ensured that the truncation of the image tree did not result in drastic loss of image information since median filtering emphasized on statistical correlation among pixels.

2.1.2 Modeling and Progressive Encoding of Image Data

The smoothed image had been encoded in a tree architecture using 4 children (one block) at each level for next level node determination. In the proposed method, consecutive blocks are selected and according to the following steps the image was encoded:-

1. Calculate the maximum intensity value in the selected block.
2. Subtract all other intensities in the block from this value.
3. The last intensity will be involved in further tree encoding at the next level, hence it is not stored at the current loop of execution.
4. The remaining three are stored using character representation (section 2.1.3).

The conceptual idea behind the encoding technique is explained in figure1.

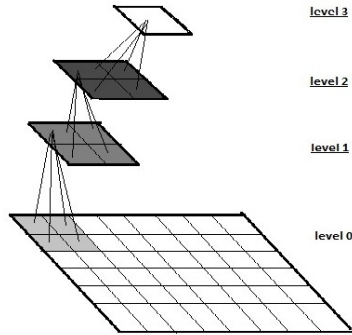


Fig. 1.

As shown in the figure, at each level the resolution of the image gets halved. The resultant tree (expanded to 2 levels) for a 256X256 image is shown in fig 2.

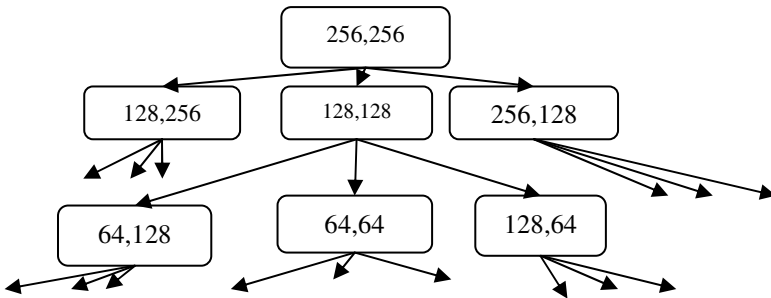


Fig. 2.

2.1.3 Residual Sample Coding

The maximum intensity, in each block, had been selected for quantization (refer to 2.1.2). Hence, residual samples were in the range of $[0 - 255]$ since only 8 bit images have considered. 8 bit ASCII character set defines 255 character symbol set. Of these, three symbols were found unfit for the purpose of intensity representation. ASCII symbols, corresponding to each intensity value, had been used in the sample encoding. The last three intensities had been assigned 2-character hybrid symbols. Also the subtraction for the local maximum intensity ensures that no negative value appears, which otherwise would have to be represented using symbols thereby reducing the resultant compression ratio.

Alternate Strategies for Achieving Greater Compression

The image tree formed can be truncated at several levels. Default strategy was to truncate the tree at level 0 because of high correlation of intensity values in 2X2

masking window. This concept was extended to level 1 tree truncation, which resulted in a smoothed image with 4X4 masking. The second approach, although notably increased the compression ratio, resulted in comparatively higher error of deviation (refer section 3, table 1 and 2).

2.2 Image Decompression

The decompression of the compressed file was done following steps discussed below:

1. Generated the symbol table as discussed in section 2.1.3.
2. Read symbols from compressed file and found the corresponding intensity values. Hence the image tree was remade.
3. Found the maximum in a parent-children block and subtract other values from the maximum value to regenerate original image matrix.
4. Use the last level leaf node values to regenerate the entire image matrix, which is the median filtered (2X2 masking) output of the original image.

3 Simulation Results

Some definitions that were used to study and compare results with other techniques:

$$\text{Bits/Pixel (BPP)} = (8 \times \text{Compressed File Size}) / (\text{Actual File Size}) \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{1}{\text{row} \times \text{col}} \left[\sum_{i=0}^{\text{row}} \sum_{j=0}^{\text{col}} \{f(i, j) - \hat{f}(i, j)\}^2 \right]} \quad (2)$$

The algorithm had been applied on standard gray-scale test images, both continuous-tone and discrete-tone types. The compression percentage achieved had varied from 96-98% for different test images. The test images were shown in figure 3.



Fig. 3. The test image set: (from left to right) lena(512X512); mandrill (256X256); discreet (64X64); peppers(256X256); nasa(512X512); jet (128X128)

The error in output generated had been shown in table 2. The scale of reference was Root Mean Square Error (RMSE), as defined above. The tabulation was done for RMSE values acquired for level 0 and level 1 image tree truncation (refer to fig1).

The experimental results suggested that the output size was predictable. For a 512X512 8bpp image, the output size was 48KB exact in each instance for level 0 truncation of image tree. For level1 truncation the size was 12.0 KB, although this

Table 1. The comparative study of detailed results of the compression output in BPP of the proposed algorithm, with level 0(L0) and level 1(L1) truncation with JPEG 2000

Image name	Lena	Mandrill	Peppers	Nasa	Discreet	Jet
JPEG 2000	0.46	0.68	0.41	0.28	0.56	0.12
PROPOSED (L0)	0.39	0.42	0.39	0.24	0.48	0.38
PROPOSED (L1)	0.09	0.13	0.10	0.11	0.05	0.09

Table 2. The RMSE result for level 0(L0) and level 1(L1) truncation of the original image tree

Image	Lena	Mandrill	Peppers	discreet	Nasa	Jet
RMSE (L 0)	15.35	12.33	17.36	25.09	5.54	6.46
RMSE (L 1)	20.24	17.35	21.54	29.09	7.82	10.98

increased the error level, as discussed in table 2. The predictability of the output image size could be used as a great advantage for network utilization by helping in pre-planning resource allocation.

4 Conclusion

In this paper a new approach of compressing image had been presented using the *modelling-coding* architecture. The computation time was comparatively much lower than the standard algorithms, which involved complex algorithm (DCT approach in JPEG, DPCM in CALIC) for image modelling. The results showed encouraging compression ratio and information loss was also within moderate range even for level 1 truncation strategy of the proposed method. The algorithm ensured fast encoding, predictive output result and simple progressive transmission based decoding which made it suitable for network data-transfer or fabrication on mobile devices.

References

- [1] Salomon, D.: Data Compression The Complete Reference, 3rd edn. Springer Press, Heidelberg
- [2] Halder, A., Chakroborty, D.: An Efficient Lossless Image Compression Using Special Character Replacement. In: ICCET 2010, Jodhpur, Rajasthan, India, November 13-14, pp. E-62 – E-67 (2010)

- [3] Acharya, T., Tsai, P.S.: JPEG 2000 standard for image compression (2000)
- [4] Wu, X.: Context-Based, Adaptive, Lossless Image Coding. *IEEE Trans. Comm.* 45(4)
- [5] Weinberger, M., Seroussi, G., Sapiro, G.: The LOCO-I Lossless Image Compression Algorithm
- [6] Cai, H., Li, J.: Lossless Image Compression with Tree Coding of Magnitude Levels, 0-7803-9332-5/05 ©2005 IEEE
- [7] Howard, P.G., Vitter, J.S.: Fast and Efficient Lossless Image Compression. In: *IEEE DCC 1993* (1993)
- [8] Watson, A.B.: Image Compression Using the Discrete Cosine Transform, NASA Ames Research Centre. *Mathematica Journal* 4(1), 81–88 (1994)
- [9] Ansari, M.A., Anand, R.S.: DWT based Context Modelling of Medical Image Compression. In: *XXXII National Systems Conference, NSC 2008, December 17-19 (2008)*

Over Load Detection and Admission Control Policy in DRTDBS

Nuparam¹ and Udai Shanker²

¹ Computer Science and Engineering Department, FGIET,
Raebareli, 229001, U.P. India

² Computer Science and Engineering Department, MMMEC,
Gorakhpur, 273010, U.P. India
{nrcua80, udaigkp}@gmail.com

Abstract. Today's Real Time System (RTS) is characterized by managing large volume of distributed data making real time distributed data processing a reality [1, 2]. The demand for real time data services is increasing in many large scale distributed real time applications. The transaction work load in DRTDBS may not be balanced and the transaction access pattern may be time varying and skewed. Hence, computation workload on large scale distributed system can lead to large number of transaction deadline misses [3, 4]. Hence, efficient database management algorithm and protocol for accessing and maintaining data are required to satisfy timing constraints of transaction supported applications. In this paper, an algorithm has been proposed for admission control in DRTDBS consisting of local controller and global load balancer working at each site, which decide whether to admit or reject the newly arrived transaction. The simulation results show that the new algorithm successfully balances the workload in DRTDBS [5, 6].

Keywords: Admission control, local controller, global load balancer, distributed data processing, computation workload.

1 Introduction

In recent year, we have seen the emergence of large scale distributed real time database system, which embedded in advance traffic control, factory automation, global environment control and nation-wide electrical power grid control. DRTDBS can relieve the difficulty of developing data intensive real time applications by supporting the logical and temporal consistence of interrelated database distributed over a computer network via transaction management. They support transaction that has explicit timing constraints which is expressed in the form of dead line. A transaction is considered to have finished executing if exactly one of two things occurs: either its primary task is completed (success fully completed) or its recovery block is completed (safe termination). Committed transaction brings a profit to the system whereas a terminated transaction brings no profit. The goal of over load detection and admission control policy employed in the system is to maximize the profit [7, 8, 9]. The presence

of multiple sites in the distributed environment raises issue that are not present in centralized system. In large scale distributed environment it is a challenging to provide data services with guarantees, while still meeting temporal requirement of transaction. One main difficulty lies in long and highly variable remote data access delays. Large scale distributed real time database system utilizing wide geographical area have to use a network that share by many participant for cost effectiveness. Second major challenge involves the complex interaction among a large number of nodes, which can incur unpredictable work load for each node. Transaction work load fluctuation causes uneven distribution of workload among the sites even if on the average, all sites receive similar amount of workload. The third challenge is the data dependent nature of transaction. End to end transaction access pattern may be time varying and skewed. This can be achieved only by timely access to remote data and timely processed of centralized data.

The system architecture consists of overload detection and admission control of scheduling transaction which provide early notification of failure to submitted transaction that are deemed not valuable or in capable of completing in time, when transaction is submitted to the system. An admission control policy is employed to decide whether to admit or reject that transaction. Once admitted, a transaction is guaranteed to finish executing before its dead line. When any site is detecting overload then it distributes the load to other site. In DRTDBS there are two type of transaction, global and local. The global transactions are distributed real time transaction executed at more than one site whereas the local transaction executes at generation site only.

2 Distributed Real Time Database Model

2.1 Real Time Database Model

We focus our study on medium scale distributed database since the load balancer need full information from every site to make accurate decision. Several applications that required distributed real time data services fall in that range for example a ship board control system which control navigation and surveillance consist of 6 distributed control unit and 2 general control consoles located throughout the platform and linked together via a ship wide redundant Ethernet to share distributed real time data and coordinate the activity [10].

2.2 Distributed Real Time Database Model

The performance of the system is evaluated by developing two simulation models for DRTDBS. The first one is for main memory resident DRTDBS which eliminate the impact of different disk scheduling algorithm on the performance. Since main memory database system, we have also developed another model followed by description of various components such as system model, network model, cohort execution model, database model [11]. In our system model, data object are divided into two type namely temporal data and non temporal data. Temporal data are the sensor data from physical world. Each temporal data object has a validity interval and is updated by periodic sensor update transaction. Non –Temporal data object do not have validity intervals and therefore there are no periodic system updates with them.

3 A New Admission Control Policy Architecture

A database system can be overloaded if many user transactions are executed concurrently. As a result computational resources such as CPU cycle and memory spaces can be exhausted. More ever many transactions can be blocked or aborted and restarted due to data contention also. Fig. 1 shows the architecture of a new admission control policy in DRTDBS. The architecture has 4 layer, remote data access layer, QoS enforcement layer, real time database management layer and DRTDBS layer. In DRTDBS layer does exact work as to admitted or reject the transaction All the transaction are submitted in arrival queue and overload admission controller check the incoming transaction as the system resources is already hold by other transaction.

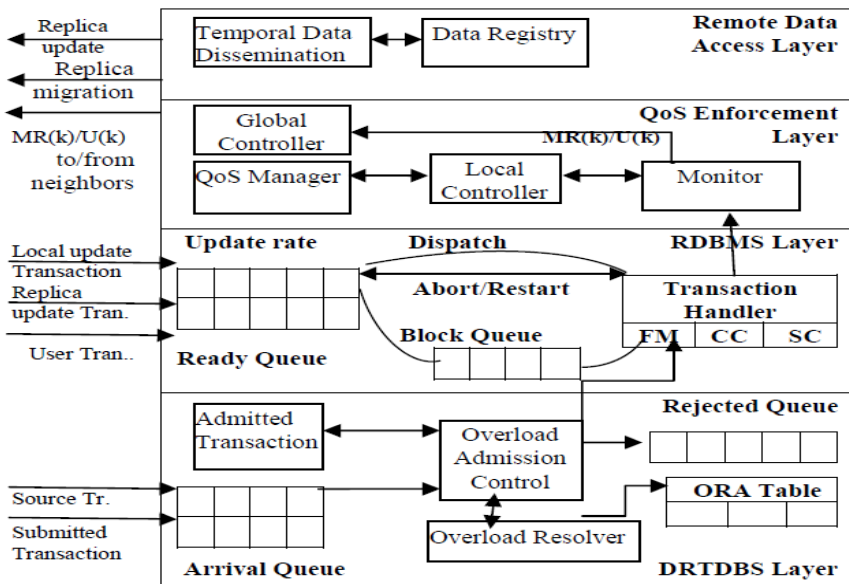


Fig. 1. The System Model for New ACP Protocol

Overload Resolver send current transaction to the ORA (Overload Resolver Array) table. After getting the response from the cohort the overload admission controller interact with the transaction handler (TH) and rejected queue.

The real time database layer does typical real time transaction handling; the incoming transaction are dispatched and processed by transaction handler. The transaction handler consists of a concurrency controller (CC) a freshness manager (FM) and a scheduler (SC). In the SC, update transaction are scheduled in the low priority queue. Update transaction are either updates from local sensors to local data objects. Within each queue, transactions are scheduled with Earliest Deadline First (EDF)[12, 13].

The third layer is guaranteed the desired miss ratio even in the presence of unpredictable workload; QoS enforcement layer exploits two feedback control loops. It has local controller (LC) which operates on the local transaction and global controller (GL) which operate on the global transaction.

The remote data access layer enables transparent access to remote data within a bounded communication time. Remote temporal data are replicated locally to provide timely access to them.

4 Algorithms for Centralized Admission Control in DRTDBS

In each node, there are a local miss ratio controller and local utilization controller. The local miss ratio controller takes the miss ratio from latest sampling period, compare them with serial execution time and compute the local miss ratio control signal δL_{misr} used to adjust the target utilization at the next sampling period. The equation used to derive δL_{misr} is as follows

$$\delta L_{misr} = \sum_{i=1}^n (I_{\frac{M_{ri}}{P}}) \times (M_{ri} - M_{rsi}) + \sum_{i=1}^n (I_{\frac{L_{tmri}}{P}}) \times (L_{tmri} - M_{rsi}) \text{ ----- (1)}$$

M_{ri} is the miss ratio of class i transaction of last period and L_{tmri} is the long term average miss ratio of class i transaction. M_{rsi} is the specified miss ratio requirement by the serial execution specification. And n is the specified serial execution level.

$I_{\frac{M_{ri}}{P}}$ and $I_{\frac{L_{tmri}}{P}}$ are two controller parameter. In order to prevent under-utilization a utilization feedback loop is added. At each sampling period, the local utilization controller compares the utilization and generates the local utilization control signal δL_{util} using equation.

$$\delta L_{util} = I^{util} \times (L_{util} - L_{utilpset}) + J^{util} \times (LT_{util} - L_{utilpset}) \text{ (2)}$$

L_{util} is the CPU utilization of last sampling period and LT_{util} is the long term average CPU utilization of the system. $L_{utilpset}$ is the preset CPU utilization threshold. I^{util} & J^{util} are controller parameter.

4.1 Load Balancing Factor

The load sharing process is guided by the load balancing factor (LBF). The LBF at each node is an array of real number which denotes the amount of workload the local node transfer to other node during the next sampling period.

$$T_{\text{remain work (a)}} \rightarrow T_{\text{trnsfer (a,i,j)}} \text{ (3)}$$

The left expression stands for remaining executing (predicted) time of transaction ‘a’ in the original node and its value is the difference between the transaction ‘average executing time and executing time of transaction ‘a’. The value of right expression stands for the time cost of transferring from node i to node j of transaction ‘a’ which is determined by:

$$T_{\text{transfer (a, i, j)}} = T_{\text{code (a, i, j)}} + T_{\text{data (a, i, j)}} \text{ (4)}$$

There into, $T_{\text{code (a, i, j)}} = \text{sizeof_code(a)}/R$, $T_{\text{data (a, i, j)}} = \text{sizeof_data(a)}/R$.

The value of sizeof_code(a) stands for the total size of executing environment parameter and log of transaction ‘a’. The value of sizeof_data(a) stands for the size of fetched data and immediate results by transaction ‘a’ in the current database server node i . R stands for the average network transferring rate.

5 Algorithm for Global Load Balancing in Decentralized DRTDBS

5.1 Work Load Transfer Test

The first step is to test whether there exist load transfer between nodes. To do that we calculate the mean deviation of M_{ri} from different node

$$\text{Mean Deviation} = \sum_{k=1}^n \text{ABS}(M_{ri}) - \text{Mean}(M_{ri}) \quad \text{-----} \quad (5)$$

Where M_{ri} is the Miss ratio of node k. $\text{ABS}(M_{ri})$ returns the absolute value of M_{ri} and $\text{Mean}(M_{ri})$ returns the mean of M_{ri} , n is the nodes in the system. The mean deviation of M_{ri} is a measure for workload balance in the system.

5.2 LBF Adjustment

The LBF adjustment is divided into two cases, depending on whether there is a load transfer among the node.

Load Imbalance -: When there is load transfer in the system i.e the mean deviation of M_{ri} is larger than the threshold, it is necessary to share the load between nodes. The load balancing algorithm at the overloaded nodes will shift some workload to the less loaded nodes. A node i is considered to be overloaded compared to other nodes if and only if the difference between its MRI and MRI mean is larger than the present mean deviation threshold. i.e. as follow

- When the difference of MRI and MRI Mean $(T_1) \geq \text{Mean Deviation Threshold } (T_2)$
- When the difference of MRI and MRI Mean < 0
- $0 \leq \text{When the difference of MRI and MRI Mean } (T_1) \leq \text{Mean Deviation Threshold } (T_2)$

5.3 Algorithm Description

In order to realize the algorithm we introduce three data structures to record the node's states namely R_{queue} , S_{queue} , and O_{queue} , and several other variables such as the maximum value of probing time to avoid too probing effect to the system performance.

5.3.1 Knocking State

A node i is considered overloaded if the difference of MRI and MRI Mean $\geq \text{Mean Deviation Threshold } T_2$. It divides into two parts: sender side and receiver side. The sender side execute the following operation when a new transaction is created in node i namely $T_i(m+1)$, which cause the node i to be sender.

- Then select a node j from LBF R_{queue} which satisfies:
 $T_{\text{transfer}(a,i,j)} = \text{Min}(T_{\text{transfer}(a,i,j)})$,
 Therein $j=1,2,3,4,\dots,n$. n is the recorded receiver number in structure LBF R_{queue} . It is necessary to judge if the node can be receive in terms of the probing resut.

- II. IF (node j is the receiver)
 Puts the transaction $T_{i(m+1)}$ into the node j LFB queue T_j and transfer the log, then the algorithm terminate
 ELSE
 Remove the node I from LBF R_{queue} and put it into the queue of LBF head of sender queue S_{queue}
- III. Select another node from LBF to repeat the process until one of the following condition satisfied. If then terminate the algorithm
 III.I R_{queue} is empty
 III.II Probing time is beyond the maximum time
 III.III Node I is no longer a sender
 If receiving probe information then the receiver's side executes the following operation.
- IV. Remove the node from the current LBF queue to the head of the S_{queue}
 Send a message about node j's states to node i.

5.3.2 Responding State

A node i is considered less overloaded if the difference of MRI and MRI Mean < 0 . It also divides into two parts: receiver's side and sender's side, the first part executes the following operation if the node i becomes a receiver when a transaction, T_{ik} , finished or be removed.

- I. The node i sends the probing message to node j which is in the head of S_{queue} . And judge if node j is a sender.
- II. IF (node j is a sender)
 IF (Exists a transaction that can be transferred to node j)
 Transfer the transaction T_{jr} which resides in node j before to the transaction LBF queue of node i then executes the T_{jr} from the beginning in the node i. The log will be sent i at the same time. After that, the algorithm terminates.
 ELSE (Remove the node j to the tail of S_{queue})
 ELSE
 Removes the node j from S_{queue} and put into the queue's head of R_{queue} or O_{queue} Select another new node to repeat I and II until one of the following conditions satisfied, if then, the algorithm terminates.
 a. S_{queue} is empty
 b. The probing time beyond the limit
 c. Node i is no longer a receiver.

When a node j receives the probe information from node i, then sender's side execute the following action:

- III IF (node j is a sender)
 Node j sends a notice message about the node j's states of node i, and evaluate the value of $T_{transfer(a,i,j)}$ by the equation 3 & 4 which is the transferring cost of the transaction. And judge if it is worthy of the transferred.
 IF (exists a transaction is worthy of the transferred)
 Node j sends a message to indicate there is an available transaction to be transfers the most cost-efficient transaction and corresponding logs. After that the algorithm terminates.

ELSE

Send a message indicating there is no available transaction can be transferred to node i.

ELSE

Remove the node i from the current queue to the queue's head of R_{queue} , and send the current states of node j and i.

5.3.3 Balance State

A node i is considered balanced if the Mean Deviation of MRI is less than the specified threshold, the LBF will reduce the load transferring factors.

$$LBF_{queue(i,j)} = LBF_{queue(i,j)} \times \mu \quad (6)$$

Where $0 < \mu < 1$. μ is called the LBF Regression Factor which regulates the load transferring factors regression process. After reducing LBF, if a LBF becomes sufficiently small (less than 0.005), it is reset to 0.

6 Performance Evaluation

In this section, we show the value of admission control policy by overload detection comparing the performance achievable through workload admission control policy. The transaction inter-arrival rate, which is drawn from an exponential distribution, is varied from 50 transactions per second up to 300 transactions per second in increments of 50, which represents light-to-medium loaded system. Each simulation was run three times, each time with a different seed, for 20000 ms. the results depicted is the average over the three runs. The settings for the user transaction workload are given in table 1. A user transaction consists of operations on both local data object and global data object to 1000 microseconds. At each node, the transaction workload consists of many periodic transactions and the average arrival rate and throughput shown in the simulation result.

Table 1. Baseline Workload Parameter

Parameter	Meaning	Value
CPUTime	CPU time per page access	2.5 ms
DBsize	Database size in pages	1,000
ArrivalRate	Transaction arrival rate	5-100 TPS
CTComp Time	Mean Compensating Task Time	10 ms
CTStdDev	St. Dev. of CT Time	0.5 T CompTime
SlackFactor	Slack Factor	2
RegFactor	Regression Factor	0 to .005
TaskSchd	Task scheduling protocol	EDF
CTSched	CT scheduling protocol	FF, LF, LMF
Thrsh	CT computation Threshold	0.125
CCntrl	Concurrency Control Protocol	OCC-BC

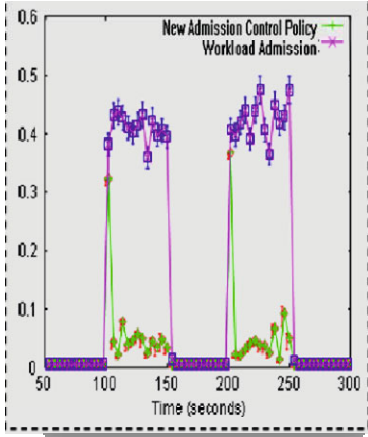


Fig. 2. Mean Deviation throughput

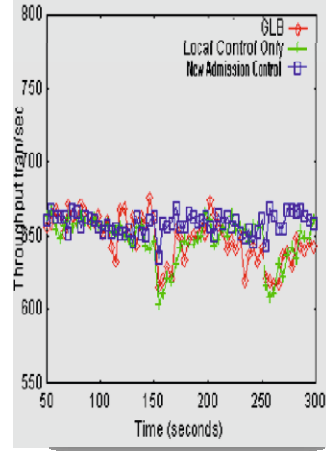


Fig. 3. Global Load Balancer with Throughput

As shown in Fig. 2 the system running best effort algorithm keeps unbalanced throughout the workload burst periods; with admission control policy the system workload become balanced (mean deviation of *MRI* becomes less than 0:1) within 5 seconds. The miss ratios at overloaded nodes are shown in Fig. 3 As we can see, for the best-effort algorithm, QoS requirements are violated and the miss ratio of class *i* transactions remains over 90%.

7 Related Work

This work differs from previous research in that our transaction model incorporates not only primary tasks, with unknown WCET, but also compensating tasks. The new admission control mechanism used admits transaction into the system with the absolute guarantee that either the primary task will successfully commit or the compensating task safely terminate. Distributed Real Time Database System (DRTDBS) have drawn research attention in recent years [14, 15]. Instead of providing strong logical consistency, DRTDBS focus on data freshness and timeliness of transaction. However most previous DRTDBS work targeted small scale system, but we extended it to large scale system in wide area network environment.

A New Admission Control Policy (NACP) and feedback mechanism could employ in variety of DRTDBS component: Transaction scheduling [15], Memory Allocation for Query Management [16], Concurrency Admission Control Management in ACCORD [8], and an Efficient Call Admission Control for Hard Real Time Communication in Differentiated Service Network [17]. The main idea of the admission control policy is to associate an local update to each submitted transaction in order to favors, when the system is overloaded, the executions of the most important transactions according to the application-transactions set. In overload conditions, each sub-transaction of the global

transaction is executed on a site that has the lowest workload among those sites that have executed the data items needed by the sub transaction.

8 Conclusion

Most previous DRTDBS studies have assumed that the only possible outcome of a transaction execution is either the commitment or the abortion of transaction. In many systems a third outcome of an outright rejection may be desirable. A process control application the outright rejection of a transaction may be safer than attempting to execute that transaction only to miss its deadline. Our system allows the system to reject a transaction because the admitted transaction holds some resources, thus making it possible utilization by other transaction, to be taken in timely fashion. Also this flexibility allows the system to relate its resources in the most profitable way, by only admitting high value transaction when the system is overloaded while being less choosy when the system is under loaded.

Our current research efforts focus on evaluating the performance of pessimistic as well as speculative CACM techniques. Moreover, our work to date has concentrated on uniprocessor systems. We are currently investigating the extension of our admission control and scheduling protocols to multiprocessor systems. A number of challenging questions arise. How are transactions, both their primary tasks and compensating tasks allocated to processors? What type of CPU scheduling discipline should be used? How valuable is the use of the WACM in a multiprocessor system? How the concurrency will be maintained in CACM?.

References

- [1] Aldarmi, S.A.: Real Time Database System, Concept and design. Department of computer science, University of York (April 1998)
- [2] Kim, Y., Son, S.: Supporting predictability in real time database system. In: Proc. 2nd IEEE Real Time Technology and Application Symposium (RTAS 1996), Boston, pp. 38–48 (1996)
- [3] Abbott, R., Garcia-Molina, H.: Scheduling real time transaction: A performance evaluation. In: Proceeding of the 14th International Conference on very large Data Bases, Los Angeles, CA, pp. 1–12 (1988)
- [4] Kang, W., Son, S.H., Stankovic, J.A.: Managing deadline miss ratio and sensor data freshness in real time databases. *IEEE Transactions on Knowledge and Data Engineering* (October 2004)
- [5] Stankovic, J.A., He, T., Abdelzaher, T., Marley, M., Tao, G., Son, S., Lu, C.: Feedback Control Scheduling in Distributed Real Time Systems Symposium (RTSS 2001), Washington, DC, USA, p. 59 (2001)
- [6] Lam, K.W., Lee, V.C.S., Hung, S.L.: Transaction scheduling in distributed real time system. *Int. J. Time – Crit. Comput. Syst.* 19, 169–193 (2000)
- [7] Lee, V.C.S., Lam, K.-W., Hung, S.L.: Concurrency control for mixed transactions in real-time data-bases. *IEEE Trans. Comput.* 51(7), 821–834 (2002)
- [8] Nagy, S., Bestavros, A.: Concurrency Admission Control Management in ACCORD. Ph.D Thesis at Boston University (1997)

- [9] Bestavros, A., Nagy, S.: Value-congnizant admission control for rtdb systems. In: RTSS 1996 the 17th Real Time System Symposium, Washington DC (December 1996)
- [10] Saab Systems Pty Ltd. “Ship Control System” in Saab System Website, <http://www.saabsystems.com.au>
- [11] Shanker, U.: Some performance issues in Distributed Real Time Database Systems. PhD Thesis, Department of Electronics & Computer Engineering, IIT Roorkee (December 2005)
- [12] Chetto, H., Chetto, M.: Some results of the earliest deadline scheduling algorithm. IEEE Transaction on Software Engineering (October 1989)
- [13] Liu, C.L., Layland, J.: Scheduling algorithms for multiprogramming in hard real time environments. Journal of the Association of Computing Machinery (January 1973)
- [14] Kang, W., Son, S.H., Stonkovic, J.A., Amirijo, M.: I/O aware deadline miss ratio management in real time embedded database. In: 28th IEEE Real Time System Symposium (RTSS) (December 2007)
- [15] Lee, V.C.S., Lam, K.W., Hang, S.L.: Transaction Scheduling in Distributed Real Time System. *Intr. J. Time Crit. Computer System* (2000)
- [16] Pang, H., Carey, M.J., Livny, M.: Managing memory for real time queries. In: Proceedings of the 1994 ACM SIGMOD Conference on Management of Data, pp. 221–232 (1994)
- [17] Baronia, P., Sahoo, A.: An efficient Call Admission Control for Hard Real Time Communication in Differentiated Services Network. *Proc. IEEE* (2003)

Wavelet Transform Based Image Registration and Image Fusion

Manjusha Deshmukh¹ and Sonal Gahankari²

¹ Asst. Prof., Electronics and Telecommunication, Saraswati college of Engg.,
Navi Mumbai, India
manju0810@yahoo.com

² Lecturer, Electronics and Telecommunication, Saraswati college of Engg.,
Navi Mumbai, India
sonalgahankari@rediffmail.com

Abstract. Image Registration is a fundamental task in image processing used to match two or more pictures taken, for example, at different times, from different sensors, or from different viewpoints. Image registration is particularly difficult when images are obtained through different sensor, (multi-modal registration). Mutual Information can be used for multimodal image registration. But this method has its own limitations one of it is of speed, method is very slow hence when time is an important constraint one cannot use this method. In this paper an attempt has been made to overcome this limitation.

Keywords: Image Registration, Image Fusion, Wavelet Transform, Mutual Information.

1 Introduction

Image registration is establishment of correspondence between images of the same scene. Many image processing applications like remote sensing for change detection, estimation of wind speed and direction for weather forecasting, fusion of medical images like PET-MRI, CT-PET etc need image registration.

A comprehensive survey of image registration methods is presented by Brown, Barbara Zitova and Jan Flusser [1, 2, and 3]. Subunku proposed various entropy based algorithms for multimodal image registration [4]. S. Chaudhari and U. Bhosale proposed methods for multispectral panoramic mosaicing and fast method for image mosaicing using geometric hashing [5,6]. Flusser classified feature based methods using special relations, methods using invariant descriptors, relaxation method and pyramid and wavelets [7, 8].

Image registration is classified into two types that are registration of images that are in same spectral band and registration of images that are in different spectral band. T. Sao proposed mutual information based method for image registration [9]. Cahill, N D Williams, C.M. Shoupu propose an approach to incorporate spatial information into the estimate of entropy to improve multimodal image registration [10]. A. survey

of medical image registration based on mutual information is presented by J. P. W. Pluim [11]. Xiaoxiang Wang and Jie Tian in their paper proposed a mutual information based registration method using gradient information [12]. Frederik Maes and Andre Collignon apply mutual information to measure the statistical dependence between the image intensities of corresponding voxels in both images [13].

2 Wavelet Domain Image Registration

Registration of image using wavelets is being discussed in this section.

Using discrete wavelet transform (DWT), a function $f(t)$ can be represented by

$$f(t) = \sum_{j,k} a_{j,k} \psi_{jk}(t) \tag{1}$$

Where $a_{j,k}$ are wavelet coefficients, $\psi_{j,k}(t)$ are basis function, j is scale, k is translation of mother wavelet $\Psi(t)$. Two dimensional DWT can be obtained by applying DWT across rows and columns of an image. The two dimensional DWT of image $f(x,y)$ is

$$f(x, y) = \sum_{j,k} C_{J_0}(k, l) \phi_{j,k,l}(x, y) + \sum_{S=H,V,D} \sum_{J=J_0}^{\infty} \sum_{k,l} D_j^S[k, l] \psi_{j,k,l}^S(x, y) \tag{2}$$

Where C_{J_0} is approximation coefficient, $\phi_{j,k,l}(x,y)$ is scaling function, D_j^S is set of detail coefficients and $\psi_{j,k,l}^S$ is set of wavelet function.

The DWT coefficients are computed by using a series of low pass filter $h[k]$, high pass filters $g[k]$ and down samplers across both rows and columns. The results are the wavelet coefficient the next scale. The filter bank approach to calculate two dimensional dyadic DWT is shown in figure 3 and dyadic representation of the DWT is shown in figure 4. The wavelet coefficients are of smaller spatial resolution as they go from finer scale to coarser scale. The coefficients are called the approximation (A), horizontal detail (H), vertical detail (V) and diagonal detail (D) coefficient.

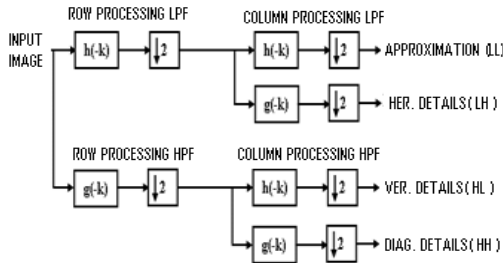


Fig. 1. Two-dimensional orthogonal wavelet decomposition

2.1 Cross Correlation as Similarity Measure

Cross correlation is a similarity measure or match metric. For template T and image I , where T is small compared to I , the two dimensional normalized cross-correlation function measures the similarity for each translation.

$$C(u, v) = \frac{\sum_x \sum_y T(x, y) I(x - u, y - v)}{\sqrt{[\sum_x \sum_y T(x, y)] [\sum_x \sum_y I(x - u, y - v)]}} \quad (3)$$

If template matches the image, then cross correlation will have it's peak.

2.2 Mutual Information as Similarity Measure

MI is an entropy-based concept and denotes the amount of information that one variable can offer to the other. Mutual Information criteria presented here states that, mutual Information of image intensity values of corresponding voxel pairs is maximum if images are geometrically aligned. Let A and B represent random variables and $P_A(a)$ and $P_B(b)$ represents its marginal probability distributions. Let $P_{AB}(a, b)$ represents joint probability distribution then A & B are independent if $P_{AB}(a, b) = P_A(a) * P_B(b)$ Mutual Information $I(A, B)$ is given by

$$I(A, B) = \sum_{a,b} P_{AB}(a, b) \log [P_{AB}(a, b) / \{P_A(a).P_B(b)\}] \quad (4)$$

Mutual Information is related to entropy by following equations.

$$I(A, B) = H(A) + H(B) - H(A, B) \quad (5)$$

3 Application of Image Registration for Image Mosaicing

We collected series of images using digital camera on a leveled tripod in front of Hiranandani complex, Kharghar, Navi Mumbai.

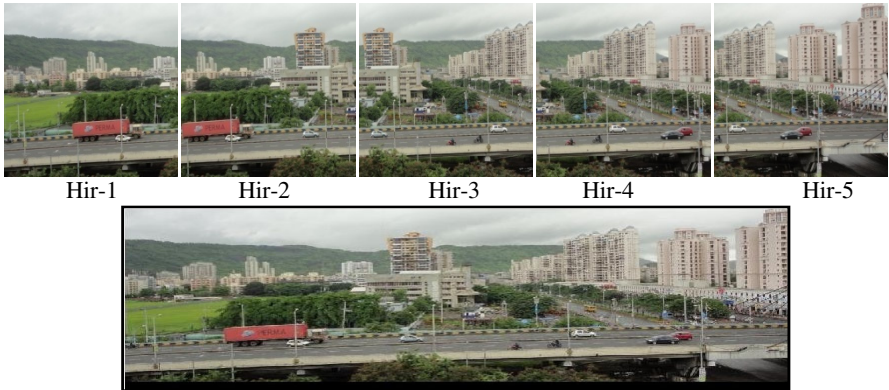


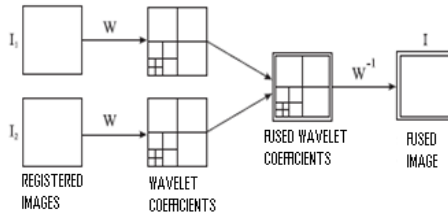
Fig. 2. Mosaic of images from Hir1 to Hir-5

Table 1. Location of Maximum Match

S.N.	Image combination	Wavelet Method Using MI	Wavelet Method Using Correlation
1	Hir-1-Hir-2	146	149
2	Hir-2-Hir-3	237	240
3	Hir-3-Hir-4	127	129
4	Hir-4-Hir-5	127	128

4 Wavelet Based Image Fusion

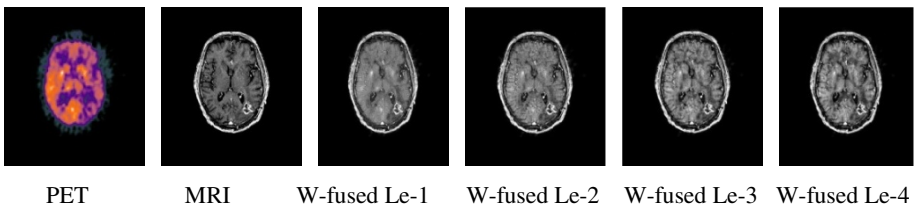
Image fusion is useful technique for merging similar sensor and multi-sensor images to enhance the information.

**Fig. 3.** Wavelet multi-dimensional fusion

4.1 Wavelet Based Algorithm

- Apply wavelet transformation separately to each source image to establish various images of
- wavelet tower shaped transformation.
- Fuse images at each transformation level.
- Apply inverse Wavelet transform on fused wavelet pyramid.

In wavelet transformation due to sampling, the image size is halved in both spatial directions at each level of decomposition process thus leading to a multi-resolution signal representation. The most important step for fusion is the formation of fusion

**Fig. 4.** Sample images and Wavelet based fused images at different levels

pyramid. We used mutual information based method for registering source images. In the process of fusion, we fused images at four different levels. In the next section, we make a quantitative evaluation of fusion at different levels.

Wavelet based fusion can deal with images of different spectral and spatial resolutions. However, this method cannot handle cases where data is scattered or when input images differ greatly in either their spectral or spatial resolution.

5 Conclusion

From experimental results it is observed that Mutual information method yields a more accurate registration. But this method has its own limitations. When images are of low resolution, when images contain little information, or when the region of overlap is small then mutual information result in mis-registration. It has one more limitation of speed, when time is an important constraint one cannot use this method. Although it has some limitations entropy and mutual information are best approaches for multimodal image registration. It is observed that combinational approach of wavelet and mutual information gives better results as compared to wavelet - correlation combination. Even wavelet mutual information combination can be used in case of multimodal image registration. Wavelet based fusion can deal with images of different spectral and spatial resolutions.

References

- [1] Brown Gottesfeld, L.: Survey of Image Registration techniques. *ACM Computing Surveys* 24(4), 325–376 (1992)
- [2] Zitova, B., Flusser, J.: Image Registration Methods: A survey. *Image and Vision Computing* 21, 977–1000 (2003)
- [3] Antoine Maintz, J.B., Vierger, M.A.: A Survey of Medical Image Registration. *Medical Image Analysis* 2(1), 1–37 (1998)
- [4] Sabuncu, M.R.: Spatial Information in Entropy – Based Image Registration. In: Gee, J.C., Maintz, J.B.A., Vannier, M.W. (eds.) *WBIR 2003*. LNCS, vol. 2717, pp. 132–141. Springer, Heidelberg (2003)
- [5] Chaudhari, S., Bhosale, U., Dutta Roy, S.: Multispectral Panoramic Mosaicing. In: International Conference on Advances in Pattern Recognition (ICPR), pp. 188–191. Indian Statistical Institute, Kolkatta (2003)
- [6] Bhosale, U., Chaudhari, S., Dutta Roy, S.: A Fast Method For Image Mosaicing Using Geometric Hashing. *IETE Journal of Research: Special Issue on Multimodal Media Processing*, 317–324 (May-August 2002)
- [7] Flusser, J., Suk, T.: Degraded Image Analysis: an Invariant approach. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 20, 590–603 (1998)
- [8] Flusser, J., Suk, T.: A Moment Based Approach to Registration of Images with Affine Geometric Distortion. *IEEE Transactions on Geoscience and remote Sensing* 32, 382–387 (1994)
- [9] Tsao, J.: Interpolation Artifacts in Multimodality Image Registration Based on Maximization of Mutual Information. *IEEE Medical imaging* 22(7), 854–864 (2003)

- [10] Cahill, Williams, N.D., Shoupu, C.M.: Biomedical Imaging, Nano to Macro. In: 3rd IEEE International Symposium, April 6-9, pp. 832–835 (2006)
- [11] Pluim, J.P.W., Maintz, J.B.A.: Mutual Information Based registration of medical images, survey. *IEEE Medical Imaging*
- [12] Wang, X., Tian, J.: Image Registration based on Maximization of Gradient Code Mutual Information. *Image Anal. Stereol.* 24, 1–7 (2005)
- [13] Maes, F., Collignon, A.: Multimodality Image Rregistration by Maximization of Mutual Information. *IEEE Transactions on Medical Imaging* 16(2) (April 1997)

60 GHz Radio Channel Characteristics in an Indoor Environment for Home Entertainment Networks

T. Rama Rao¹, S. Ramesh², and D. Murugesan²

¹Dept. of Telecommunication Engineering,
SRM University, Kattankulathur – 603203, TN, India
ramarao@ieee.org

²Dept. of ECE, Valliammai Engineering College
Kattankulathur – 603203, Tamil Nadu, India
rameshsvk@gmail.com, mdmurugesan@gmail.com

Abstract. 60 GHz radio channel propagation characteristics in a typical indoor environment are addressed in this paper using a simple deterministic 2-Dimensional model utilizing ray-tracing technique based on geometrical optics (GO) and image principle. Values of the received power, rms delay spread (DS) and power delay profile (PDP) for horn, omni-directional and isotropic antennas are presented.

Keywords: Radio channel characterization, Indoor Environments, Millimeter Waves, Ray-Tracing Technique, Antennas, Received Power, Delay Spread, Wireless Networks.

1 Introduction

In today's world, 'wireless' is the new electricity and wireless technologies are allowing us to enter an age where information becomes contextualized geographically as well as personally. The last few years have seen a growth in the demand for wireless broadband access. The reason for this growth can be seen as the emergence of multimedia applications, demands for ubiquitous high-speed Internet connectivity, the massive growth in the wireless and mobile communications and the deregulation in the telecommunications industry. Recently, applications of millimeter (mm) waves for high-speed broadband WLAN (Wireless Local Area Network) WPANs (Wireless Personal Area Network) communication systems such as multimedia equipment, home appliances, video signal transmissions, and personal computers in indoor environment are increasingly gaining importance due to the spectrum availability and wider bandwidth requirements [1, 2]. Within these mm wave bands, various kinds of wireless data transmissions are expected to develop with/without license.

Propagation characteristics of mm waves, especially at 60 GHz are very different from the ones at lower frequencies. Strong attenuation over the free space due to the smaller wavelengths, oxygen absorption and severe attenuation by walls allow frequency reuse and user privacy [3, 4, 5, 6]. This makes 60 GHz an attractive proposition for high-speed indoor WLAN & WPANs [7, 8]. It is important to know the

propagation characteristics of the wireless radio channel at 60 GHz for designing efficient WLAN/WPAN systems. Next generation WLAN/WPAN's at 60 GHz is now a subject of intensive research topic.

In the present work, in order to describe the radio signal propagation at 60 GHz in a typical indoor environment, we present a deterministic approach of a simple 2-Dimensional model utilizing ray-tracing technique based on classical geometrical optics (GO) and image method [9, 10, 11] to account for the direct ray and reflected rays from wall, floor and ceiling, respectively. The key measurable parameters characterizing the wideband indoor radio channel at 60 GHz are the received power, rms delay spread and power delay profile. These are evaluated for horn, omni-directional and isotropic antennas utilizing Matlab simulations. This paper is organized as follows; Section 2 deals with 60 GHz propagation channel model describing the geometry of the environment under consideration, the proposed ray-tracing model and the simulation procedure. Section 3 deals with results obtained in our preset work and discussions. Finally, Section 4 gives conclusions.

2 Propagation Channel Modeling

2.1 Description of the Indoor Environment

The considered simulation environment is a long corridor with dimensions 40.0 x 4.0 x 3.0 m³. The left and right wall surfaces are made of brick and plasterboard (relative permittivity $\epsilon_r = 5.0$ [12, 13]). In order to simplify the simulation procedure we assume the surface as a uniform wall made of brick and plasterboard. The floor is concrete covered with marble ($\epsilon_r = 4.0$ [12, 13]) and furred ceiling is made of aluminum ($\epsilon_r = 1.0$ [12, 13]). The beginning and the end of the corridor are open areas and are not taken into account in the simulations.

2.2 Modeling

The radio channel propagation modeling at mm wave frequencies can be realized based on ray-tracing theory. The ray-tracing method is among the available methods for the relatively accurate estimation of field strengths to deal with the type of complex layout that is often found in indoor environments [14] and allows fast computation of single & double reflection processes. In 60 GHz region the diffraction phenomenon can be neglected and the sum of the direct ray and the reflected rays are enough to describe the behavior of the propagation channel with great accuracy [1, 15]. In this present work, the proposed simple 2D model is a general case of the two-ray model [9, 16]. The reflected components may exhibit single or double reflection from a plane surface. If we know the geometry of the environment where the signal propagates and the surface reflection coefficients, one may calculate the propagation losses [11]. The total received power (R_R) of the multi-rays are calculated by the summation of 'X' single reflected and 'W' double reflected rays given by

$$R_R = T_R \left(\frac{\lambda}{4\pi} \right)^2 a_i a_r \left| \frac{e^{-jkd_1}}{d_1} + \sum_{i=1}^X R(\theta_0) \frac{e^{-jkd_2}}{d_2} + \sum_{j=1}^W R(\theta_1) R(\theta_2) \frac{e^{-jkd_3}}{d_3} \right|^2 \quad (1)$$

where λ is the wave length; k is the wave number; d_1 is the distance of the direct path; d_2 is the distance of the single reflected path; d_3 is the distance of the double reflected path; a_t , a_r are the antenna functions; $R(\theta_0)$ is the reflection coefficient of the single reflected ray on the reflecting surface; $R(\theta_1)$, $R(\theta_2)$ are the reflection coefficient of the double reflected rays on respective reflecting surfaces; and T_R is the transmitted power. For isotropic antennas ($a_t = a_r = 1$) the total received power (R_R) is

$$R_R = T_R \left(\frac{\lambda}{4\pi} \right)^2 \left| \frac{e^{-jkd_1}}{d_1} + \sum_{i=1}^X R(\theta_0) \frac{e^{-jkd_2}}{d_2} + \sum_{j=1}^W R(\theta_1)R(\theta_2) \frac{e^{-jkd_3}}{d_3} \right|^2 \quad (2)$$

2.3 Simulation Procedure

In our simulations, we considered the direct ray, floor reflected ray, wall reflected rays, wall-floor reflected rays, ceiling reflected ray and ceiling-floor reflected ray, and wall-ceiling reflected ray. The simulations are conducted with Matlab script. The initial transmitter position is at the beginning of the corridor and the receiver is moving away from T_x with 1 m initial separation and we collect a signal sample as a function of distance every 0.00125 m ($\lambda/4$). During the entire simulation procedure vertical polarization is assumed. Hence, for the rays reflected from vertical walls we use the perpendicular reflection coefficient, whereas for the rays from floor and ceiling surfaces we use the parallel reflection coefficient. Both reflection coefficients are calculated using the equations described in [16]. To examine how the antenna radiation patterns affect the signal propagation in the indoor environment, we assumed three different transmission systems with different antenna characteristics and transmitted power. The systems are 1) Horn antenna with 10 dBm output power, $a_t = a_r = 20.8$ dBi, 2) Omni directional antennas with 20 dBm output power, $a_t = a_r = 8.5$ dBi, and 3) Isotropic antennas with 20 dBm output power. Further in our simulations, we are not considered the factors such as the diffraction loss, atmospheric propagation losses, third/fourth order reflections and the non-uniformities of the surface materials, as these are almost negligible and not contribute to the total received power.

3 Results and Discussions

3.1 Received Power

The received power provides an indication of the received signal strength with respect to receiver position. It can be used by system designers to gauge what transmitted power is required in order to achieve the desired performance parameters. The received power is vital to determining cell sizes and co-channel interference effects; it also provides a preliminary measure of maximum raw data rates prior to modulation. The antenna radiation patterns, have a significant impact on the received power, the simulated results differ significantly for isotropic, omni-directional and horn antennas. The uniform radiation characteristics of isotropic and omni-directional antennas, allow signals to be transmitted and received equally well from all directions. This results in all the multipath components to reach the receiver, making the multipath effect much more severe. Whereas for the horn antennas the narrow antenna bandwidths suppress more of the multipath rays, hence in our simulation we discovered that using horn

antennas only first order reflected rays had a significant contribution to the received power, with second order reflections being suppressed.

Table 1. Comparison of total received power values for different antenna configurations

	Received Power for Horn Antenna (dBm)	Received Power for Omni-directional Antenna (dBm)	Received Power for Isotropic Antenna (dBm)
Min	-79.73	-94.33	-109.33
Max	-13.44	-28.04	-43.04
Mean	-36.47	-51.07	-66.07
Median	-36.61	-51.22	-66.22
SD	7.26	7.26	7.26

Table 1 summarizes the received power statistics for isotropic, omni and horn antennas. It is observed that the mean received power is -66.07, -51.07 and -36.47 dBm for isotropic, omni-directional and horn antenna respectively. An increase of 29.7 dB in received power between horn and isotropic antennas can be attributed to the ability of the horn antenna to minimize its response to unwanted signals not in the favored direction of the antenna. Inspecting the above table, we observe that the received power is far greater when a horn antenna is employed in comparison to the isotropic and omni-directional antennas.

3.2 Power Delay Profile (PDP)

The wideband channel is characterized by the time or space-variant channel impulse response (CIR) [16]. In an indoor propagation environment where the typical time-varying factors are human movement, it can be assumed that the channel is quasi-stationary. The phase variations are assumed to be mutually independent random variables, which have a uniform distribution over $[-\pi, \pi]$. Thus we considered only the amplitude and delay components in our simulations. The most significant parameter derived from the wideband channel model is the power delay profile [16] which is a representation of the individual contributions of each multipath component with respect to excess time delay and provides a visualization of vital delay statistics. The PDP can be expressed as

$$P_D(\tau) = \sum_{i=1}^N R_R(d) \delta(\tau - \tau_i) \quad (3)$$

where $R_R(d)$ is the received signal power of the i^{th} multipath component, τ_i is the excess delay which is the relative delay of the i^{th} component as compared to the first arriving component and N is the total number of equally spaced multipath components. During our quantization process we discretize the delay axis to have a time resolution of 1 ns [16]. In this process we assigned each multipath echo to the nearest value of delay equal to a multiple of 1 ns. The average received power in each bin is normalized to the direct ray component. In the simulation process we obtain a PDP for each T_x - R_x distance traversed, resulting in a set of PDPs. Each PDP is the individual power contributions of each multipath component at that particular distance. To gain an

overall characterization of the channel we take the averages of these PDP's. The PDPs of horn, isotropic and omni antennas followed respective trends with large delays since the entire multipath components contribute to the total received signal power.

3.3 rms Delay Spread (DS)

The important parameter derived from the PDP is the rms (root mean square) delay spread (DS), is defined as square root of the second central moment of the average PDP [16]. At mm wave frequencies the channel dispersion is smaller when compared to values encountered at lower frequencies because echo paths are shorter on average. Simultaneous measurements at 5 and 60 GHz indicate a difference of a factor 1.5 – 2 [17]. The rms DS of the channel may range from a few to 100 ns if linear polarisation is used. It is expected to be highest if omni-directional antennas are used in large reflective indoor environments [18]. When, instead, high gain antennas are used, the rms DS may be limited to a few ns only [18], but this is only the case when the antennas are exactly pointed towards each other. In our simulations we discovered that when isotropic antennas were employed, an average rms DS of approximately 2.2 ns was obtained, this increase in the rms DS can be attributed to the fact that for isotropic antennas all the multipath components contribute to the total received power. Obtained rms DS of 1.59 ns with horn antennas due to its narrow antenna beamwidths suppresses some of the multipath components, reducing the rms DS, and hence increasing the channel performance. With Omni directional antennas we obtained rms DS of 1.21 ns. Comparing our results with similar research works [11, 19, 20, 21], we find that our simulated results are in agreement. The close correlation with our simulated observations is due to the similarity in configuration methodologies.

4 Conclusions

With the view to analyze 60 GHz wireless indoor scenario for WLAN & WPAN applications, radio channel characterization has been made in an indoor environment using a simple deterministic 2D model utilizing ray-tracing technique based on geometrical optics and image principle. From our results we observed that the antenna radiation patterns have a significant effect on received power and the use of directive horn antennas greatly improved the power performance in comparison to its isotropic/omni-directional counterpart. In our simulations we observed that the PDP is dominated by 1st order reflected components, whilst 2nd order reflected components have a relatively insignificant contribution and the PDP follows an exponentially decaying relationship with respect to excess delay. Further, we observed that the rms DS values decreased with the use of high gain directive antennas, due to suppression of unwanted signal components by the narrow antenna beam width. It is believed that the huge demand for bandwidth and higher data rate services will make the 60 GHz channel an inevitable eventuality. However due to the complex nature of mm wave propagation, there are still a lot of unknowns that need to be quantified before a working standard is achieved. With this in mind the work presented in this paper serves as a contribution to the deployment of mm wave based WLAN & WPANs.

References

1. Correia, L.M., Prasad, R.: An overview of wireless broadband communications. *IEEE Commun. Mag.*, 28–33 (January 1997)
2. Smulders, P.F.M.: Exploiting the 60 GHz band for local wireless multimedia access: prospects and future directions. *IEEE Communications Magazine* 40(1) (January 2002)
3. Andrisano, O., Chiani, M., Tralli, V.: Millimeter wave short range communications for advanced transport telematics. *European Transactions on Telecommunications* (July-August 1993)
4. Smulders, P.F.M., Wagemans, A.G.: Wideband Indoor Radio Propagation Measurements at 58 GHz. *Electron. Lett.* 28(13), 1270–1272 (1992)
5. Dardari, D., Minelli, L., Tralli, V., Andrisano, O.: Wideband Indoor Communication Channels at 60 GHz. In: *Proceedings of PIMRC 1996, Taiwan, October 15-18 (1996)*
6. Prasad, R.: Overview of Wireless Personal Communications: Microwave Perspective. *IEEE Communications Mag.*, 104–108 (April 1997)
7. Xiao, S.-Q., Zhang, M.-T.Z.Y.(eds.): *Millimeter Wave Technology for Wireless LAN, PAN and MAN*. Auerbach Publications (2008)
8. Yong, S.K., Chong, C.-C.: An Overview of Multi gigabit Wireless through Millimeter Wave Technology: Potentials and Technical Challenges. *EURASIP Journal on Wireless Communications and Networking*, ArticleID 78907 (2007)
9. Betroni, H.L.: *Radio Propagation for Modern Wireless Systems*. Prentice Hall, Englewood Cliffs (2000)
10. Hammoudeh, A.M., Graham, A.: Millimetric wavelengths radiowave propagation for LoS Microcellular mobile communications. *IEEE Trans.* 44(3) (August 1995)
11. Nektarios, M., Philip, C.: Propagation Modeling at 60 GHz for Indoor Wireless LAN Applications. *IST Mobile & Wireless Telecommunications Summit* (2002)
12. Sato, K., et al.: Measurements of the complex refractive index of concrete at 57.5 GHz. *IEEE Trans. Antennas Propa.* 44(1), 35–39 (1996)
13. Sato, K.: Measurements of reflection and transmission characteristics of interior structures of office building in the 60-GHz band. *IEEE Trans. Ant. Prop.* 45 (December 1997)
14. Imai, T., Fujii, T.: Indoor micro cell area prediction system using ray-tracing for mobile communication systems. In: *Proc. PIMRC 1996, vol. 1, pp. 24–28 (1996)*
15. Hübner, J., Zeisberg, S., Koora, S., Finger, A.: Simple Channel model for 60 GHz indoor wireless LAN Design Based on Complex Wideband Measurements. In: *IEEE 47th Vehicular Technology Conference*, pp. 1004–1008 (1997)
16. Rappaport, T.S.: *Wireless Communications*. Prentice Hall, Englewood Cliffs (2000)
17. Plattner, A., Prediger, N., Herzig, W.: Indoor and outdoor propagation measurements at 5 and 60 GHz for radio LAN application. *IEEE MTT-S Digest.*, 853–856 (1993)
18. Smulders, P.F.M.: *Broadband wireless LANs: a feasibility study*. Ph.D. Thesis, Eindhoven University of Technology, The Netherlands (1995); ISBN 90-386-0100-X
19. Dardari, D., Minelli, L., Tralli, V., Andrisano, O.: Wideband indoor, communication channels at 60 GHz. In: *PIMRC 1996, vol. 3, pp. 791–794 (October 15-18, 1996)*
20. Ghobadi, C., Shepherd, P.R., Pennock, S.R.: 2D ray-tracing model for indoor radio propagation at millimetre frequencies, and the study of diversity techniques. In: *IEE Proc. Microw. Ant. Prop.*, vol. 145(4), pp. 349–353 (August 1998)
21. Chiu, C.C., Wang, C.P.: A Comparison of Wideband Communication Characteristics for Various Corridors at 57.5 GHz. *Wireless Personal Communications* (12), 71–81 (2000)

Improved Back Propagation Algorithm to Avoid Local Minima in Multiplicative Neuron Model

Kavita Burse¹, Manish Manoria², and Vishnu Pratap Singh Kirar¹

¹ Department of Electronics and Communication, Truba Institute of Engineering and Information Technology, Bhopal, India
kavitaburse14@gmail.com, vishnupskirar@live.com

² Department of of Computer Science and Engineering, Truba Institute of Engineering and Information Technology, Bhopal, India
manishmanoria@rediffmail.com

Abstract. The back propagation algorithm calculates the weight changes of artificial neural networks, and a common approach is to use a training algorithm consisting of a learning rate and a momentum factor. The major drawbacks of above learning algorithm are the problems of local minima and slow convergence speeds. The addition of an extra term, called a proportional factor reduces the convergence of the back propagation algorithm. We have applied the three term back propagation to multiplicative neural network learning. The algorithm is tested on XOR and parity problem and compared with the standard back propagation training algorithm.

Keywords: Three term back propagation, multiplicative neural network, proportional factor, local minima.

1 Introduction

Artificial Neural Network (ANN) consists of a number of interconnected processors known as neurons, which are identical to the biological neural cells of the human brain. Neural network is defined by its architecture, neuron model and the learning algorithm. Architecture refers to a set of neurons and the weighted links connecting the layers of neurons. Neuron model refers to information processing unit of the neural network. The weights are adjusted during the training process. A learning algorithm is used to train the NN by modifying the weights in order to model a particular learning task correctly on the training examples. Learning is a fundamental and essential characteristic of ANN. ANN training usually updates the weights iteratively using the negative gradient of a Mean Squared Error (MSE) function. The error signal is then back propagated to the lower layers.

The back propagation (BP) algorithm was developed by Rumelhart, Hinton and Williams in 1986. Efficient learning by the BP algorithm is required for many practical applications. The BP algorithm calculates the weight changes using a two-term algorithm consisting of a learning rate and a momentum factor. The major drawbacks of the two-term BP learning algorithm are the problems of local minima

and slow convergence speeds. The addition of an extra term called a proportional factor (PF) to the two-term BP algorithm was proposed in 2003 by Zweiri and has outperformed standard two-term BP in terms of low complexity and computational cost [1]. BP is a method for calculating the first derivatives or gradient of the cost function required by some optimization methods. It is certainly not the only method for estimating the gradient. However, it is the most efficient [2]. The major limitations of this algorithm are the existence of temporary, local minima resulting from the saturation behavior of the activation function. A number of approaches have been implemented to avoid the local minima which are based on selection of dynamic variation of learning rate and momentum, selection of better activation function and better cost function.

In [3] the learning rate and momentum coefficient are adapted according to the coefficient of correlation between the downhill gradient and the previous weight update. In [4] modification is based on the solving of weight matrix for the output layer using theory of equations and least squares techniques. Drago *et al.* have proposed an adaptive momentum BP for fast minimum search [5]. A randomized BP algorithm is proposed by Chen *et al.* It is obtained by choosing a sequence of weighting vectors over the learning phase [6]. A new generalized BP algorithm is proposed in [7] to change the derivative of the activation function so as to magnify the backward propagated error signal, thus the convergence rate can be accelerated and the local minimum can be escaped. An adaptive BP algorithm is proposed in [8] which can update learning rate and inertia factor automatically based on dynamical training error rate of change. In [9] an improved BP is proposed where each training pattern has its own activation function of neurons in hidden layer to avoid local minima Wang *et al.* have proposed an individual inference adjusting learning rate technique to enhance the learning performance of the BP neural network [10]. In [11] Conjugate Gradient (CG) algorithm which is usually used for solving nonlinear functions is combined with the modified Back Propagation (BP) algorithm yielding a new fast training multilayer algorithm. The proposed algorithm improved the training efficiency of BP-NN algorithms by adaptively modifying the initial search direction.

This paper is organized as follows. In the next section we propose the learning rule with improved BP algorithm to avoid local minima in multiplicative neuron model. In section 3, we illustrate the basic results of our paper. Section 4 concludes the paper.

2 Learning with Improved BP Algorithm

The McCulloch-Pitts model initiated the use of summing units as the neuron model, while neglecting all possible nonlinear capabilities of the single neuron and the role of dendrites in information processing in the neural system. It is widely agreed that there is only minor correspondence between these neuron models and the behavior of real biological neurons. In particular, the interaction of synaptic inputs is known to be essentially nonlinear. In search for biologically closer models of neural interactions, neurobiologists have found that multiplicative-like operations play an important role in single neuron computations. For example, multiplication models nonlinearities of dendritic processing and shows how complex behavior can emerge in simple

networks. In recent years evidence has accumulated regarding specific neurons in the nervous system of several animals compute in a multiplicative manner. Multiplication increases the computational power and storage capacity of neural networks is well known from extensions of ANN where this operation occurs as higher order units [12].

The back-propagation learning algorithm with multiplicative neural networks (MNN) has to be more efficient as both single-units and also in networks. The goal of learning is to update the network weights iteratively to minimize globally the difference between the actual output vector of the network and the desired output vector. The rapid computation of such a global minimum is a rather difficult task since, in general, the number of network variables is large and the corresponding non convex multimodal objective function possesses multitudes of local minima and has broad flat regions adjoined with narrow steep ones. The architecture of the MNN is described as follows. The basic building block of the MNN is a single neuron or node as depicted in Fig. 1[13].

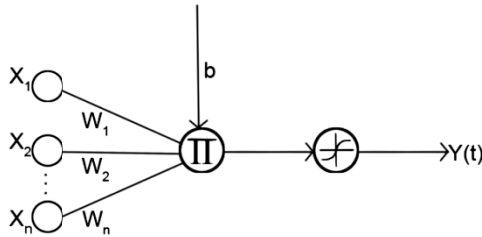


Fig. 1. Node structure of MNN

A node receives a number of real inputs x_1, x_2, \dots, x_n which are then multiplied by a set of weights w_1, w_2, \dots, w_n and bias terms b_1, b_2, \dots, b_n are added. The resultant values are multiplied to form a polynomial structure. This output of the node is further

subjected to a nonlinear function f defined as $f = \frac{1 - e^{-x}}{1 + e^{-x}}$.

In the MNN a number of nodes described above are arranged in layers. A multidimensional input is passed to each node of the first layer. The outputs of the first layer nodes then become inputs to the nodes in the second layer and so on. The output of the network is the output of the nodes of the final layer. Weighted connections exist from a node to every node in the succeeding node but no connections exist between nodes of the same layer. As the number of layers in the MNN is increased decision regions are formed which are considerably more complex and have highly nonlinear boundaries. Fig. 2 shows a general model of a feed forward MNN where the summation at each node is replaced by the product unit.

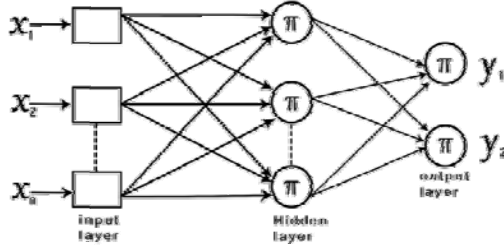


Fig. 2. Architecture of feed forward MNN

The MNN is trained using supervised learning where a set of input and the corresponding target vector is used to adjust the scalar parameters weight and bias. The network is trained incrementally so that the weights and biases are updated after each input is presented. The MNN is trained using supervised learning where a set of input and the corresponding target vector is used to adjust the scalar parameters weight and bias. Output of the node u before applying activation function is given by [14].

$$u = \prod_{i=1}^n (w_i x_i + b_i) \quad (1)$$

The bipolar sigmoidal activation function f is given by

$$y = f(u) = \frac{1 - e^{-u}}{1 + e^{-u}} \quad (2)$$

An error back propagation based learning rule is used for training. The MSE is given by

$$E = \frac{1}{2N} \sum_{p=1}^N (y^p - d^p)^2 \quad (3)$$

Where, p is the number of input patterns.

The weight update equation for single layer algorithm is given by

$$\begin{aligned} \Delta w_i &= -\eta \frac{\partial E}{\partial w_i} \\ &= -\frac{1}{2} \eta (y - d)(1 + y)(1 - y) \frac{u}{(w_i x_i + b_i)} x_i \end{aligned} \quad (4)$$

Where, η is the learning rate and d is the desired signal. If η is large, learning occurs quickly, but if it is too large it may lead to instability and errors may even increase.

The bias is updated as

$$\begin{aligned} \Delta b_i &= -\eta \frac{\partial E}{\partial b_i} \\ &= -\frac{1}{2} \eta (y - d)(1 + y)(1 - y) \frac{u}{(w_i x_i + b_i)} \end{aligned} \quad (5)$$

The standard BP algorithm calculates the new weights and biases as

$$w_i^{new} = w_i^{old} + \Delta w_i \quad (6)$$

$$b_i^{new} = b_i^{old} + \Delta b_i \quad (7)$$

The standard algorithm is further modified by adding the momentum term and proportional factor term. The momentum term is a fraction of the previous weight change. The momentum term prevents extreme changes in the gradient due to anomalies and suppresses oscillations due to variations in the slope of the error surface [15] and prevents the network to fall into shallow local minima. The convergence still remains relatively slow because of the saturation behavior of the activation function. In the saturation area of the output activation function, the corresponding gradient descent takes very small value leading to small changes in weight adjustments. The problem of slow convergence is solved by adding a term proportional to the difference between the output and the target. The improved BP weight update is calculated as

$$\Delta w_i^{improved} = \Delta w_i + \beta \Delta w_i^{old} + \gamma(y - d) \quad (8)$$

$$\Delta b_i^{improved} = \Delta b_i + \beta \Delta b_i^{old} + \gamma(y - d) \quad (9)$$

β is the proportional term

Δw_i^{old} is the previous weight change

γ is the proportional term

$(y - d)$ is the difference between the output and the target at each iteration

Δb_i^{old} is the previous bias change

The error function optimization depends on three independent quantities. The three term back propagation function as a PID controller used in control application [16].

3 Simulation and Results

We have tested the convergence of the three term BP algorithm for the MNN network on the XOR problem which is the most used nonlinear pattern classification problem as compared to other logic operations. The architecture of the network is a single layer MNN with 2 inputs, 3 hidden layer neurons and 1 output neuron. The convergence curve for the three term BP algorithm is compared with the standard BP algorithm in Fig. 3. The convergence of the improved BP algorithm with momentum and PF factor is five times faster as compared to the standard BP algorithm. Table 1 compares the testing performance for the XOR problem.

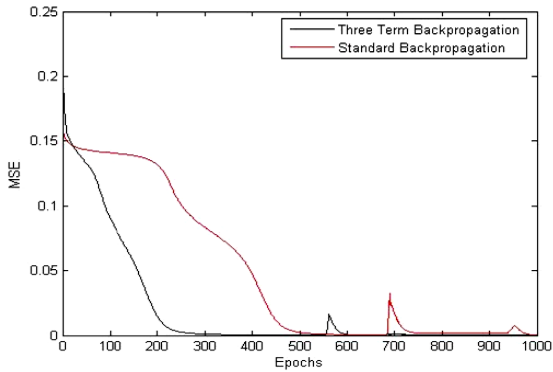


Fig. 3. Convergence curves for XOR problem

Table 1. Testing performance for XOR problem

Input	Target	Output with MNN trained with standard BP algorithm	Output with MNN trained with three term BP algorithm
0 0	0	0.0004	0.0001
0 1	1	0.9841	0.9942
1 0	1	0.9835	0.9997
1 1	0	0.0002	0.0001

Table 2. Testing performance for 3 bit parity problems

Input	Target	Output with MNN trained with standard BP algorithm	Output with MNN trained with three term BP algorithm
0 0 0	0	0.0312	0.0032
0 0 1	1	0.8921	0.9978
0 1 0	1	0.9876	0.9886
0 1 1	0	0.0214	0.0021
1 0 0	1	0.8953	0.9778
1 0 1	0	0.0021	0.0041
1 1 0	0	0.0032	0.0034
1 1 1	1	0.9873	0.9921

We have further tested the three term BP algorithm for the MNN network on the three bit parity problem which maps 3 bit binary numbers onto its parity. The parity bit output is 1 for odd number of 1 else it is 0. Table 2 compares the testing performance for parity problem.

4 Conclusion

In this paper we have proposed an improved BP algorithm to avoid local minima and for faster convergence of multiplicative neural network training algorithm. We have tested the algorithm for XOR and three bit parity problem and compared the result with standard BP multiplicative neural network algorithm. The addition of PF term helps in convergence of the algorithm five times faster.

References

1. Zweiri, Y.H., Whidborne, J.F., Althoefer, K., Seneviratne, L.D.: A three term back propagation algorithm. *Neurocomputing* 50, 305–318 (2003)
2. Edward, R.J.: An Introduction to Neural Networks. In: A White paper. Visual Numerics Inc., United States of America (2004)
3. Yam, Y.F., Chow, T.W.S.: Extended back propagation algorithm. *Electronics Letters* 29(19), 1701–1702 (1993)
4. Verma, B.K., Mulawka, J.J.: A modified back propagation algorithm. In: IEEE World Congress on Computational Intelligence, pp. 840–844 (1994)
5. Drago, G.P., Morando, M., Ridella, S.: An adaptive momentum back propagation. *Neural Computing and Application* 3, 213–221 (1995)
6. Chen, Y.Q., Yin, T., Babri, H.A.: A stochastic back propagation algorithm for training neural networks. In: International Conference on Information, Communications and Signal Processing, Singapore, pp. 703–707 (1997)
7. Ng, S.C., Leung, S.H., Luk, A.: Fast convergent generalized back propagation algorithm with constant learning rate. *Neural Processing Letters* 9, 13–23 (1999)
8. Wen, J.W., Zhao, J.L., Luo, S.W., Han, Z.: The improvements of BP neural network learning algorithm. In: ICSP 2000, pp. 1647–1649 (2000)
9. Wang, X.G., Tang, Z., Tamura, H., Ishii, M., Sun, W.D.: An improved back propagation algorithm to avoid the local minima problem. *Neuro Computing* 56, 455–460 (2004)
10. Wang, C.H., Kao, C.H., Lee, W.H.: A new interactive model for improving the learning performance of back propagation neural network. *Automation in Construction* 16(6), 745–758 (2007)
11. Bayati, A.Y., Al, S.N.A., Sadiq, G.W.: A modified conjugate gradient formula for back propagation Neural Network Algorithm. *Journal of Computer Science* 5(11), 849–856 (2009)
12. Mel, B.: Information processing in dendritic trees. *Neural Computing* 6, 1031–1085 (1994)
13. Yadav, R.N., Kalra, P.K., John, J.: Time series prediction with single multiplicative neuron model. *Applied Soft Computing* 7, 1157–1163 (2007)
14. Yadav, R.N., Singh, V., Kalra, P.K.: Classification using single neuron. In: IEEE Int. Conf. on Industrial Informatics, Banff, Alberta, Canada, pp. 124–129 (2003)
15. Yu, C.C., Liu, B.D.: A back propagation algorithm with adaptive learning rate and momentum coefficient. In: The International Joint Conference on Neural Networks, IJCNN 2002, pp. 1218–1223 (2007)
16. Zweiri, Y.H.: Optimization of a Three-Term Backpropagation Algorithm Used for Neural Network Learning. *International Journal of Engineering and Mathematical Sciences* 3(4), 322–327 (2007)

Technical White Paper on “Time and Frequency Synchronization in OFDM”

Anagha Rathkanthiwar¹ and Mridula Korde²

¹ Department of Electronics Engineering
Priyadarshini College of Engineering, Nagpur
anagharathkanthiwar@yahoo.co.in

² Department of Electronics Engineering
Shri Ramdeobaba Kamala Nehru College of Engineering
Nagpur
mridulakorde@yahoo.com

Abstract. Orthogonal frequency division multiple access (OFDMA) is an emerging standard for broadband wireless access. Synchronization in OFDMA represents one of the most challenging issues and plays a major role in the physical layer design. Aim of this paper is to provide an overview of various frequency and time synchronization errors in OFDM systems. This paper also discusses the effect of timing and frequency errors on system performance. It focuses on time and frequency error estimation algorithms for OFDM based systems in downlink transmission as well.

Keywords: Downlink synchronization, frequency correction, frequency estimation, orthogonal frequency division multiple access (OFDMA), orthogonal frequency division multiplexing (OFDM), timing estimation.

1 Introduction

The demand for wireless communications is growing today at an extremely rapid pace and this trend is expected to continue in the future. The common feature of many current wireless standards for high-rate multimedia transmission is the adoption of a multi-carrier air interface based on orthogonal frequency division multiplexing (OFDM). OFDM is a special case of multi-carrier transmission, where a single data stream is transmitted over a number of lower rate (thus increased symbol duration) sub-carriers. The increased symbol duration of OFDM symbol improves the robustness to channel delay spread and use of orthogonal sub-carriers makes efficient use of the spectrum by allowing overlap. It eliminates Inter Symbol Interference (ISI) and Inter Block Interference (IBI) through use of a cyclic prefix to some extent. Furthermore, it provides larger flexibility by allowing independent selection of the modulation parameters (like the constellation size and coding scheme) over each sub-carrier. OFDM Modulation can be realized with Inverse Fast Fourier Transform (IFFT). Due to all these favorable features, many digital transmission systems have adopted OFDM as the modulation technique such as digital video broadcasting

terrestrial TV (DVB-T), digital audio broadcasting (DAB), terrestrial integrated services digital broadcasting (ISDB-T), digital subscriber line (xDSL) etc. IEEE802.11a standard is the first one to use OFDM in packet-based communications, while the use of OFDM until now was limited to continuous transmission systems [1]. Now it is being used in other packet based systems like multimedia mobile access communications (MMAC), and the fixed wireless access (FWA) system in IEEE 802.16.3 standard [2]. It has become fundamental technology in the future 3GPP LTE and 4G-multimedia mobile communications systems.

Despite its appealing features, the design of an OFDM system poses several technical challenges. One basic issue is related to the stringent requirement on frequency and timing synchronization [4]. In OFDM downlink transmission, each terminal has to perform timing and frequency synchronization by exploiting the broadcast signal transmitted by the BS. Synchronization is performed in two phases one is acquisition phase and other is tracking phase. During acquisition step reference blocks are exploited to get coarse estimates of synchronization parameters i.e. frequency and timing errors. These estimates are then refined during tracking step.

1.1 Effect of Timing Errors

This section describes the effect of uncompensated timing error on system performance. Timing error occurs because of multipath dispersion. Due to this timing error the receiver’s time-domain FFT window spans samples from two consecutive OFDM symbols. This results in inter-OFDM symbol interference. Additionally, even small misalignments of the FFT window result in an evolving phase shift in the

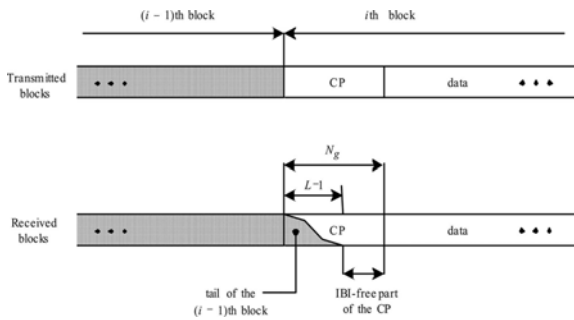


Fig. 1. Partial overlapping between received blocks due to multipath dispersion

frequency domain symbols, leading to BER degradation [5]. As shown in Fig. 1 the tail of each received block extends over the first $L - 1$ samples of the successive block as a consequence of multipath dispersion. If cyclic prefix is greater than Channel Impulse Response (CIR) duration (0 to $L-1$) then there is certain interval which is not affected by previous block. As long as the DFT window starts anywhere in this interval, no IBI is present at the DFT output. This situation occurs whenever the timing error $\Delta\theta$ belongs to interval $-N_g + L - 1 \leq \Delta\theta \leq 0$ and only results in a cyclic shift of the received OFDM block [6]. Thus, recalling the time-shift property of the

Fourier transform and assuming perfect frequency synchronization the DFT output over the n th subcarrier takes the form

$$R_i(n) = e^{j2\pi n\Delta\theta/N} H(n)d_i(n) + W_i(n)$$

Here

$$H(n) = \sum_{l=0}^{L-1} h(l)e^{-j2\pi nl/N} \quad (2)$$

where $h(l)$ is the channel impulse response and d_i is the i^{th} OFDM block. The equation of $R_i(n)$ indicates that timing error $\Delta\theta$ appears as a linear phase across subcarriers and it can be compensated by the channel equalizer. On the other hand, if the timing error is outside the interval $-Ng + L - 1 \leq \Delta\theta \leq 0$, samples at the DFT input will be contributed by two adjacent OFDM blocks. In addition to IBI, this results in a loss of orthogonality among subcarriers which, in turn, generates ICI. In this case, the n^{th} DFT output is given by

$$R_i(n) = e^{j2\pi n\Delta\theta/N} \alpha(\Delta\theta)H(n)d_i(n) + I_i(n,\Delta\theta) + W_i(n) \quad (3)$$

where $\alpha(\Delta\theta)$ is an attenuation factor while $I_i(n,\Delta\theta)$ accounts for IBI and ICI and can reasonably be modeled as a zero-mean random variable with power $\sigma_i^2(\Delta\theta)$. The loss in SNR parameter can be obtained by using

$$\gamma(\Delta\theta) = \frac{SNR^{ideal}}{SNR^{real}} \quad (4)$$

Where SNR^{ideal} is the SNR of a perfectly synchronized system and SNR^{real} is the SNR in the presence of a timing offset.

For a normalized channel response with unit average power $\gamma(\Delta\theta)$ is given by

$$\gamma(\Delta\theta) = \frac{1}{\alpha^2(\Delta\theta)} \left[1 + \frac{\sigma_i^2(\Delta\theta)}{\sigma_w^2} \right] \quad (5)$$

1.2 Effect of Frequency Errors

This section describes the effect of uncompensated carrier frequency error on system performance. Carrier frequency errors occurs due to offset between the incoming waveform and the local references used for signal demodulation. This error result in a shift of the received signal in the frequency domain. If the frequency error is an integer multiple of the sub-carrier spacing, then the received frequency domain modulated sub-carriers are shifted by sub-carrier positions. The sub-carriers are still mutually orthogonal but the received data symbols are now in the wrong position in the demodulated spectrum, causing increase in BER. If the carrier frequency error is not an integer multiple of the sub-carrier spacing, then it results in loss of mutual orthogonality between the sub-carriers. ICI is then observed between the sub-carriers, which deteriorates the BER performance of the system. To better explain this concept, assume ideal timing synchronization and compute the DFT output corresponding to

the i^{th} OFDMA block in the presence of a frequency error ε [6]. In case when frequency error ε is integer multiple of subcarrier spacing, the DFT output is given by

$$R_i(n) = e^{j\psi_i} H((n - \varepsilon|_N)d_i) + W_i(n) \quad (6)$$

Here $(n - \varepsilon|_N)$ is the value of $n - \varepsilon$ reduced to interval $[0, N-1]$. This equation shows that even though the received symbols appear in a wrong position at the DFT output, no ISI is present as orthogonality among subcarriers is preserved. The situation is drastically different when frequency error ε is not integer valued. In this case, the subcarriers are no longer orthogonal and DFT output can conveniently be rewritten as

$$R_i(n) = e^{j\psi_i} H(n)d_i f_N(\varepsilon) + I_i(n, \varepsilon) + W_i(n) \quad (7)$$

Here $I_i(n, \varepsilon)$ is a zero mean ICI term. Loss in SNR and $f_N(x)$ are given by (8) and (9) as follows

$$\gamma(\varepsilon) = \frac{1}{|f_N(\varepsilon)|^2} \left\{ 1 + \frac{D_2}{\sigma_w^2} \left[1 - |f_N(\varepsilon)|^2 \right] \right\} \quad (8)$$

$$f_N(x) = \frac{\sin(\pi x)}{N \sin(\pi x / N)} e^{j\pi x(N-1)/N} \quad (9)$$

2 Synchronization Algorithms for Downlink Transmission

As discussed OFDMA system is extremely sensitive to timing errors and carrier frequency offsets between the incoming waveform and the local references used for signal demodulation. Inaccurate compensation of the frequency offset destroys orthogonality among subcarriers and produces ICI. Timing errors result in IBI & ISI and must be counteracted to avoid severe error rate degradations. Therefore, synchronization is extremely crucial to the OFDM systems.

Synchronization algorithms can be divided into following categories:

1. Non data aided methods :based on use of internal structure of OFDM symbols.
2. Data aided methods : based on training symbols or pilots.

In continuous mode transmission systems, there is no stringent requirement on acquisition time hence averaging method can be used to improve estimation accuracy. It is appropriate to apply non data aided methods which makes use of CP for this mode. However in the burst packet mode, synchronization ought to be established at any time because when data streams are ready to transmit is unknown hence there is stringent requirement on synchronization time.

The report [1] described use of CP (non data aided method) for both timing and frequency synchronization considering that first T_G (duration of CP) seconds of each symbol is identical to last part. This algorithm correlates a T_G long part of the signal with a part that is T (symbol duration) seconds delayed. Frequency offset is then estimated by averaging correlation output function over interval T_G and finding phase of this output function.

The conventional algorithms for the coarse symbol timing synchronization in time domain are MLE (Maximum Likelihood Estimation) utilizing the cyclic prefix of the OFDM symbols. This technique of synchronization was proposed by J. J. Van de

Beek [2]. However, good performance is achieved only for AWGN channel. For multipath fading channels data is badly contaminated by ISI and there is significant influence on carrier frequency offset (CFO).

To improve performance of MLE under multipath fading channels, a novel scheme was proposed in [3]. This scheme utilizes both redundancy in CP & Pilots symbols for coarse symbol timing synchronization. This estimator allow a shorter CP and thus a more spectrally efficient system.

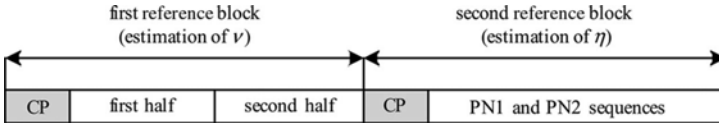


Fig. 2. Training Symbol of S & C algorithm

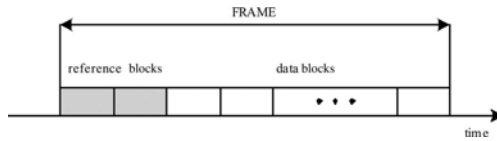


Fig. 3. Training Symbols as reference blocks placed in the beginning of frame

But because of limited number of pilots used for estimation, non negligible fluctuation still exists. In order to mitigate the fluctuation T.M. Schmidl and D.C. Cox [4] (S & C) introduced the method allowing large acquisition range for the carrier frequency offset by using one unique symbol composed of two identical halves of half the length of symbol which is transmitted at the beginning of each frame as shown in Fig. 3. The symbol is shown in Fig.2. The algorithm is based on concept that a training symbol with two identical halves in the time domain, will remain identical after passing through the channel, except that there will be a phase difference between them caused by the carrier frequency offset. Consider the training symbol where the first half is identical to the second half (in time order), except for a phase shift caused by the carrier frequency offset. If the conjugate of a sample from the first half is multiplied by the corresponding sample from the second half (seconds later), the effect of the channel should cancel, and the result will have a phase of approximately $\varphi = \pi T \Delta f$. At the start of the frame, the products of each of these pairs of samples will have approximately the same phase, so the magnitude of the sum will be a large value. As long as the CP is longer than the CIR duration, the two halves of the reference block will remain identical after passing through the transmission channel except for a phase shift induced by the CFO and the received samples corresponding to the first half as

$$r(k) = s(R)(k)e^{j2\pi ck/N} + w(k), \quad \theta \leq k \leq \theta + N/2 - 1 \tag{10}$$

Received samples corresponding to second half are given by

$$r(k + N/2) = s(R)(k)e^{j2\pi ck/N} e^{j\pi c} + w(k + N/2), \quad \theta \leq k \leq \theta + N/2 - 1 \tag{11}$$

Timing estimate is

$$\hat{\theta} = \arg \max \left\{ \Gamma(\tilde{\theta}) \right\} \tag{12}$$

Where

$$\Gamma(\tilde{\theta}) = \frac{\sum_{q=\tilde{\theta}}^{\tilde{\theta}+\frac{N}{2}-1} r(q+N/2)r^*(q)}{\sum_{q=\tilde{\theta}}^{\tilde{\theta}+\frac{N}{2}-1} |r(q+N/2)|^2}$$

The above method is rapid and suitable for continuous transmission or a burst operation over a frequency-selective channel. Unfortunately metric of this algorithm exhibits a large plateau that may greatly reduce the estimation accuracy. The start of the frame and the beginning of the symbol can be found, and carrier frequency offsets of many subchannels spacing can be corrected. The algorithms operate near the Cramer–Rao lower bound for the variance of the frequency offset estimate, and the inherent averaging over many subcarriers allows acquisition at very low signal-to-noise ratios (SNR’s).

JungJu Kim; Jungho Noh; KyungHi Chang, presented [5] a preamble timing synchronization method for OFDMA timing estimation as modification to T. M. Schmidl and D. C. Cox to reduce uncertainty due to timing metric plateau. The training symbol used in this method consists of four segments in which first & third segments are identical and second & fourth segments are symmetric conjugate with first segment as shown in figure 4. Therefore, the proposed efficient timing synchronization method guarantees the better performance of the initial timing synchronization of OFDMA systems.

Thomas Keller, Lorenzo Piazzi, Paolo Mandarini, and Lajos Hanzo in their paper [7] suggested use of cyclic postamble to mitigate effect of timing errors. The reference symbols used by their algorithm consists of repetitive copies of a synchronization pattern as shown in figure 5 and proposed algorithm rely on evaluation of correlation functions as given in equations (14), (15) one corresponding to OFDM symbol and other corresponding to reference symbol.



Fig. 4. Training symbol used by algorithm in [5]

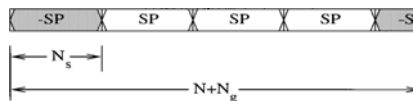


Fig. 5. Repetitive copies of a synchronization pattern

$$G(J) = \sum_{m=0}^{N_g-1} z(j-m).z(j-m-N)^*$$

$$R(J) = \sum_{m=0}^{N+N_g-N_s-1} z(j-m).z(j-m-N)^*$$

$G(J)$ is used both for symbol synchronization and frequency tracking. Where as $R(J)$ is used for frequency acquisition. Also they proposed in the same paper joint frequency and time synchronization acquisition algorithm using periodicity of CP of OFDM symbol.

Two algorithms for timing synchronization using preamble with special properties consisting of two part each of one symbol duration were described by M. Gertou, G. Karachalios, D. Triantis, K. Papantoni and P. I. Dallas [8]. One algorithm utilizes only first part of preamble and take average of two detected positions obtained in the defined metric function giving better timing estimation. Second algorithm uses first part of preamble for coarse and second part for fine symbol synchronization.

Moose, P.H in his paper [9] describes a technique to estimate frequency offset using repeated data symbols with MLE. It has been shown that for small error in the estimate, the estimate is conditionally unbiased and is consistent in the sense that the variance is inversely proportional to the number of carriers in the OFDM signal. Furthermore, both the signal values and the IC1 contribute coherently to the estimate so that it is possible to obtain very accurate estimates even when the offset is too great. Since the estimation error depends only on total symbol energy, the algorithm works equally well in multipath spread channels and frequency selective fading channels. However, it is required that the frequency offset as well as the channel impulse response be constant for a period of two symbols.

The synchronization schemes described in [1], [7], [2], [9] are Non Data Aided methods and the schemes described in [3], [4], [5], [8] are Data Aided methods of synchronization.

3 Conclusion

As mentioned in the earlier section, synchronization methods can be divided into two categories: DA and NDA algorithms. Pilots, training symbols or the combination of them are generally applied to the DA-type of methods achieving synchronization in less time but at the expense of the reduced bandwidth efficiency. For the NDA methods, data used for the estimation may be contaminated by ISI, resulting in the inaccurate estimation. Throughput and power efficiency are improved but time taken to establish synchronization is increased. Synchronization time, algorithm complexity, the required system performance and etc. are all the factors that should be considered when choosing the synchronization scheme for the particular system.

After studying and analyzing various available OFDM synchronization algorithms a suitable algorithms can be developed. Simulation of these algorithms can be performed for various parameters like probability of missed detection, probability of

false alarm, for different values of threshold. Also simulation of estimation error probability vs SNR can be done. Complexity of algorithms vs performance kind of simulation can also be done.

Design and appropriate use of reference symbols to be used for synchronization is important. So suitable reference symbol can be developed for estimation of time and frequency errors.

References

1. Intini, A.L.: Orthogonal frequency division multiplexing for wireless networks standard IEEE 802.11a. University Of California Santa Barbara (December 2000)
2. van de Beek, J.J., Sandell, M., Borjesson, P.O.: ML estimation of time and frequency offset in OFDM systems. *IEEE Transactions on Acoustics, Speech and Signal Processing* 45(7), 1800–1805 (1997)
3. Landström, D., Wilson, S.K., Van de Beek, J.J., Odling, P., Börjesson, P.O.: Symbol time offset estimation in coherent OFDM systems. In: *Proc. Int. Conf. On Communications, Vancouver, BC, Canada*, vol. 1, pp. 500–505 (June 1999)
4. Schmidl, T.M., Cox, D.C.: Robust frequency and timing synchronization for OFDM. *IEEE Trans. on Commun.* 45(12), 1613–1621 (1997)
5. Kim, J., Noh, J., Chang, K.: An efficient timing synchronization method for OFDMA system. In: *IEEE/ACES International Conference on Wireless Communications and Applied Computational Electromagnetics*, April 3-7, pp. 1018–1021 (2005)
6. Morelli, M., Kuo, C.-C.J., Pun, M.-O.: Synchronization Techniques for Orthogonal Frequency Division Multiple Access (OFDMA): A Tutorial Review. *Proceedings of the IEEE* 95(7), 1394–1427 (2007)
7. Keller, T., Piazza, L., Mandarini, P., Hanzo, L.: Orthogonal Frequency Division Multiplex synchronization techniques for frequency-selective fading channels. *IEEE Journal on selected areas in communications* 19(6) (June 2001)
8. Gertou, M., Karachalios, G., Triantis, D., Papantoni, K., Dallas, P.I.: Synchronization Approach for OFDM based Fixed Broadband Wireless Access Systems. *INTRACOM S.A*
9. Moose, P.H.: A technique for orthogonal frequency division multiplexing frequency offset correction. *IEEE Transactions on Communications* 42(10), 2908–2914 (1994)

Cell-ID Based Vehicle Locator and Real-Time Deactivator Using GSM Network

Nilesh Dubey¹, Vandana Dubey¹, and Shivangi Bande²

¹ Sanghvi Institute of Management and Science, Indore-453332

² Institute of Engineering and Technology DAVV, Indore-452009

Abstract. The work “Cell-ID Based Vehicle Locator and Real-Time Deactivator Using GSM Network” (VLRD) uses GSM network for locating and real-time controlling on the vehicle by using the cell-id sent by guardian software and real time decoding of the encoded commands which are communicated through the circuit switched network. The system gets the command from the user and acts according to a pre-programmed set of instructions in real time to avoid the problems of delayed or undelivered commands in the existing system and improves the assurance of command execution. Removing the GPS usage can be reached through detecting the moments that Cell-ID is providing sufficient accuracy for the envisaged application.

Keywords: Cell-ID, DTMF, Embedded, GSM, Immobilizer.

1 Introduction

NICB (National Insurance Crime Bureau) of United States had a report which stated that there were more than one million cars stolen in 2006. The value of the stolen cars was about ninety seven hundred millions dollars, and every twenty six seconds there was a car being stolen. Even advanced technologies such as electric locks and electronic immobilizers have been applied to vehicles; cars are still stolen by thieves to resell the components or the whole car.

The presently used systems are very costly and need high maintenance cost as well. Presently used system uses GPS tracking which can only track the vehicle but cannot send control command to vehicle. Some systems allow immobilizing the vehicle through sending the control command through SMS.

But **the problem with the presently used technique is**, they uses the GPS technology for getting the position of the vehicle and sends the GPS data via GSM/GPRS modem through SMS.

1. These units are very expensive approximately Rs.16000 to 25000 in Indian market.
2. Second problem is monthly rental of the service provided which is approximately Rs.500 per month + SMS charges + GPRS charges.
3. The strength of the GPS signals, which is very weak in the urban areas, under the tree, basement of building or any type of shadows. The GPS almost in non working condition in case of weak signals

4. GPS need a long antenna for receiving the signals from the satellites.

Not all the tracking unit provides the function of vehicle immobilization and which are providing are facing the problem of non real-time controlling on the vehicle because of the undelivered or delayed SMS command.

The work suggests solution of the above problems. It uses only GSM network for the whole working of the system. The mobile phone market lacks a satisfactory location technique that is accurate, but also economical and easy to deploy. Current technology provides high accuracy, but requires substantial technological and financial investment.[4]

The GSM signals are much stronger than the GPS signals and no need of any external antenna in the unit. It is very cost effective and it sends the location data only when we need.

The user needs not to pay any monthly charges. It gives the real time controlling on the vehicle because of the DTMF commands over the voice channel which has high priority than the data channel which used for sms service.

VLRD uses GSM network for locating the vehicle. It uses the database of the GSM service provider for getting the location of the cell in which the vehicle is currently running. The hidden unit inside the vehicle sends the Cell-ID of the current location when we need. The same network used to send the DTMF (Dual tone multiple frequency) commands from the remote side to the hidden unit. The DTMF command activates pre-programmed tasks to De-activate the vehicle and for other actions like alarming, photo capturing, Video recording, central car locking etc. The VLRD also provides the facility of full duplex communication between driver and remote user. The VLRD is useful in case of vehicle jacking and lifting and also used in transportation business for finding the location for managing the time and to keep eyes on your vehicles.

DTMF is a generic communication term for touch tone (a Registered Trademark of AT&T). The tones produced when dialing on the keypad on the phone could be used to represent the digits, and a separate tone is used for each digit.

DTMF dialing uses a keypad with 12/16 buttons. Each key pressed on the phone generates two tones of specific frequencies. One tone is generated from a high frequency group of tones and the other from low frequency group[3]. The frequencies generated on pressing different phone keys are shown in the Table 1.

Table 1. DTMF Frequency Assignment

Button	Low freq. (Hz)	High Freq. (Hz)
1	697	1209
2	697	1336
3	697	1477
4	770	1209
5	770	1336
6	770	1477
7	852	1209
8	852	1336
9	852	1477
0	941	1209
*	941	1336
#	941	1477

2 System Design and Methodology

The concept is, whenever a user wants to immobilize or locate the vehicle which is installed with the VLRD. He just have to make a call on the mobile attached with the system, when call received he has to dial a password to enter in the system operate mode.

Then he has to press the key according to which action he wants to perform on the vehicle either of immobilization or location. If he wants location than mobile sends a SMS which contains Cell-ID data.

The Cell-ID data contains LAC(Location Area Code), MCC(Mobile Country Code), NAC(Network Area Code) and Operator code.

The work is basically divided in two parts one is Hardware which is an Embedded system shown in fig. 1 and second is Software which is a Web based service.

2.1 Hardware Section

The Embedded system is interfaced with the Symbian 60 series of mobile phone here we used Nokia N-72. The mobile is on auto answer mode so that when a call make from remote side the call will automatically received.

After call received any key pressed from remote mobile sends a unique DTMF signal to the system’s mobile according to the Table-1 which will be decoded in BCD through a DTMF to BCD Decoder CM8870 [5] this decoded commands are than send to the Microcontroller AT89c51 [6]. The outputs of the Microcontroller activate/deactivates the relays using relay driver ULN 2003 [7] according to the command input and the embedded program.

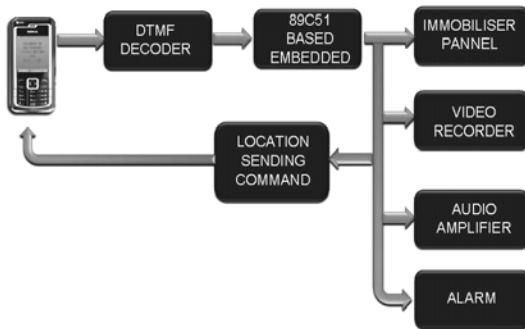


Fig. 1. Hardware System Block Diagram

These relays then perform different actions.

1. The VLRD’s one of the most important works is to immobilize the vehicle and this task done by the immobilizer panel. The immobilizer immobilizes the vehicle by many types. The VLRD activates and deactivates some other systems to immobilize the vehicle like deactivating the ignition system, deactivating fuel pump, locking gear, locking steering, etc.

2. A video recording facility is available in the VLRD for legal evidence and recognition of theft/ unknown person.
3. It provides full duplex voice communication facility between driver and the remote user using car's audio system.
4. The Project VLRD has an loud alarm system which can be activated remotely through a mobile phone. The purpose of the alarm is to alert the surrounding persons about the theft car and also useful to locate the car in a particular cell.
5. The VLRD sends the Cell-ID of the particular GSM cell in which the vehicle is running using Guardian software which is free available on internet. Every time the user powers up the phone, Guardian automatically starts and check if the inserted SIM card is present in the authorized list. If this is not authorized the software will send a notification SMS to the number previously stored. [9]

The restarting of system's mobile is done through microcontroller. The circuit is shown in Fig. 2.

Cell-ID: The mobile operators keep the locations of GSM masts a secret. Every GSM mast sends its Cell ID shown in fig. 3, and it is possible to read this Cell ID of the nearest-by mast on your GSM mobile. By gathering these Cell IDs with accompanying locations and by putting them into a database, it should be possible to read out Mobile's current location. With this information it is possible to make location based applications using GSM Mobile.

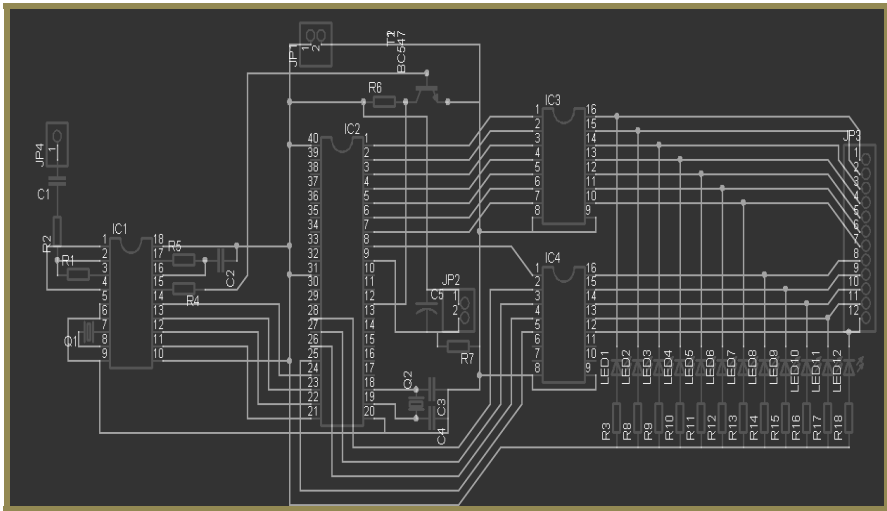


Fig. 2. Complete circuit diagram

2.2 Software Section

The Geolocation API is an abstraction for various location APIs that currently exist on mobile platforms (GPS-based, network/cellid-based). Geolocation implementations could be straightforward mappings to native APIs (e.g the S60 Location Acquisition

API) or have a more complex design that combines several location providers (e.g. a GPS-based provider and a cell id-based provider) and returns the location from the most accurate provider at any given time.[8].

The cell-id based location service is a Web based application which locates the vehicle on Google Map used in Website according to the Cell-ID data entered. It compares the input data to the cell ID database which is free provided by Google. The Google has a worldwide cell id database which can be fetched by using Google API on any website.

This database has the Geo co-ordinates of the Cell specified by Cell-ID.

3 Conclusion

A novel method of designing a low-cost, compact theft control system for a vehicle was designed demonstrated in this paper and also practically implemented by us. This work is an ultimate threat for vehicle thieves. By this work which is presented in this paper, it is very easy to track the vehicle at a higher degree of reliability, since it is based on GSM Technology, which is very developed now.

References

1. Nagaraja, B.G., Rayappa, R., Mahesh, M., Patil, C.M., Manjunath, T.C.: Design & Development of a GSM Based Vehicle Theft Control System. In: IEEE International Conference on Advanced Computer
2. Li, S., Xiong, Z., Li, T.: Distributed Cooperative Design Method and Environment for Embedded System. In: Proceedings of the 9th International Conference on Computer Supported Cooperative Work in Design, pp. 956–960.
3. Ladwa, T.M., Ladwa, S.M., Kaarthik, R.S., Ranjan, A., Dhara, N.D.: Control of Remote Domestic System Using DTMF. In: IEEE ICICI-BME 2009 Bandung, Indonesia (2009)
4. Trevisani, E., Vitaletti, A.: Cell-ID location technique, limits and benefits: an experimental study. In: Proceedings of the Sixth IEEE Workshop on Mobile Computing Systems and Applications, WMCSA 2004 (2004)
5. CM8870C DTMF Decoder Datasheet, <http://www.calmicro.com/products/data/pdf/cm8870.pdf>
6. Microcontroller 89c51 datasheet, <http://www.atmel.com/atmel/acrobat/doc0265.pdf>
7. <http://pdf1.alldatasheet.co.kr/datasheetpdf/view/25575/STMICROELECTRONICS/ULN2003.html>
8. <http://code.google.com/p/gears/wiki/GeolocationAPI>
9. <http://www.guardianmobile.com/usage.pdf>

A Novel Design of Reconfigurable Architecture for Multistandard Communication System

T. Suresh¹ and K.L. Shunmuganathan²

¹ Research Scholar, R.M.K Engineering College,
Anna University, Chennai, India
fiosuresh@yahoo.co.in

² Professor & Head, Department of CSE, R.M.K Engineering College,
Anna University, Chennai, India
kls_nathan@yahoo.com

Abstract. The goal of future mobile communication systems will be to incorporate and integrate different wireless access technologies and mobile network architectures in a complementary manner so as to achieve a seamless wireless access infrastructure. To support this seamless user mobility across different wireless access technologies it is needed to design reconfigurable multistandard receiver architecture. This paper presents the system-level design of a wireless receiver's baseband architecture, which supports two wireless access technologies: Wideband Code Division Multiple Access (WCDMA) and Orthogonal Frequency Division Multiplexing (OFDM). In this paper, efficient method of Fast Fourier Transform (FFT) algorithm for OFDM standard and Rake Receiver design for WCDMA standard were implemented. This architecture efficiently shares the resources needed for these two standards while reconfiguring. The proposed architecture is simulated using ModelSimSE v6.5 and mapped onto a Spartan 3E FPGA device (3s5000epq208) using the tool Xilinx ISE 9.2. Simulation results show that the proposed architecture can be efficiently reconfigured in run-time and proved as area efficient.

Keywords: FPGA, WCDMA, OFDM, FFT/IFFT, Rake Receiver, Reconfigurable.

1 Introduction

The need to support several standards in the same handheld device, associated with the power consumption and area restrictions, created the necessity to develop a portable, power efficient, integrated solution [1]. Users carrying an integrated open terminal can use a wide range of applications provided by multiple wireless networks, and access to various air interface standards. The continuous evolution of wireless networks and the emerging variety of different heterogeneous, wireless network platforms with different properties require integration into a single platform [2]. This has lead to an increased interest in the design of reconfigurable architecture. The idea of the reconfigurable architecture is that it should be possible to alter the functionality of a mobile device at run-time by simply reusing the same hardware for different

wireless technologies and ultimately for users to connect to any system that happens to be available at any given time and place.

2 Reconfigurable Architecture

Multistandard wireless communication applications demand high computing power [3], flexibility, and scalability. An Application-Specific Integrated Circuit (ASIC) solution would meet the high computing power requirement, but is inflexible [4] and the long design cycle of ASICs makes them unsuitable for prototyping. On the other hand, general purpose microprocessors or Digital Signal Processing (DSP) chips are flexible, but often fail to provide sufficient computing power. Field Programmable Gate Arrays (FPGAs) [5] signal processing platforms are now widely being accepted in base_station designs. However, low power and form factor requirements have prevented their use in handsets. Reconfigurable hardware for Digital Base-Band (DBB) [6] processing is rapidly gaining acceptance in multi-mode handheld devices that support multiple standards. In Reconfigurable Hardware tasks that are required initially can be configured in the beginning. When another task is required, the configuration to load it can then be triggered. In this paper we present the design methodology for reconfigurable baseband signal processor architecture that supports WCDMA and OFDM wireless LAN standards.

During recent years, a number of research efforts focused on the design of new reconfigurable architectures. In [7] the flexibility of the MONTIUM architecture was verified by implementing HiperLAN/2 receiver as well as a Bluetooth receiver on the same architecture. In [8], a broadband mobile transceiver and a hardware architecture which can be configured to any cyclic-prefix(CP) based system reconfigurable architecture for multicarrier based CDMA systems is proposed. Reconfigurable Modem (RM) Architecture targeting 3G multi-standard wireless communication system was proposed in [3]. This architecture targeted two 3G wireless standards WCDMA and CDMA 2000 and the design objectives are scalability, low power dissipation and low circuit complexity. It is seen that though different functions can be reconfigured on a reconfigurable hardware, the major challenge is to have an efficient system configuration and management function which will initiate and control the reconfiguration as per the different application requirements.

3 Reconfigurable Receiver Architecture

Figure 1 shows the Block Diagram of Reconfigurable Receiver System. This Receiver System is able to reconfigure itself to the WCDMA or OFDM Wireless LAN (WLAN) standard. The Proposed architecture comprises functional blocks, which is in the form of reusable, reconfigurable [9-12] functional blocks for use in implementing different algorithms necessary for OFDM and WCDMA standards. One or more reusable functional blocks as given in Fig. 1, can be configured to implement a process including multiplication, addition, subtraction and accumulation. By accommodating the above mentioned capabilities, the architecture should be configured to support WCDMA and WLAN OFDM Standards. For example Fast Fourier Transform (FFT) (basic butterfly function) for WLAN OFDM and Rake.

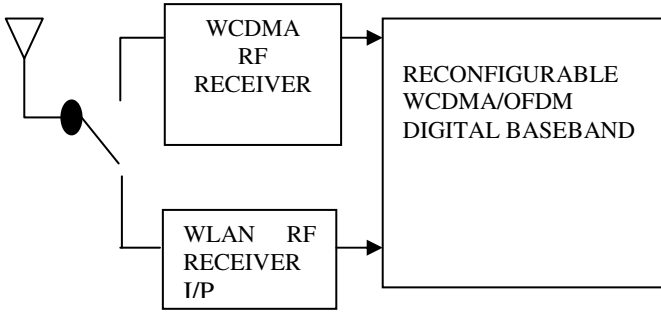


Fig. 1. Block Diagram of Reconfigurable Receiver System

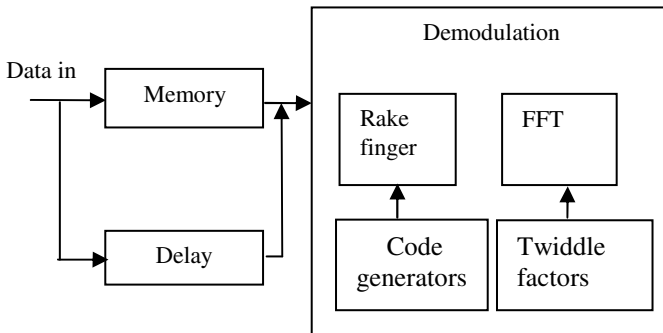


Fig. 2. Block Diagram of Reconfigurable Receiver Architecture

Receiver algorithms (multiply and accumulate select function) for WCDMA are implemented in the architecture as shown in Fig.2. This architecture allows for transformation of the chip from WCDMA chip to WLAN Wi-Fi chip on-demand wherein new algorithms can be accommodated on-chip in real time via different control sets.

3.1 Rake Finger Implementation

In WCDMA receivers, the demodulation is performed in the Rake fingers by correlating the received signal with a spreading code over a period corresponding to the spreading factor. The output of the i^{th} Rake finger can be expressed as

$$O_i(n) = \sum_{i=0}^{L_{sf}-1} C_s(i + nL_{sf})R(i + nL_{sf}). \tag{1}$$

where C_s is the combined spreading and scrambling code and R is the received signal and both are complex numbers [3]. Since the scrambling and spreading codes are always of +/-1, the multiplication and addition of each correlation stage are simplified. So the equation(1) is simplified to

$$(R_r + jR_i)(C_{sr} - jC_{si}) = (R_r C_{sr} + R_i C_{si}) + j(R_i C_{sr} - R_r C_{si}). \tag{2}$$

If the value +1 is represented as logic ‘0’ and the value -1 is represented as logic ‘1’, the equation (4) is simplified as follows

$$= \begin{cases} R_r+R_i+j(R_i-R_r), & \text{when } C_{sr}=0, C_{si}=0 \\ R_r-R_i+j(R_i+R_r), & \text{when } C_{sr}=0, C_{si}=1 \\ -(R_r-R_i)-j(R_i-R_r), & \text{when } C_{sr}=1, C_{si}=0 \\ -(R_r+R_i)-j(R_i-R_r), & \text{when } C_{sr}=1, C_{si}=1 \end{cases} \quad (3)$$

Since the code input is binary valued, the complex multiplication in the correlations is simplified to one real addition/subtraction and one imaginary addition/subtraction. Selection of addition or subtraction is done with the help of multiplexer. So the total resources required to implement Rake Receiver using (3) are two adders, two subtractors and one multiplexer.

3.2 FFT Implementation

In OFDM, the demodulation is performed by applying 64-point FFT. The twiddle factor is calculated and put in a table in order to make the computation easier and can run simultaneously. The Twiddle Factor table is depending on the number of points used. During the computation of FFT, this factor does not need to be recalculated since it can refer to the Twiddle factor table, and thus it saves time. Figure 3 shows the 2 point Butterfly structure [13] where multiplication is performed with the twiddle factor after subtraction.

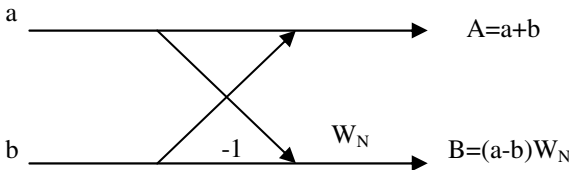


Fig. 3. 2 Point Butterfly Structure

Multiplication is certainly the most vital operation in Communication processing, and its implementation in an integrated circuit component requires large hardware resources and significantly affects the size, performance, and power consumption of a system [14]. So an efficient way of multiplier reduction in FFT processing is done as follows. Consider the problem of computing the product of two complex numbers R and W

$$\begin{aligned} X = RW &= (R_r+jR_i)(W_r+jW_i) \\ &= (R_rW_r-R_iW_i)+j(R_rW_i+R_iW_r) \end{aligned} \quad (4)$$

From equation (4), the direct architectural implementation requires total of four multiplications and one real subtraction and one imaginary addition to compute the complex product. However, by applying the Strength Reduction Transformation we can reformulate equation (4) as:

$$X_r=(R_r-R_i)W_i+R_r(W_r-W_i) \tag{5a}$$

$$X_i=(R_r-R_i)W_i+R_i(W_r+W_i) \tag{5b}$$

As can be seen from Equations (5a) and (5b), by using the Strength Reduction Transformation the total number of real multiplications is reduced to only three. This however is at the expense of having two additional subtractors and one adder.

4 Processing Element

Figure 4 shows the Processing Element(PE) and its resources required for the implementation of FFT in WLAN OFDM and figure 5 shows the Processing Element(PE) and its resources required for the implementation of Rake finger in WCDMA. It is shown that the two adders and subtractors(red coloured) are shared by both the standards. So the proposed Reconfigurable Architecture consists of processing units ,their computational elements are shared by both the Rake Receiver operation of WCDMA and FFT operation of OFDM. The processing units perform the multiply-accumulate operation in the Rake mode as described in section 3.1 and butterfly operations in the FFT mode as described in section 3.2. The computational resources required by the proposed architecture are 5 adders, 4 subtractors, 3 multipliers and multiplexers.

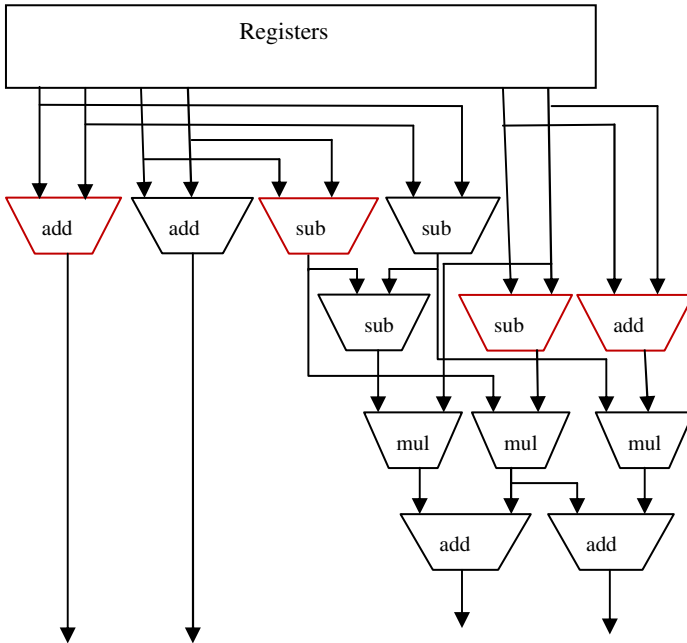


Fig. 4. PE and its Resources of WLAN OFDM

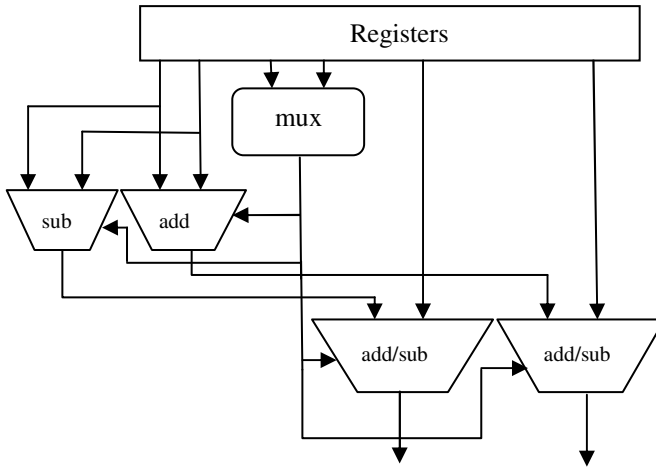


Fig. 5. PE and its Resources of WCDMA

5 Results and Discussion

The proposed reconfigurable architecture described in section 3 and 4 were simulated using ModelSimSE v6.5 and mapped onto a Spartan 3E FPGA device (3s5000epq208) with speed grade (-5) using the tool Xilinx ISE 9.2 and synthesized. The proposed Reconfigurable Architecture with Resource sharing is compared with Reconfigurable Architecture without Resource sharing. Table 1 and Figure 6 show the Resources utilized by the proposed Architecture(Reconfigurable Architecture with Resource sharing) and the Reconfigurable Architecture without Resource sharing. From the results presented above it seems that there is a significant reduction in large number of computational resources which forms the proposed architecture which is more efficient than the conventional Architecture in terms of area.

Table 1. Resource utilization of Reconfigurable Architecture without and with Resource sharing

Resources Utilized	Reconfigurable Architecture without Resource Sharing	Reconfigurable Architecture with Resource Sharing
Number of Slices	1172 out of 4656 (25%)	1070 out of 4656 (23%)
Number of 4 input LUTs	2195 out of 9312 (23%)	2048 out of 9312 (22%)
Number of IOBs	140 out of 158 (88%)	140 out of 158 (88%)

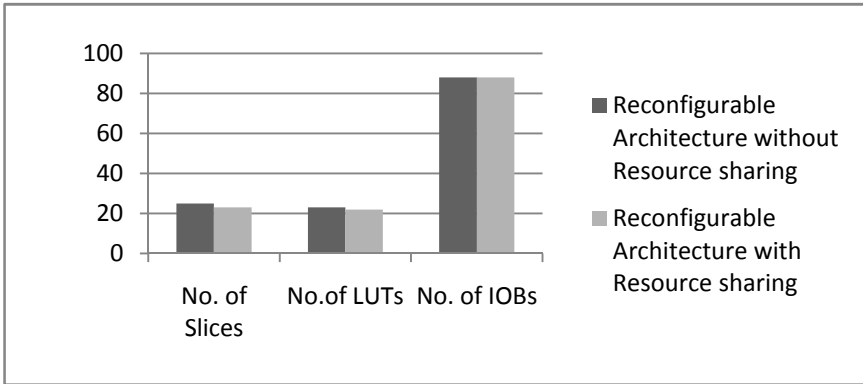


Fig. 6. Comparison of the percentage of resources utilized by Reconfigurable Architecture without Resource Sharing and with Resource Sharing

6 Conclusion

An architecture which can reconfigure itself to wireless LAN OFDM and WCDMA standards, was presented in this paper. While configuring these two standards, it was also presented to implement FFT operation for OFDM and Rake Receiver functioning for WCDMA efficiently. To lower the number of multipliers in FFT and eliminate the multipliers in Rake Receiver, we adopted Strength Reduction Transformation technique and multiplier-less technique. The proposed architecture was simulated using ModelSimSE v6.5 and mapped onto a Xilinx Spartan 3E FPGA device and synthesis report was generated. Simulation results demonstrated that the proposed architecture can reduce hardware overhead, enhance circuit efficiency and significantly reduce area. Moreover, the proposed architecture can be improved to reconfigure to various other advanced wireless standards.

References

1. Atallah, J.G., Ismail, M.: Future 4G front-ends enabling smooth vertical handover. *IEEE Circuits and Devices Magazine* XXII, 6–15 (2006)
2. Liljana Gavrilovska, M., Vladimir Atanasovski, M.: Interoperability in Future Wireless Communications systems: A Roadmap to 4G. *Microwave Review*, 19–28 (2007)
3. Lee, J.-S., Ha, D.S.: FleXilicon: a Reconfigurable Architecture for Multimedia and Wireless Communications. In: *IEEE International Symposium on Circuits and Systems*, pp. 4375–4378 (2006)
4. Kim, J., Ha, D.S.: A New Reconfigurable Modem Architecture for 3G Multi-Standard Wireless Communication Systems. In: *IEEE International Symposium on Circuits and Systems*, pp. 1051–1054 (2005)
5. David, R., Chillet, D., Pillement, S., Sentieys, O.: A compilation framework for a dynamically reconfigurable architecture. In: Glesner, M., Zipf, P., Renovell, M. (eds.) *FPL 2002*. LNCS, vol. 2438, pp. 153–194. Springer, Heidelberg (2002)

6. Harju, L., Nurmi, J.: A Programmable Baseband Receiver Platform for WCDMA/OFDM Mobile Terminals. In: IEEE Wireless Communications and Networking Conference, USA, pp. 33–38 (2005)
7. Rauwerda, G.K., Smit, G.J.M., van Hoesel, L.F.W., Heysters, P.M.: Mapping Wireless Communication Algorithms to a Reconfigurable Architecture. *The Journal of Supercomputing* 30, 263–282 (2004)
8. Liang, Y.-C., Naveen, S., Pilakkat, S.K., Marath, A.K.: Reconfigurable Signal Processing and Hardware Architecture for Broadband Wireless Communications. *EURASIP Journal on Wireless Communications and Networking* 3, 323–332 (2005)
9. Hauck, S., Fry, T.W., Hosler, M.M., Ko, J.P.: The Chimaera Reconfigurable Functional Unit. *IEEE Transactions on Very Large Scale Integration(VLSI) Systems* 12(2), 206–217 (2004)
10. Parizi, H., Niktash, A., Kamalizad, A., Bagherzadeh, N.: A Reconfigurable Architecture for Wireless Communication Systems. In: Third International Conference on Information Technology: New Generations, pp. 250–255 (2006)
11. Hartenstein, R.: Coarse Grain Reconfigurable Architectures. In: Conference on Asia South Pacific Design Automation, pp. 564–570 (2001)
12. Qu, Y., Tiensyrj, K., Soininen, J.-P., Nurmi, J.: Design Flow Instantiation for Run-Time Reconfigurable Systems: A Case Study. *EURASIP Journal on Embedded Systems* 11 (2008)
13. Heysters, P., Smit, G., Molenkamp, E.: A Flexible and Energy-Efficient Coarse- Grained Reconfigurable Architecture for Mobile Systems. *The Journal of Supercomputing* 26, 283–308 (2003)
14. Hinkelmann, H., Zipf, P., Li, J., Liu, G., Glesner, M.: On the design of reconfigurable multipliers for integer and Galois field multiplication. *Microprocessors & Microsystems* 33(1), 2–12 (2009)

Two Novel Long-Tail Pair Based Second Generation Current Conveyors (CCII)

Amisha Naik¹ and N.M. Devashrayee²

¹ Asst. Prof, EC Department, Institute of Technology, Ahmedabad-382481, Gujarat, India
a_p_niak@yahoo.com

² Asso. Prof, EC Department, Institute of Technology, Ahmedabad-382481, Gujarat, India
nalin_deepika@yahoo.com

Abstract. Two novel long tail pair based CCII are proposed in this paper. The first OTA CCII offer 1.98GHZ current transfer bandwidth and 10MHZ voltage transfer bandwidth. The second proposed design of CCII is independent against bias current variation. The second proposed design offer voltage transfer bandwidth of 10MHZ and current transfer bandwidth of 1.2GHZ with a very accurate voltage and current copy at corresponding x and Z nodes. The none idealities of first CCII are also measured using Spice simulation with TSMC 180nm model parameters.

Keywords: OTA CCII, Voltage transfer bandwidth, Current transfer bandwidth.

1 Introduction

The current-mode approach considers the information flowing on time-varying currents. Current-mode techniques are characterized by signals as typically processed in the current domain. The current-mode approach is also powerful if we consider that all the analog IC functions, which are traditionally designed in the voltage-mode, can also be implemented in current-mode. In voltage mode circuits, the main building block used to add subtract, amplify, attenuate, and filter voltage signals is the operational amplifier. A current-mode approach is not just restricted to current processing, but also offers certain important advantages when interfaced to voltage-mode circuits. Since the introduction of conveyors in early 70's, lot of research has been carried out to prove usefulness of this CCII. The CCII is a functionally flexible and versatile, rapidly gaining acceptance as both a theoretical and practical building block. Internal architecture of CCII is voltage follower cascaded with current follower as shown in figure:1.

2 OTA CCII

The simple current Mirror based CCII can be improved by replacing Simple diode connected level shifter in a flipped voltage follower based simple current mirror CCII

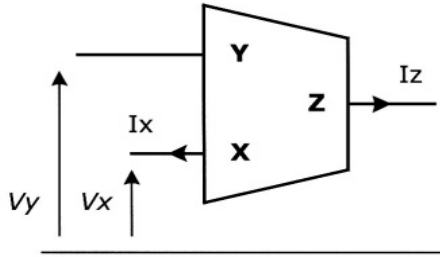


Fig. 1. CCII Block Representation

proposed by A.J.Lopez martin etal by OTA level shifter. The resulting circuit is depicted in figure:2. The proposed circuit is a combination of a OTA and FVF cell. In the proposed circuit(fig:2), transistors M1 to M4 form differential pair. Transistor M7 and M8 forms Flipped voltage follower. The X terminal voltage is connected to the gate of M2 transistor which is controlled by source voltage of M7. So, transistor M1-M4 and from terminal a through M7 forms a level shifter, the voltages at X and Y terminals follow each other. The drain voltage of M7 is also used to control M8 and M11 transistors. So, the current flowing at X terminal is conveyed to Z terminal via M8. Here M7 and M8 ensure low resistance at X node. While Y input is at gate of M1 which gives very high input impedance at Y input. The circuits of proposed CCII in figure 2 was simulated using $0.18 \mu\text{m}$ CMOS technology with NMOS and PMOS threshold voltages of approximately 0.4 V and -0.39v . The Transistor aspect ratios are shown in Table 1. Bias voltage was $\pm 1 \text{ V}$, and bias current I_B was $70 \mu\text{A}$. First, its time response was evaluated by configuring the the circuit as unity-gain voltage amplifiers. In order to do so, ports X and Z were loaded with $15 \text{ k}\Omega$ resistances. The input voltage, a 100KHZ , 100 mVpp , sinusoid, was applied to the Y port Figure2 and the result is tabulated in Table:2 The AC small-signal frequency response for the circuit is subsequently obtained, using the same load resistors. The simple structure of Figure 2 has a unity bandwidth of 100 MHz , as expected. Table 2 also compares simulation results of the proposed circuit with low-voltage current conveyor reported in the literature[6]. The advantages in terms of power dissipation ,offset and compactness at reduced circuit complexity can be clearly evidenced.

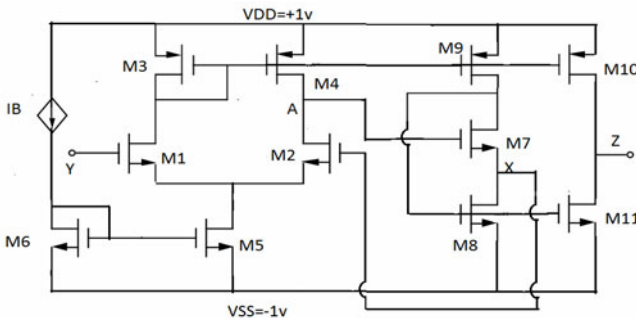


Fig. 2. OTA CCII based on differential pair

Table 1. Aspect Ratio for Fig:2

Transistor	W/L
M1,M2,M7	30u/.9u
M3,M4,M9,M10	15u/.9u
M5,M6,M8,M10	13.5u/2.7u

The small signal terminal impedances at X,Y and Z node is as follows.

$$Z_x = r_{o8} \parallel \left[\left(\frac{r_{o7}}{1 + g_{m7} r_{o7}} \right) + r_{o9} \right], Z_y = \gamma * W * L * c_{ox}, R_z = r_{o10} \parallel r_{o11},$$

The small signal ratios V_x/V_y and I_z/I_x is given as follows. The non-idealities of CCII shown in figure is measured using TSMC 180nm model parameters using spice simulation. The non ideal matrix is shown below.

$$\frac{V_x}{V_y} = \frac{r_{o7} g_{m7} (r_{o2}) g_{m1}}{1 + r_{o7} g_{m7} (1 + (r_{o2}) g_{m2})}, \frac{I_z}{I_x} = \frac{(g_{m11} * r_{o11} * r_{o8})}{(r_{o11} + r_{o10})(1 + g_{m8} r_{o8})}$$

$$\begin{bmatrix} I_y \\ V_x \\ I_z \end{bmatrix} = \begin{bmatrix} 0.33 \times 10^{-7} & 0 & 0.195 \times 10^{-3} \\ 1 & 1.8k & 0.310 \\ 0.09 & \pm 1 & 0.005 \times 10^{-3} \end{bmatrix} \begin{bmatrix} V_y \\ I_x \\ V_z \end{bmatrix}$$

3 Cascode OTA and FVF Based CCII

The problem with the above CCII is the output offset is a function of I_b , technology parameters and input voltage. The mathematically it is given by,

Table 2. Simulation Results For OTA CCII

Characteristic Parameters	Proposed FV mirror based Arch-1	A.J.L.Martin etal
Voltage Supply	+1V	1.5V
Power Cons.	0.58mwatt	0.75mwatt
I_z/I_x transfer BW	1.98Ghz	20Mhz
V_x/V_y transfer Bw	10Mhz	100Mhz
IBIAS	90uA	100uA
Offset	-23.72mv	300mv
I_z / I_x	1	1.1
V_x / V_y	1	1
Y Para. Imp.	6G Ω	80k
X Para. imp	1.4k Ω	10K Ω
Z Para. imp.	300k Ω	11K Ω
THD	1.21% @ 100Mhz	1%

$$V_{offset} = \sqrt{\frac{2I_b}{\beta}} [V_{DS1} - V_{DS2}]$$

The variation with I_b is plotted in figure:3 shows that the offset varies with biasing current I_b . From figure it is clear that as biasing current increases from 10mA to 170mA the output offset varies from -400mv to 0v. The circuit can be made independent from biasing current by adding one more pair of pMOS current mirror load on the top of the pMOS current mirror load in OTA based CCII in figure:2. The resulting CCII is shown figure:4 .

Table 3. Simulation Results For Cascoded OTA CCII

Characteristic Parameters	Proposed FV mirror based Arch-1	Cascoded OTA CCII	A.J.L.Martin etal
Voltage Supply	+1V	+/-1.25V	1.5V
Power Cons.	0.354mwatt	0.9mV	0.75mwatt
current transfer bandwidth	2Ghz	1.2Ghz	100Mhz
IBIAS	70uA	60uA	100uA
offset	Function of Ibias	Independent of Ibias	300mv
Iz / Ix	1	1	1.1
Vx / Vy	1	1	1

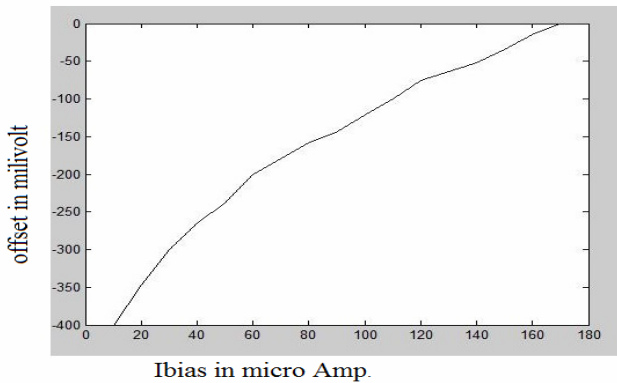


Fig. 3. Plot of biasing current dependency on output offset of OTA CCII in figure:2

5 Conclusion

Two novel CCII topologies are proposed simulated and compared with the present state of art design. The topologies are very compact, low power and wideband. The

second topology is a high precision with zero offset and independent of biasing current. The results are tabulated in Table-2 and 3.

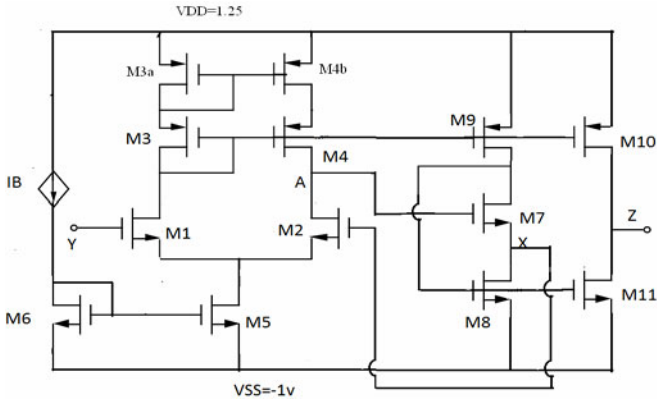


Fig. 4. OTA CCII based on differential pair

References

- [1] Rajput, S.S., Jamar, S.S.: Low voltage, low power, high performance current conveyors. In: Proc. ISCAS 2001, Sydney, pp. 1123–1726 (May 2001)
- [2] Ramirez-Angulo, J., Carvajal, R.G., Torraiba, A., Galan, A., Vega-Leal, A.P., Tombs, I.: The flipped voltage follower: a useful cell for low-voltage low-power circuit design. In: Proc. ISCAS 2002, Phoenix, AZ, pp. III 615–III 618 (May 2002)
- [3] Lopez Martin, A.J., Ramirez-Angulo, J., Carvajal: Low voltage Low power wideband CMOS current conveyors based on a flipped voltage follower. Proc. IEEE, III 801– III 804 (2003)
- [4] Lopez Martin, A.J., Angulo Sheetal Gupta, J.R., Carrvajal, R.G.: Comparison of conventional and New flipped voltage structure with increased input and output signal swing and current sourcing /Sinking capacity. IEEE proceedings (2005)

Generating Testcases for Concurrent Systems Using UML State Chart Diagram

Debashree Patnaik¹, Arup Abhinna Acharya¹, and Durga P. Mohapatra²

¹ School of Computer Engineering
KIIT University

debashree.patnaik@gmail.com, arupacharya.kiit@gmail.com

² Department of Computer Science and Engineering
National Institute of Technology, Rourkela
durga@nitrkl.ac.in

Abstract. Communication and concurrency are the major factors needed for the construction of a concurrent system. In concurrent environment systematic testing becomes a complex task. Generating test cases in concurrent environment is a difficult task because of arbitrary meddling of the concurrent thread. The interference of the concurrent thread may lead to a deadlock. In this paper we propose a methodology to generate the test cases for the conformance of deadlock in concurrent systems using UML State Chart Diagram. For system specification we have used the UML State Chart Diagram, from which event tree is generated. Along with a case study an algorithm is proposed to generate the test suite and confirm whether it is free from deadlock.

Keywords: Concurrency, State Chart Diagram, Deadlock, Event Tree, Test Sequences.

1 Introduction

The design of the concurrent software system is a complex activity leading to issues in testing. Environment is distributed means there are many machines distributed over the network which runs concurrently. In object-oriented system, objects become the means of communication. In concurrent execution environment we have multiple objects communicating concurrently [1]. In one class we can have many objects which can be active at the same time behaving differently from each other (multiple role playing). The concurrency of the objects has to be properly tested to establish the desired confidence in the functionality of any object. Testing the concurrent object-oriented system has to deal with the complexities arising from the physical distribution and parallel execution of the objects. An object has a definite state at a particular instant of time [1]. The state of the object changes due to the reaction of particular event. So event becomes the basic unit to observe the behavior of the object. States do not qualify to become the basic unit of observation because of the possibility of being complex and containing complex sub-states.

The main challenges in designing concurrent programs are ensuring the correct sequence of interactions between different computational processes, and coordinating the access of resources that are shared among processes [3].

The rest of the paper is organized as follows: Section-2 discusses deadlock issue in concurrent system. Section-3 describes deadlock analysis in UML State Chart. The proposed model is discussed in Section 4. Section-5 discusses the Future work. And finally Section-6 concludes the paper.

2 Deadlock: An Issue in Concurrent System

There are many challenges for a concurrent system. We have a number of different processes running together at a time. To synchronize and coordinate between the processes is very contending.

A deadlock is a situation where two or more competing actions are waiting for each other to finish and neither of the processes finishes [5]. It is a lock having no keys. Deadlock refers to a specific condition when two or more processes are each waiting for each other to release a resource, or more than two processes are waiting for resources in a circular chain. Deadlock is a common problem in multiprocessing where many processes share a specific type of mutually exclusive resource. Computers intended for the time-sharing and/or real-time markets are often equipped with a hardware lock guarantying exclusive access to processes, forcing serialized access. There is no general solution to avoid deadlocks.

3 Deadlock Analysis in UML State Chart Diagram

Deadlock is a common problem in multiprocessing where many processes share a specific type of mutually exclusive resource known as a software lock or soft lock. Sometimes, two or more transactions in their course of operation may attempt to access the same table or resource in a database. This prevents both the transactions from proceeding forward. This is called a Deadlock.

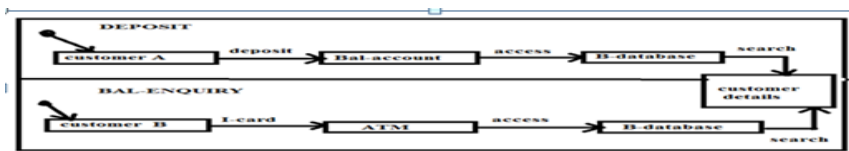


Fig. 1. Concurrency Leading To Deadlock In A Banking System

3.1 Deadlock in State Chart

In FIG.1 deadlocks is represented in terms of State Chart Diagram. *Deposit* and *Balance Enquiry* are the two states. *Customer A*, *Customer B*, *Bank account*, *Bank database*, *ATM*, *Customer Details* are the sub-states. When the event is fired from one state it reaches another state, it is called as transition. In deposit the Customer A is in initial state, it wants to carry out the deposit action. When certain sum of money is to be deposited by the customer, first we verify whether the account is a valid account. If the account is a valid then the bank database is accessed and then the customer details are searched. On the other hand in balance enquiry customer B wants to find a mini

statement of his account. He inserts the card into the ATM and types the pin. The pin number is verified. If the pin code is OK then the database is accessed and the customer details are searched. There may be an instant where the *Deposit* and the *Balance Enquiry* are trying to access the same customer details at same instant of time. In this case deadlock occurs; the two actions try to take place concurrently.

4 Proposed Method

In this paper the authors have proposed a model to generate Test cases, for a concurrent system. The main issue for a concurrent system is deadlock. The concurrent systems are modeled by using UML State Chart Diagram from which event tree are generated. Finally event trees are traversed to generate the test cases.

4.1 Creation of Event Tree

Using **Chow’s Algorithm [1]** the authors have generated the event tree for individual State Charts Diagram. In Fig.2 *Deposit* and *Balance Enquiry* are the generated event trees from the State Chart Diagram. Even though both sub states run concurrently they are represented in different graphs [6]. To generate the test sequence we combine both the event tree, and generate a graph for the whole system and then find the test Suite for the scenario. In the system graph, we have to take all possible scenarios and generate the test cases.

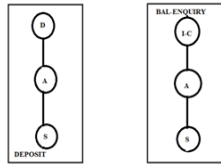


Fig. 2. Event tree of two sub-states (Deposit and Balance Enquiry) D-DEPOSIT, A-ACCESS, S-SEARCH, I-C-INSERTCARD

4.2 Traversing Event Graph

With the backtracking algorithm we make an effort to visit every node. All nodes are initialized with value 0, when a node is discovered the value of the node changes to 1. If the node is visited then the value of the node is changed to 2. If all the nodes are traversed then finally we get a graph with value 2. If there is deadlock as shown in Fig.3, then the transactions are rolled back.

We have applied backtracking algorithm for traversal of the tree and generation of Test Cases and Test Sequences.

V= {account, withd, aaccess, baccess, search,p_val,n_bal,open,close,found, n_found}

Node status= {0, 1, 2}

0 – Nodes that are Unexplored, 1 - Nodes that are discovered, 2 – Nodes that are explored.

ALGORITHM: BACKTRACKING_ALGORITHM

INPUT: Event Tree

OUTPUT: Test Sequences

Step1: Start

Step2: For each u that belongs to V [G]

Value [u] = 0

Pie[u] = nil // records the event by setting u's predecessor field pie[u]

time \leftarrow 0

// all the vertices are initialized to value 0 and their pie fields to nil. Time global counter) is set to 0.

Step3: For each u

If (value [u] == 0)

DFS visit (u) // Depth First Search

Value[u] = 1

time \leftarrow time + 1

//When a vertex with value 0 is found it visits using the DFS visit. Vertex is discovered and set to value 1, global counter is incremented.

Step4: Find each vertex V which is adjacent to u.

D[v] \leftarrow u

If (value [u] == 0) // (reaches the last node)

Pie[v] \leftarrow u

DFS visit (v)

Value[u] \leftarrow 2

time \leftarrow time + 1

//The discovered node are explored and while leaving the vertex the value of the vertex are set to 2, with the increment of the global counter).

Step5: Stop

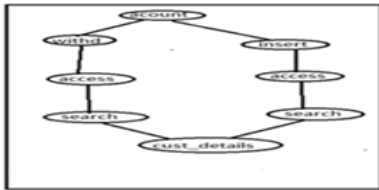


Fig. 3. Event Graph representing Deadlock

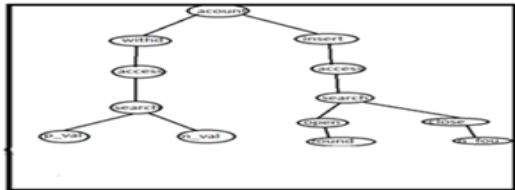


Fig. 4. Event graph free from deadlock

Account(Root node)- as it enhances a situation to enter into the State Chart Diagram, *Withd*-Withdrawal, *insert*-inserting the ATM card, *access*-accessing the bank database, *Search*-Searching the required account number, *cust_details*- accessing the same customer details, *p_val*- permitted value, *n_val*-no balance, *open-account* is active, *close-account* is inactive, *found-customer* details is found, *n_fou*-account is not found.

4.3 Generation of Test Sequences

The test sequences are generated by the backtracking traversal algorithm. In Fig.4 the system is free from deadlock and the test sequences generated are given in table 1. In

Fig. 3 the system shows a deadlock event as at same instant of time the access to customer details is required.

Table 1. Test Sequences Generated for Conformance of Deadlock

Test Sequence #1	Test Sequence #2	Test Sequence #3	Test Sequence #4
withd aaccess search p_val	withd aaccess search n_val	insert baccess search open found	insert baccess search close n_fou

4.4 Generation of Test Cases

The test cases generated using the proposed methodology is listed in Table 2. As deadlock is a major concern in concurrent systems, only a subset of test cases are generated are shown here to address this issue.

Table 2. Generated Test Cases

Step #	Action Performed	Expected Results
1	Card is inserted in the ATM machine	The ATM prompts us to enter the 4 digit PIN/Transaction code
2	Key in the 4 digit PIN/Transaction code	On entering the PIN code the System starts a action or transaction for checking the Customer details and account details
3	System searches for the Customer details attached to this card	Correct Customer is found from the customer table
4	Once the PIN is validated. The system prompts for the required action, we select the "Balance Inquire"	On selecting Balance Inquire the system searches for the Account details in the Account Details table
5	The system checks for the status of the "Account Details" table	<p>If another Banking transaction is trying to simultaneously access the "Account Details" table, a Deadlock situation is faced by both of the transactions and our transaction is ends abruptly giving a message as "Your transaction cannot be processed currently. Please try again later".</p> <p>If the "Account Status" table is available/free for the transaction then the "Account Search" action is initiated.</p>

5 Future Work

The deadlock analysis can be determined with the help of a mathematical tool called as petrinets. Both static and dynamic analysis can be determined by petrinets concurrently. Graphs can be constructed avoiding the interleaving. The authors intend to

analyze and explore on the idea we have established in a wider range of the distributed systems.

6 Conclusion

Problems that arise from concurrency and deadlock are discussed in context to distributed environment. The paper emphasizes on testing the behavioral aspects of the distributed objects using events to stimulate the object's behavior. A method of deadlock analysis is presented through UML State chart diagrams. From the State Chart diagram an event tree is generated with event as node. The event tree is further transformed to an event graph generating the test cases and the test sequences. For the traversal of the event graph the authors have implemented the back-tacking algorithm.

References

- [1] Adnan Bader, A.S.M., Sajeev, S.R.: Testing concurrency and communication in distributed objects
- [2] Asaadi, H.R., Khosravi, R., Mousavi, M., Noroozi, N.: Towards Model-Based Testing of Electronic Funds Transfer Systems
- [3] Hessel, A., Larsen, K.G., Mikucionis, M., Nielsen, B., Pettersson, P., Skou, A.: Testing real-time systems using UPPAAL. In: Hierons, R.M., Bowen, J.P., Harman, M. (eds.) FORTEST. LNCS, vol. 4949, pp. 77–117. Springer, Heidelberg (2008)
- [4] Hierons, R.M., Bogdanov, K., Bowen, J.P., Cleaveland, R., Derrick, J., Dick, J., Gheorghe, M., Harman, M., Kapoor, K., Krause, P., Lüttgen, G., Simons, A.J.H., Vilkomir, S.A., Woodward, M.R., Zedan, H.: Using formal specifications to support testing. *ACM Computing Surveys* 41(2) (2009)
- [5] Koopman, P.W.M., Plasmeijer, R.: Testing reactive systems with GAST. In: Post-Proceedings of TFP 2003, Intellect, pp. 111–129 (2003)
- [6] Mikucionis, M., Nielsen, B., Larsen, K.G.: Real-time system testing on-the-fly. In: Proceedings of NWPT, pp. 36–38 (2003)

Intelligent Agent Based Resource Sharing in Grid Computing

V.V. Srinivas¹ and V.V. Varadhan²

¹ Department of Computer Science and Engineering,
National Institute of Technology – Tiruchirappalli
srinivas15j1988@gmail.com

² Department of Information Technology,
Madras Institute of Technology - Chennai
varadhan3n90@gmail.com

Abstract. Most of the resource present in grid are underutilized these days. Therefore one of the most important issue is the best utilization of grid resource based on users request. The architecture of intelligent agent proposed to handle this issue consists of four main parts. We discuss the need and functionality of such an agent and propose a solution for resource sharing which satisfies problems faced by today's grid. A J2EE based solution is developed as a proof of concept for the proposed technique. This paper addresses issues such as resource discovery, performance, security and decentralized resource sharing which are of concern in current grid environment.

Keywords: grid, resource sharing, intelligent agent, decentralization.

1 Introduction

Grid computing is distributed, large-scale cluster computing, as well as a form of network-distributed parallel processing. Each computer present in grid has computational power and resources such as memory, printer etc., which are underutilized. In order to utilize resources and provide service to customers resource sharing was introduced. Resource sharing [1] provides access to a particular resource on a computer to be accessed by clients on grid. The need for resource sharing arises in case of complex mathematical modeling and simulations like the network simulation or simulation of automatic test pattern algorithms, virtual supercomputing or DNA mapping.

Resource sharing involves three main process namely: resource discovery, resource management and resource allocation. Resource discovery is finding resources available in grid. This problem is solved in tools such as Globus and Condor Matchmaker [2]. Resource management involves collecting resource. The challenge involved is finding the right quantity of resource [3].

1.1 Contribution

There are a number of ongoing research in the field of grid computing most of them trying to address some of the challenges faced in grid environment. This paper

addresses aspects like: resource discovery in grid, security, decentralized resource sharing, scheduling of resources and threshold based resource allocation.

The rest of this paper is organized as follows: Section 2 discusses ongoing research followed by section 3 discussing proposed architecture. Section 4 describes experimental results. In section 5, the paper deals with application and advantages followed by future research and conclusion.

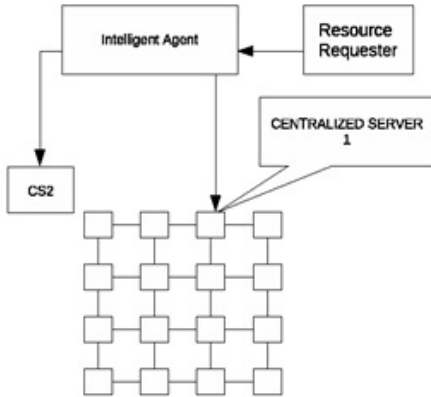


Fig. 1. Resource discovery

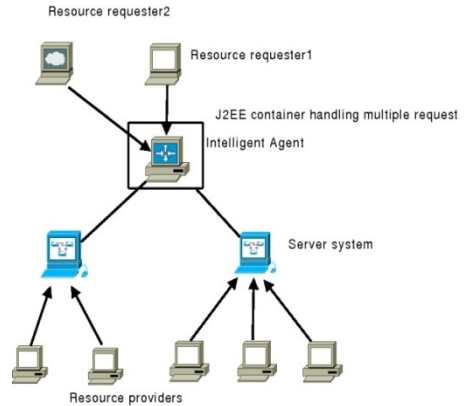


Fig. 2. Grid model of entire system

2 Ongoing Research

2.1 Globus Toolkit

Globus is an open source toolkit that is used for construction of grids. It provides access to resource present within the grid located in any geographical area [5]. Globus uses GSI (Grid Security Infrastructure), GRAM (Grid Resource Allocation and Management) and MDS (Monitoring and Discovery Services) for resource management and discovery. Globus provides a remote front end to multiple batch systems. Our paper tries to preserve all the advantages of Globus and eliminates all complex configuration and installation of number of tools.

2.2 Meta-broker Architecture

The meta-broker architecture focuses on how to allocate a particular resource present in some other network or grid to a user requesting for resource [6]. Previous works deals with MESS (Multi Engine Search Services) and ISS (Internet Search Service) based on CORBA [7] and meta-broker architecture for management of grid resources [8].

2.3 Negotiation Algorithms

This work mainly deals with negotiation protocols between the client and provider. The key focus is on contracts [9]. One important algorithm used is G-Negotiation algorithm. Our paper discusses a simple mechanism for secure communication.

3 Proposed Architecture

The entire work is split up into several modules which are discussed in detail. First is the *server module* in which client registers to provide resource. The client is provided with an address. The details registered include resource type, amount of resource and time duration when the resource would be available. *Resource discovery* [4] module is used to keep check on parameters such as processor time, print queue, system threads, disk queue length and cpu usage. The values are obtained from performance logs and alerts. The retrieved values are stored in .csv format. These obtained values

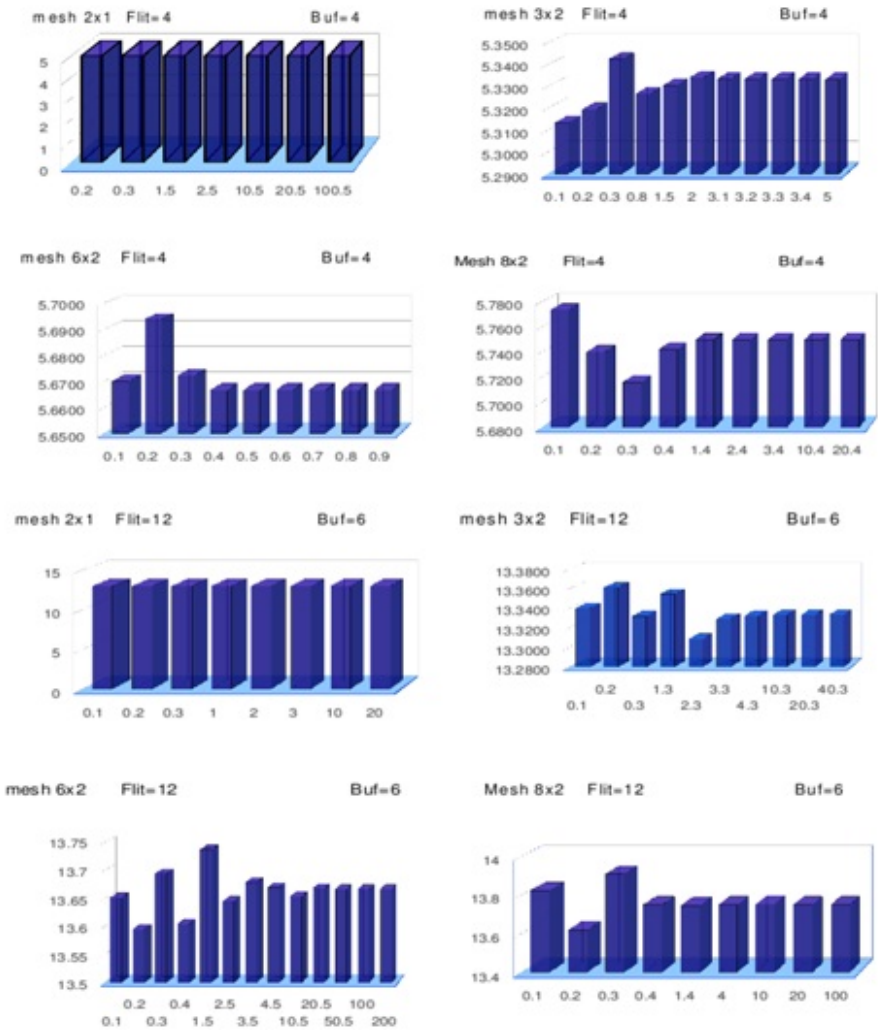


Fig. 3. Four program instances run on 2, 3, 6 and 8 nodes. X axis represents injection rate and Y axis represents latency. Buffer size are taken as 4 and 6.

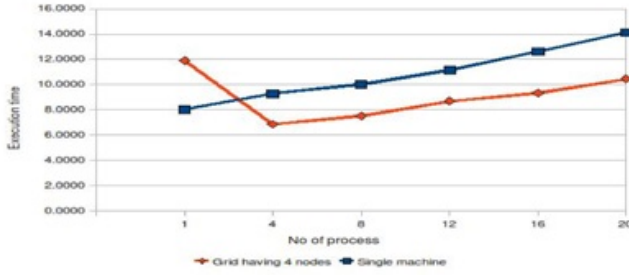


Fig. 4. Performance analysis in terms of execution time

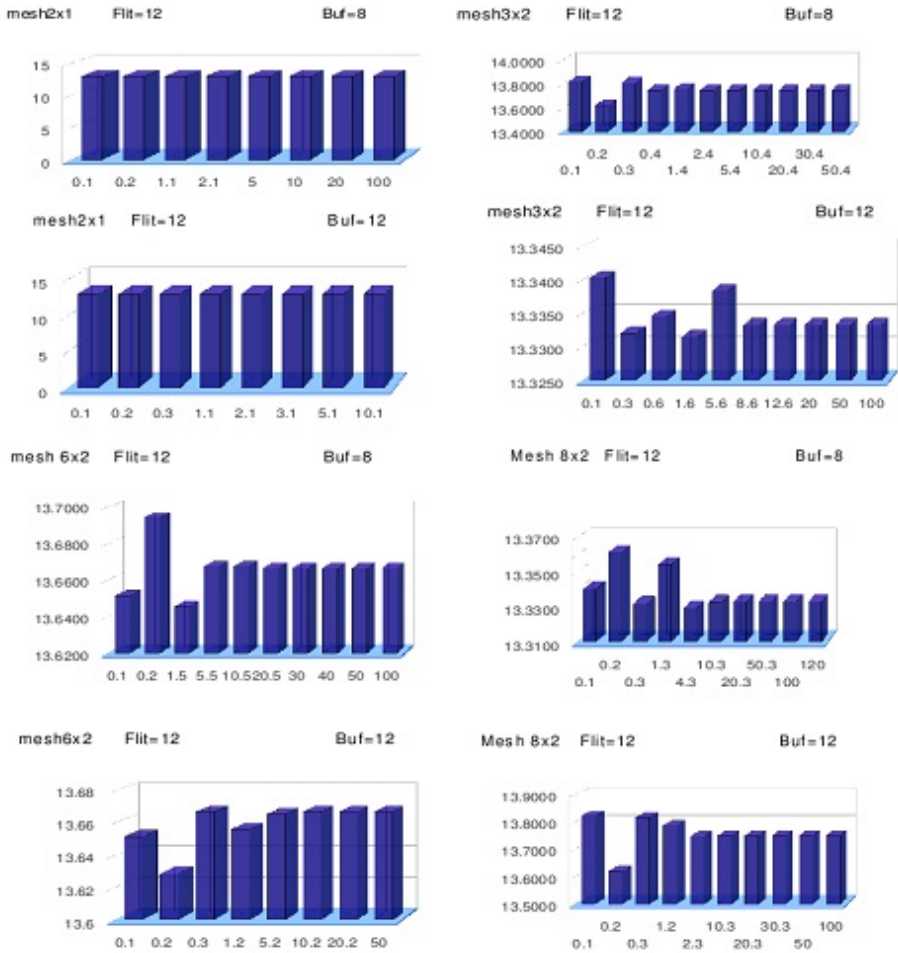


Fig. 5. Four program instances run on 2, 3, 6 and 8 nodes with buffer size are taken as 8 and 12

are compared with threshold values. The third is the *intelligent agent module* which reads the files obtained from the resource discovery module running on each machine. The resource files are set to refresh after t_k time. The resource requester requests the agent for a particular resource. The intelligent agent searches and retrieves the various resources it has from various clients. Once the right resource is obtained, a secret key is transmitted to the resource requester and provider. Along with the secret key, the resource requester receives the providers address and vice-versa. This ensures security.

4 Experimental Results

The performance was evaluated between the number of process and execution time. We took the number of resource providers in the grid to be 4 and the number of process was taken as multiples of 4. A graph was plotted for the process running on single system to the process split among the four systems and the resultant graph is shown in Fig. 4.

5 Conclusion

Our paper discusses a solution for resource sharing at the same time preserving features like security, authentication, resource discovery and decentralization. This paper defines a simple solution to implement intelligent agent in grid environment.

References

1. Cruz-Perez, F.A., Ortigoza-Guerrero, L.: Equal resource sharing allocation with QoS differentiation conversational services in wireless communication networks. *IEEE Proceedings Communications*, 150, 391–398 (2003)
2. De Smet, A.: Computer Science Department. University of Wisconsin Madison, <http://www.cs.wisc.edu/condor>
3. Li, Y., Wolf, L.: Adaptive Resource Management in Active Nodes. In: 8th IEEE International Symposium on Computer and Communication (2003)
4. Giovanni, A., Massimo, C., Italo, E., Maria, M., Silvia, M.: Resource and Service Discovery in the iGrid Information Service. In: Gervasi, O., Gavrilova, M.L., Kumar, V., Laganá, A., Lee, H.P., Mun, Y., Taniar, D., Tan, C.J.K. (eds.) ICCSA 2005. LNCS, vol. 3482, pp. 1–9. Springer, Heidelberg (2005)
5. Ian, F., Carl, F.: The Globus project: A Status Report. In: Heterogeneous Computing Workshop, pp. 4–18 (1998)
6. Kertesz, A., Kaeushk, P.: Meta-broker for future generation grids: A new approach for high level inter operable resource management. In: CoreGrid workshop (2007)
7. Yue-Shan, C., Hsin-Chun, H., Shyan-Ming, Y., Lo, W.: An agent based search engine based on Internet search service on CORBA. In: Proceedings of International Symposium on Distributed Objects and Applications, pp. 26–33 (1999)
8. Kertesz, A., Kacsuk, P.: Grid Interoperability Solutions in Grid Resource Management. *Systems Journal* 3, 131–141 (2009)
9. Antoine, P., Phelipp, W., Oliver, W., Wolfgang, Z., Dynamic, S.L.A.: negotiation based on WS agreement. In: Core Grid Technical Report TR-0082 (2007)
10. Karl, W.: The Management of Change among Loosely Coupled Elements. In: Making Sense of the Organization (1982)

Wideband Miniaturized Patch Antenna Design and Comparative Analysis

Sanket Patel^{1,*}, Yogeshwar Kosta², Himanshu Soni¹, and Shobhit Patel²

¹ G.H. Patel College of Engineering and Technology, V.V. Nagar, Gujarat, India
sanket_patel_2020@yahoo.co.in

² Charotar University of Science and Technology, Changa, Gujarat, India

Abstract. A novel kind of miniaturized wideband patch antenna is designed and comparative analysis is presented. The antenna is having size of 2.1cm X 2.1cm X 1.25cm and patch is having area of 153mm². Single or dual bands frequency response can be obtained by varying the feed location, height of the substrate and the geometric specifications of the antenna. This proposed antenna is very small in size although it provides good impedance behavior, return loss S₁₁ behavior and VSWR which is very much nearer to 1 at each band. Bandwidths up to more than 56% can be obtained. Far-field radiation pattern and field distributions on the coaxial probe feed patch have been analyzed.

Keywords: Dual Bands, Return Loss, VSWR.

1 Introduction

Microstrip antennas have gained extensive applications in recent times due to their light weight, small size, easy reproduction and integration ability with the circuitry [1-10]. There is a tremendous growth in demand for wireless RF systems in applications such as local area networks, point-to-point communications and applications in medical and industrial sectors. By time patch structure is modified to have application specific resonating frequencies and to have higher gain and bandwidth response. [11-13]. Rectangular, circular and triangular patch geometries are the most extensively analyzed antenna geometries in recent years and now these geometries are modified to improve their performance [14-16]. Lu [15] analyzed a circular patch antenna and its arrays with a pair of L-shaped slots for broadband dual-frequency operation. Wong and Hsu [16] applied a V shaped slot in an equilateral triangular microstrip antenna. In this article, novel patch antenna is designed and comparative analysis is done with the normal patch antenna. By varying proposed antenna parameters one can have single or dual frequency bands of operation. Antenna provides enhanced bandwidth. Comparison is done based on the geometric specifications as well as in terms of responses.

* Sanket S. Patel is currently pursuing M.E. in Communication Engineering from G.H. Patel College of Engineering and Technology, Gujarat Technological University.

2 Antenna Design

The geometry of the antenna is shown in Fig.1. The antenna parameters are also given in Fig.1. The antenna is mounted on a duroid (tm) substrate and fed by a coaxial transmission line.

Simulations were performed using HFSS™[17]. HFSS (High Frequency Structure Simulator), is the industry-standard simulation tool for 3D full-wave electromagnetic field simulation. HFSS provides E and H-fields, currents, S-parameters and near and far radiated field results. It integrates simulation, visualization, solid modeling, and automation. Ansoft HFSS employs the Finite Element Method (FEM) for EM simulation by developing/ implementing technologies such as tangential vector finite elements and adaptive meshing.

In this proposed antenna patch and ground plane are made of copper having relative permittivity as 1. The substrate material duroid (tm) has relative permittivity as 2.2 and dielectric loss tangent 0.0009.

Convergence was tested for each case separately in terms of evaluating S11 (dB) at selected frequencies for a number of times. Once convergence was obtained, simulations were conducted in order to obtain swept frequency response extending from 1 to 10 GHz. The swept response gave the S11, which was used to calculate the VSWR. After that radiation pattern was computed. Here three different designs are presented and analyzed.

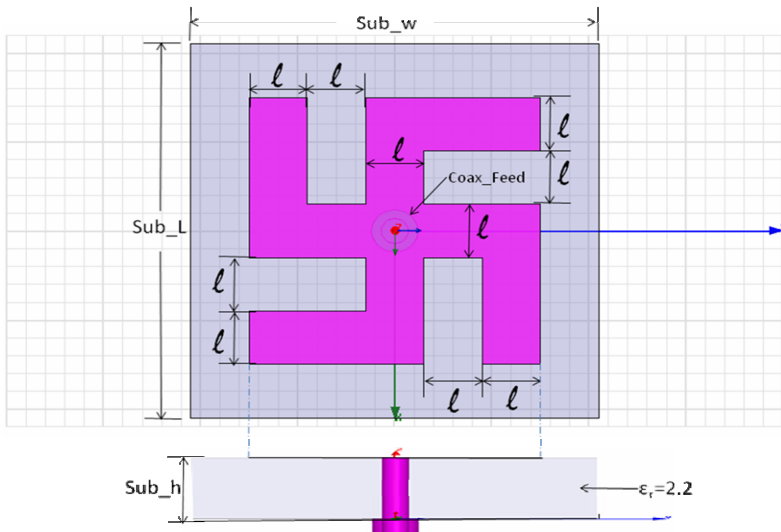


Fig. 1. Antenna Geometrical Configuration. Parameters are: $sub_w=21mm$, $sub_L=21mm$, $sub_h=12.5mm$, $l=3mm$.

3 Comparative Analysis

The tabular analysis is given in the Table 1. Table compares the proposed miniaturized antenna with the conventional design. Comparison is done considering geometrical specifications. Table narrates that the dimensions of the proposed antenna are very much smaller than the conventional antenna. Proposed antenna has the substrate area is about 20 times smaller than the conventional antenna. It should be noted that patch area is about 8 times smaller than the conventional design results in low conducting material requirement and so low cost. Volume of the proposed antenna is 5 times lesser than the conventional patch antenna so proposed antenna can be encapsulated even in small size instruments. Substrate dielectric material used for the both the antennas are having same relative permittivity as 2.2 and same dielectric loss tangent 0.0009.

Considering specific feed locations the return loss S_{11} behavior of both the antennas are compared. Conventional antenna provides single band of operation and %bandwidth at 10dB is 2.54% (60MHz). While in second proposed antenna, dual band of operation is seen respective to feed locations. For feed location (0.7mm,0.7mm) achieved dual bands have the %BW's of 40.30%(730MHz) and 5.43%(500MHz). For feed location (2,0) achieved dual bands have the %BW's of 40.30%(730MHz) and 7.26%(660MHz). So the antenna can be used for UWB antenna. Corresponding VSWR plots are compared, and VSWR much nearer to 1 is achieved at each frequency band. Far Field Radiation Pattern for specific angles is compared. Pattern depicts the directivity of the antennas. Total Gain (3D Polar Plot) is generated and is matching with the directivity 2D plots as gain is directly proportional to directivity. It should be noted that here in both antennas the material used for the patch and for the ground plane is the copper. Mesh refinement and E-field distribution on the patch for both the antennas are also depicted in Table 1.

Table 1. Comparative analysis of conventional antenna and proposed antenna

Parameter	Conventional Design	Proposed design
Top View		
Front View		
Dimensions	Sub_L=100mm, sub_w=90mm, L=40mm, w=30mm, sub_h=3.2mm	Sub_L=21mm, sub_w=21mm, sub_h=12.5mm, l=3mm
Ground Plane area	9000mm ²	441mm ²

Table 1 (continued)

Patch Area	1200mm ²	153mm ²
Volume of Antenna	28.8cm ³	5.5125cm ³
Dielectric Material of the substrate	Material: Rogers RT/duroid 5880 (tm) Relative Permittivity 2.2 Loss tangent 0.0009	Material: duroid (tm) Relative Permittivity 2.2 Loss tangent 0.0009
Feed Location (x,y)	(-0.5mm,0mm)	Case:1 (0.7mm,0.7mm) Case:2 (2mm,0mm)
Return Loss S11		
VSWR		
Far Field Radiation Pattern (Directivity)		
Total Gain (Polar Plot) G=kD		
Patch/Ground Plane Material	Copper	Copper

Table 1 (continued)

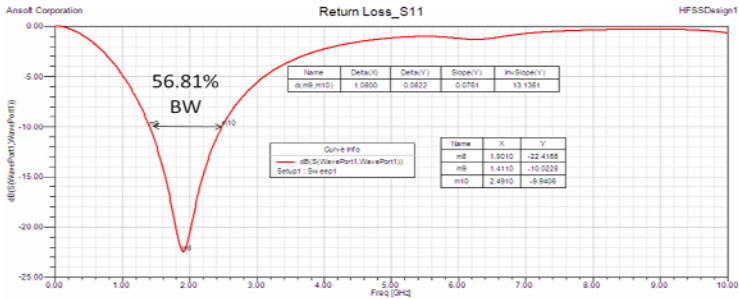
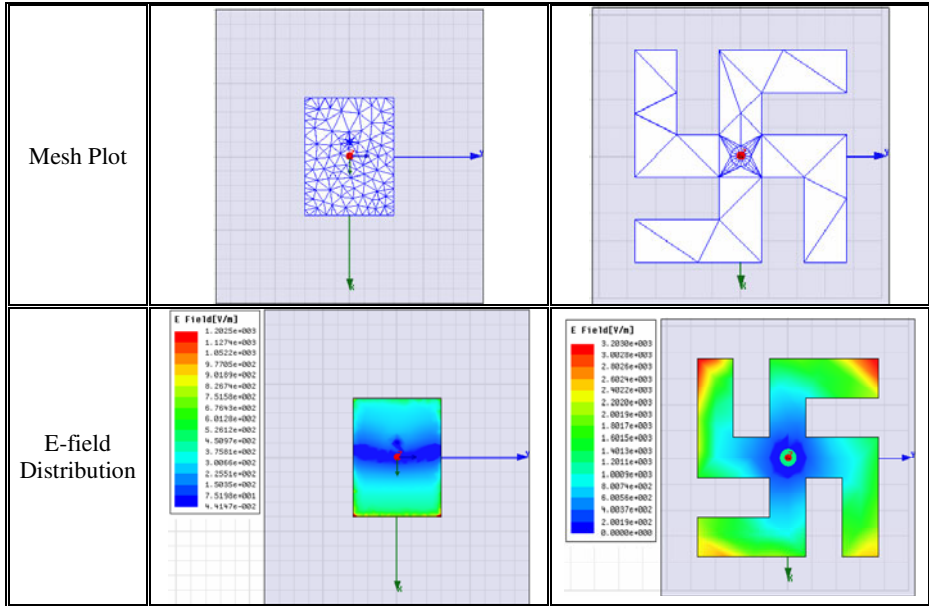


Fig. 2. Return Loss and VSWR plots for proposed wideband antenna

The case when the feed location of the proposed antenna is at (0mm,0mm) position and the substrate height is 10mm then the return loss behavior and the VSWR plot is shown in the Fig. 2. Resonant frequency is at 1.90GHz, which can be used for GSM 1900 transmit frequency (uplink) band (1850 MHz –1910MHz). Bandwidth achieved at 10dB is 56.81% which is equivalent to 1.088GHz of band span as shown in Fig. 2.

4 Conclusion

Proposed antenna provides large bandwidth, good return loss and VSWR behavior. It is very much small in size compare to the conventional antenna. This proposed

antenna is the efficient design in terms of bandwidth, number of bands of operation and the cost.

References

1. Carver, K.R., Mink, J.W.: Microstrip antenna technology. *IEEE Trans. Antennas Propag.* AP-29, 2–24 (1981)
2. Mailloux, R.J., McIlvenna, J.F., Kemweis, N.P.: Microstrip array technology. *IEEE Trans. Antennas Propag.* AP-29, 25–37 (1981)
3. Bahl, I.J., Bhartia, P.: *Microstrip Antennas*. Artech House, Dedham, MA (1980)
4. James, J.R., Hall, P.S., Wood, C.: *Microstrip Antenna Theory and Design*
5. James, J.R., Hall, P.S.: *Handbook of Microstrip Antennas*. Peter Peregrinus, London, U.K (1989)
6. Richards, W.F., Lo, Y.T., Harrison, D.: An improved theory for microstrip antennas and applications. *IEEE Trans. Antennas Propag.* AP-29, 38–46 (1981)
7. Pozar, D.M.: Considerations for millimeter wave printed antennas. *IEEE Trans. Antennas Propag.* AP-31, 740–747 (1983)
8. Schaubert, D.H., Pozar, D.M., Adrian, A.: Effect of microstrip antenna substrate thickness and permittivity: Comparison of theories and experiment. *IEEE Trans. Antennas Propag.* 37, 677–682 (1989)
9. Pozar, D.M.: *Microstrip Antennas*. *Proceedings of the IEEE* 80(1) (January 1992)
10. Garg, R., Bhartia, P., Bahl, I.J., Ittipiboon, A.: *Microstrip antenna design handbook*. Artech House, New York (2001)
11. Carver, K.R.: Practical analytical techniques for the microstrip antenna. In: *Proc. Workshop Printed Circ. Antennas*, pp. 7.1–7.20. New Mexico State University (1979)
12. Binu Paul, S., Mridula, C.K., Aanandan, P., Mohanan: A new microstrip patch antenna for mobile communications and Bluetooth applications. *Microwave and Opt. Technol. Lett.* 33(4), 285–286 (2002)
13. Tiwari, V.K., Kimothi, A., Bhatnagar, D., Saini, J.S., Saxena, V.K.: Theoretical and experimental investigation of circular sector microstrip antenna. *Indian J. Radio and Space Physics* 35, 206–211 (2006)
14. Bhardwaj, D., Bhatnagar, D., Sancheti, S., Soni, B.: Design of square patch antenna with a notch on FR4 substrate. *IET Microwaves, Antennas & Propagation* 2(8), 880–885 (2008)
15. Lu, J.H.: Broadband Dual-Frequency Operation of Circular Patch Antennas and Arrays with a Pair of L-Shaped Slots. *IEEE Transactions on Antennas and Propagation* 51(5), 1018–1023 (2003)
16. Wong, K.L., Hsu, W.S.: Broadband triangular microstrip antenna with V-shaped slot. *Electron. Lett.* 33(25), 2085–2087 (1997)
17. Ansoft HFSS, Ansoft Corporation, <http://www.ansoft.co.jp/hfss.htm>

Predicting Number of Zombies in a DDoS Attack Using ANN Based Scheme

B.B. Gupta^{1,2,*}, R.C. Joshi¹, M. Misra¹, A. Jain², S. Juyal²,
R. Prabhakar², and A.K. Singh²

¹ Department of Electronics and Computer Engineering, Indian Institute of Technology
Roorkee, Roorkee, India
gupta.brij@gmail.com

² Department of Computer Science and Engineering, Graphic Era University, Dehradun, India

Abstract. Anomaly based DDoS detection systems construct profile of the traffic normally seen in the network, and identify anomalies whenever traffic deviate from normal profile beyond a threshold. This deviation in traffic beyond threshold is used in the past for DDoS detection but not for finding zombies. In this paper, two layer feed forward neural networks of different sizes are used to estimate number of zombies involved in a DDoS attack. The sample data used to train the feed forward neural networks is generated using NS-2 network simulator running on Linux platform. The generated sample data is divided into training data and test data and MSE is used to compare the performance of various feed forward neural networks. Various sizes of feed forward networks are compared for their estimation performance. The generalization capacity of the trained network is promising and the network is able to predict number of zombies involved in a DDoS attack with very less test error.

1 Introduction

Denial of service (DoS) attacks and more particularly the distributed ones (DDoS) are one of the latest threat and pose a grave danger to users, organizations and infrastructures of the Internet. A DDoS attacker attempts to disrupt a target, in most cases a web server, by flooding it with illegitimate packets, usurping its bandwidth and overtaxing it to prevent legitimate inquiries from getting through [1,2]. Anomaly based DDoS detection systems construct profile of the traffic normally seen in the network, and identify anomalies whenever traffic deviate from normal profile beyond a threshold [3]. This extend of deviation is normally not utilized. Therefore, this extends of deviation from detection threshold and feed forward neural networks [4-6] are used to predict number of zombies. A real time estimation of the number of zombies in DDoS scenario is helpful to suppress the effect of attack by choosing predicted number of most suspicious attack sources for either filtering or rate limiting. We have assumed that zombies have not spoof header information of out going packets. Moore et. al [7] have already made a similar kind of attempt, in which they have used backscatter

* Corresponding author.

analysis to estimate number of spoofed addresses involved in DDoS attack. This is an offline analysis based on unsolicited responses.

Our objective is to find the relationship between number of zombies involved in a flooding DDoS attack and deviation in sample entropy. In order to predict number of zombies, feed forward neural network is used. To measure the performance of the proposed approach, we have calculated mean square error (MSE) and test error. Training and test data are generated using simulation. Internet type topologies used for simulation are generated using Transit-Stub model of GT-ITM topology generator [8]. NS-2 network simulator [9] on Linux platform is used as simulation test bed for launching DDoS attacks with varied number of zombies and the data collected are used to train the neural network. In our simulation experiments, attack traffic rate is fixed to 25Mbps in total; therefore, mean attack rate per zombie is varied from 0.25Mbps to 2.5Mbps and total number of zombie machines range between 10 and 100 to generate attack traffic. Various sizes of feed forward neural networks are compared for their estimation performance. The result obtained is very promising as we are able to predict number of zombies involved in DDoS attack effectively.

The remainder of the paper is organized as follows. Section 2 contains overview of artificial neural network (ANN). Intended detection scheme is described in section 3. Section 4 contains simulation results and discussion. Finally, Section 5 concludes the paper.

2 Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) [4-6] is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true for ANNs as well. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and answer "what if" questions.

3 Detection of Attacks

Here, we will discuss propose detection system that is part of access router or can belong to separate unit that interact with access router to detect attack traffic. Entropy based DDoS scheme [10] is used to construct profile of the traffic normally seen in the network, and identify anomalies whenever traffic goes out of profile. A metric that

captures the degree of dispersal or concentration of a distribution is sample entropy. Sample entropy $H(X)$ is

$$H(X) = -\sum_{i=1}^N p_i \log_2(p_i) \quad (1)$$

where p_i is n_i/S . Here n_i represent total number of bytes arrivals for a flow i in $\{t - \Delta, t\}$ and $S = \sum_{i=1}^N n_i, i = 1, 2, \dots, N$. The value of sample entropy lies in the range $0 - \log_2 N$.

To detect the attack, the value of $H_c(X)$ is calculated in time window Δ continuously; whenever there is appreciable deviation from $X_n(X)$, various types of DDoS attacks are detected. $H_c(X)$, and $X_n(X)$ gives Entropy at the time of detection of attack and Entropy value for normal profile respectively.

4 Results and Discussion

4.1 Training Data Generation

Neural network has to be trained by giving sample inputs and corresponding output values and a training algorithm will adjust the connection weight and bias values until a minimum error or other stopping criteria is reached. The training data has to be taken carefully to consider the complete input range. Normalization and other pre-processing of the data improve the training performance.

In our paper, in order to predict number of zombies (\hat{Y}) from deviation ($H_c - H_n$) in entropy value, training data samples are generated using simulation experiments in NS-2 network simulator. Simulation experiments are done at the same attack strength 25Mbps in total and varying number of zombies from 10-100 with increment of 5 zombies i.e. mean attack rate per zombie from 0.25Mbps-2.5Mbps. The data obtained is divided into two parts, 78.95% of the data values are used for training. The remaining data values which are selected randomly are used for testing.

4.2 Network Training

For the prediction of the number of zombies in a DDOS attack, three feed forward neural networks have been tested. The feed forward networks used have different sizes. The size of a network refers to the number of layers and the number of neurons in each layer. There is no direct method of deciding the size of a network for a given problem and one has to use experience or trial error method. In general, when a network is large, the complexity of the function that it can approximate will also increase. But as the network size increase, both training time and its implementation cost increase and hence optimum network size has to be selected for a given problem. For the current

Table 1. Training results of various feed forward networks

<i>Network used</i>	<i>Network size</i>	<i>Number of Epochs</i>	<i>MSE in training</i>
2 layer network	5-1	400	6.86
	10-1	400	0.36
	15-1	400	0.0025

problem, two layer feed forward networks with 5, 10 and 15 neurons are selected. The training algorithm used is the Levenberg-Marquardt back propagation algorithm of MATLAB's neural network toolbox. The training results are given in Table 1.

4.3 Network Testing

Table 2 shows the result of the testing of the networks using the test data values.

Table 2. Test results of various feed forward networks

<i>Network used</i>	<i>Network size</i>	<i>MSE in Testing</i>
2 layer network	5-1	2.91
	10-1	2.59
	15-1	3.14

From the result of table 1, we can see that the MSE in training decreases linearly as the network size increase. This is as expected. But in table 2, we can see that in spite of the smaller MSE in training and the increase in network size, the test result for the feed forward network having 15 hidden layer neurons is greater than the networks having 5 and 10 neurons. One reason for this is, for a good network performance the ration of number of tunable parameters to that of training data size has to be very small and in here network size has increased but training data size is the same. For the last network, the number of tunable parameters is 31 and ration is 1.63. And because of this over fitting has occurred and the generalization performance of the last network is poor though it has good training performance. The training performance is measured using the mean square error (MSE). MSE is the difference between the target and the neural network's actual output. So, the best MSE is the closest to 0. If MSE is 0, this indicates neural network's output is equal to the target which is the best situation. Number of zombies of the individual networks can be compared with actual number of zombies for each test data values and the results are given in figure 1, 2 and 3. The simulation results show that two layer feed forward networks with 10 neurons performs best. Two layer feed forward networks with 10 neurons is able to predict number of zombies involved in a DDoS attack with very less error.

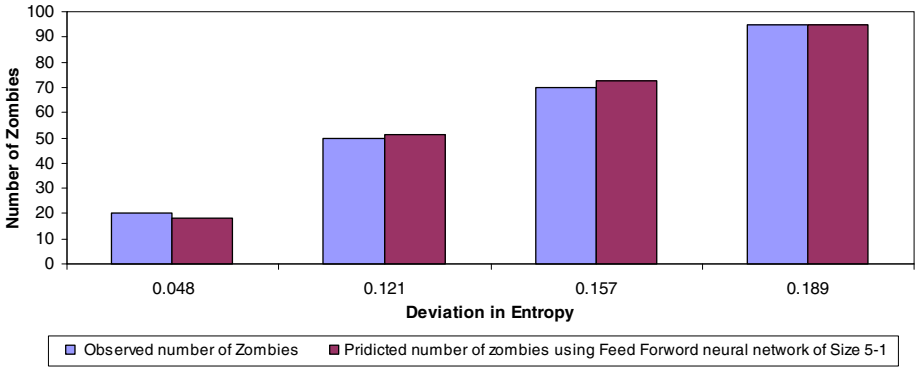


Fig. 1. Comparison between actual number of zombies and predicted number of zombies using feed forward neural network of size 5-1

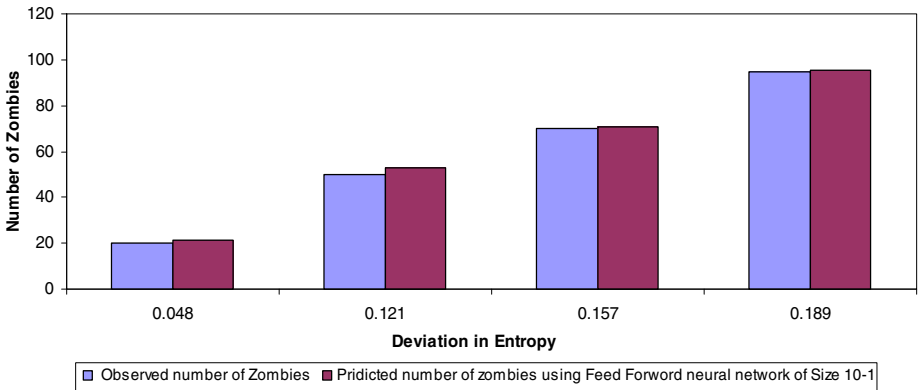


Fig. 2. Comparison between actual number of zombies and predicted number of zombies using Feed forward neural network of size 10-1

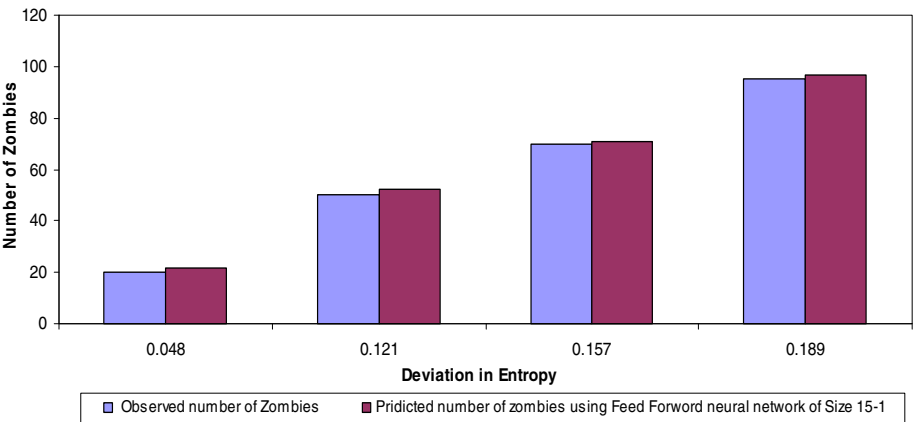


Fig. 3. Comparison between actual number of zombies and predicted number of zombies using Feed forward neural network of size 15-1

5 Conclusion and Future Work

The potential of feed forward neural network for predicting number of zombies involved in a flooding DDoS attack is investigated. The deviation ($H_c(X) - X_n(X)$) in sample entropy is used as an input and MSE is used as the performance measure. Two layer feed forward networks of size 5, 10 and 15 have shown maximum mean square error (MSE) of 2.91, 2.59 and 3.14 respectively in predicting the number of zombies. Therefore, total number of predicted zombies using feed forward neural network is very close to actual number of zombies. However, simulation results are promising as we are able to predict number of zombies efficiently, experimental study using a real time test bed can strongly validate our claim.

References

1. Gupta, B.B., Misra, M., Joshi, R.C.: An ISP level Solution to Combat DDoS attacks using Combined Statistical Based Approach. *International Journal of Information Assurance and Security (JIAS)* 3(2), 102–110 (2008)
2. Gupta, B.B., Joshi, R.C., Misra, M.: Defending against Distributed Denial of Service Attacks: Issues and Challenges. *Information Security Journal: A Global Perspective* 18(5), 224–247 (2009)
3. Gupta, B.B., Joshi, R.C., Misra, M.: Dynamic and Auto Responsive Solution for Distributed Denial-of-Service Attacks Detection in ISP Network. *International Journal of Computer Theory and Engineering (IJCTE)* 1(1), 71–80 (2009)
4. Burns, R., Burns, S.: *Advanced Control Engineering*. Butterworth Heinemann (2001)
5. Dayhoff, U.E., DeLeo, J.M.: *Artificial neural networks*. *Cancer* 91(S8), 1615–1635 (2001)
6. Yegnanarayana, B.: *Artificial Neural Networks*. Prentice-Hall, New Delhi (1999)
7. Moore, D., Shannon, C., Brown, D.J., Voelker, G., Savage, S.: Inferring Internet Denial-of-Service Activity. *ACM Transactions on Computer Systems* 24(2), 115–139 (2006)
8. GT-ITM Traffic Generator Documentation and tool,
<http://www.cc.gatech.edu/fac/EllenLegura/graphs.html>
9. NS Documentation, <http://www.isi.edu/nsnam/ns>
10. Shannon, C.E.: A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communication Review* 5, 3–55 (2001)
11. Gibson, B.: TCP Limitations on File Transfer Performance Hamper the Global Internet. White paper (2006),
<http://www.niwotnetworks.com/gbx/TCPLimitsFastFileTransfer.htm>

A Novel Biometric Watermarking Approach Using LWT- SVD

Meenakshi Arya¹ and Rajesh Siddavatham²

¹ Lecturer, ² Associate Professor

Department of CSE & IT,

Jaypee University of Information Technology, Wagnaghat,

Solan, Himachal Pradesh, India

meenakshi.arya@juit.ac.in, srajesh@juit.ac.in

Abstract. The lifting wavelet transform (LWT) is a recent approach to wavelet transform, and singular value decomposition (SVD) is a valuable transform technique for robust digital watermarking. While LWT allows generating an infinite number of discrete biorthogonal wavelets starting from an initial one, singular values (SV) allow us to make changes in an image without affecting the image quality much. This paper presents an approach which tries to amalgamate the features of these two transforms to achieve a hybrid and robust digital image watermarking techniques. Certain performance metrics are used to test the robustness of the method against common image processing attacks.

Keywords: Biorthogonal wavelets, Biometric watermarking, Lifting wavelet transform, Singular value decomposition.

1 Introduction

Biometrics based authentication systems are becoming increasingly popular as they offer enhanced security and user convenience as compared to traditional token-based (I.D. card) and knowledge based (password) systems. Biometric watermarking refers to embedding a biometric trait like fingerprint [1], face [2], handwritten signature [3] etc. for the purpose of content authentication.

In [7], LSB and DWT have been synergistically combined to embed face template in fingerprint image while the same idea was applied in [4] for embedding offline handwritten signature in a host image. Vatsa et al. [6] developed a v- SVM based biometric watermarking method which was further revised by Cheng et al. [5] for biometric watermarking based on offline handwritten signature.

In the current literature, neither lifting wavelet transform nor singular value decomposition have been used for biometric watermarking. In this paper, a novel LWT-SVD based biometric watermarking technique for offline handwritten signatures has been proposed. The lifting scheme has been used firstly for separating the significant pixels of the host image from the insignificant ones and then the Singular Value Decomposition (SVD) is applied. The watermark is embedded at this level using a gain factor (k). The watermarked image is then obtained by taking inverse LWT transform. The proposed algorithm gives excellent results for various attacks on the host image. The rest of the paper is organized as follows: Section 2 explains the theoretical

framework of SVD and LWT while Section 3 presents the proposed method. In Section 4, the significance measures PSNR and SSIM have been described to assess the quality of the watermarked image and the recovered signature image. The efficiency of the proposed method along with the results have been presented in Section 5. Section 6 concludes the work.

2 Theoretical Framework of Lifting Wavelet Transform (LWT) and Singular Value Transform (SVD)

2.1 Lifting Wavelet Transform (LWT)

The basic idea of wavelet transforms is to exploit the correlation structure present in most real life signals to build a sparse approximation. The lifting scheme is a technique for both designing fast wavelets and performing the discrete wavelet transform. The technique was introduced by Swelden [8, 9]. While the discrete wavelet transform applies several filters separately to the same signal, the signal is divided like zipper for the lifting scheme. Then a series of convolution-accumulate operations across the divided signals is applied. Generally speaking, lifting scheme includes three steps that are splitting, prediction and update. The basic idea of lifting is described here briefly:

Split: The original signal is divided into two disjoint subsets. Although any disjoint split is possible, we will split the original data set $x[n]$ into $x_e[n]$ $x[2n]$, the even indexed points and $x_o[n]$ $x[2n+1]$, the odd indexed points.

Predict: The wavelet coefficients $d[n]$ is generated as error in predicting $x_o[n]$ from $x_e[n]$ using prediction operator P .

$$d[n] = x_o[n] - P(x_e[n]). \quad (1)$$

Update: $x_e[n]$ and $d[n]$ are combined to obtain scaling coefficients $c[n]$ that represent a coarse approximation to the original signal $[n]$. This is accomplished by applying an update operator U to the wavelet coefficients and adding the result to $x_e[n]$:

$$c[n] = x_e[n] + U(d[n]). \quad (2)$$

These three steps form a lifting stage. Iteration of the lifting stage on the output $c[n]$ creates the complete set of DWT scaling and wavelet coefficients $c_j[n]$ and $d_j[n]$. At each scale we weight the $c_j[n]$ and $d_j[n]$ with k_e and k_o respectively as shown in Fig. 1. This normalizes the energy of the underlying scaling and wavelet functions.

The lifting steps are easily inverted even if P and U are nonlinear, space-varying, or noninvertible. Rearranging equation (1) and (2) we have

$$x_e[n] = c[n] - U(d[n]). \quad (3)$$

$$x_o[n] = d[n] + P(x_e[n]). \quad (4)$$

The original signal will be perfectly reconstructed as long as the same P and U are chosen for the forward and the inverse transforms.

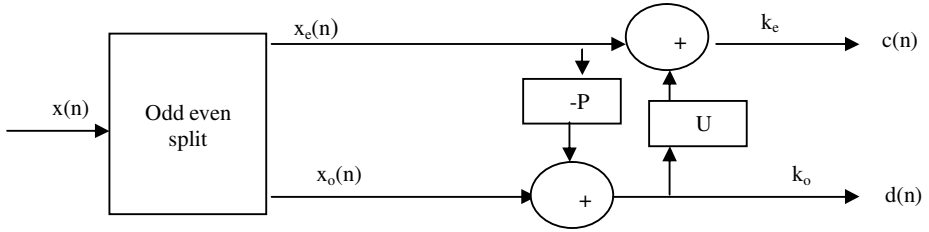


Fig. 1. Lifting Steps

2.2 Singular Value Decomposition (SVD)

The Singular Value Transform (SVD) was explored a few years ago for watermarking purposes. In recent years, SVD has been used in watermarking as a different transform as it is one of the most powerful tools of linear algebra with several applications in watermarking[10,11,12,13,14]. Singular values are the luminance values of SVD image layer, changing these values slightly do not affect the image quality much .The purpose of singular value decomposition is to reduce a dataset containing a large number of values to a dataset containing significantly fewer values, but which still contains a large fraction of the variability present in the original data. SVD analysis results in a more compact representation of these correlations, especially with multi-variate datasets and can provide insight into spatial and temporal variations exhibited in the fields of data being analyzed. SVD is optimal matrix decomposition in a least square sense packing the maximum signal energy into a few coefficients as possible [11]. The SVD theorem decomposes a digital image A of size $M \times N$, as:

$$AV = U\Sigma. \tag{5}$$

$$A^T U = V\Sigma. \tag{6}$$

Since U and V are orthogonal, this becomes the singular value decomposition

$$A = U\Sigma V^T. \tag{7}$$

The full singular value decomposition of an $(M \times N)$ matrix involves an $(M \times M)$ U , an $(M \times N)$ Σ , and an $(N \times N)$ V . In other words, U and V are both square and Σ is the same size as A . The singular value decomposition is the appropriate tool for analyzing a mapping from one vector space into another vector space, possibly with a different dimension.

3 Proposed Technique

An image comprises of certain high frequency components (edges) known as the approximation coefficients and low frequency components (smooth areas) known as the detailed coefficients. Most of the previous SVD and DWT-based watermarking

techniques treat different parts of the image in the same way. Therefore, the edges and the smooth areas of the image, related to different sub-bands, accept similar effects. The HVS is less sensitive to noise on edges, hence making similar changes to perceptually significant and insignificant areas of the image consequently lead to noticeable alternation in smooth areas, thereby causing a significant degradation to the image quality.

The paper proposes a novel biometric watermarking technique with imperceptible image quality alteration. Additional advantages of the presented technique could be highlighted as high capacity and robustness of the method against different types of common attacks. Since LWT provides high redundancy in transform domain, the high capacity of the transformed host could utilize as the beneficial point to scatter the watermark data.

3.1 Biometric Feature Processing

To employ offline handwritten signature as watermark, the preprocessing algorithm as depicted in Fig 3 is applied on the signature image. Initially, the signature image is binarized and resized to an image of 300 pixels x 200 pixels. This is to isolate single stroke or a cluster of separated strokes of a handwritten signature from the background. Median filter is applied to this binary image to eliminate noise which might be present in the form of speckles, smears, scratches etc. that might thwart feature extraction. Hough transform (HT) is then applied to the signature image for projection into feature space. The step is followed by applying Principle Components Analysis (PCA) is to compress the feature space generated by HT without losing the significant attributes [15]. Lastly, PCA feature is statistically discretized into binary representation signature code as proposed in [16].

3.2 Watermark Embedding

The following steps explain the embedding phase.

- (i) Let I_{original} be the host image of size $N \times N$.
- (ii) The Lifting Wavelet Transform $I_{\text{LWT}(i,j)}$ of the host image is calculated according to the selected decomposition level (L), sub-bands of size $\frac{N}{2^L} \times \frac{N}{2^L}$ can be achieved.
- (iii) Let S_{original} be the original offline handwritten signature of $m \times n$ where $m \leq n$. Resize the signature image such that size $(I_{\text{original}}) = \text{size}(S_{\text{original}})$
- (iv) Calculate $S_{\text{LWT}(i,j)}$, the corresponding wavelet transform of the signature image.
- (v) At $L=2$, apply SVD to the horizontal detailed sub-band of the cover image as well as to the signature image.
- (vi) The singular values of the cover image sub-band are modified with the singular values of the signature sub-band obtaining modified LWT coefficient at the 2nd level.

$$[I_{wm}(2, h)]_{\text{singular}} = [I_{\text{original}}(2, h)]_{\text{singular}} + k * [S_{\text{original}}(2, h)]_{\text{singular}} \quad (8)$$

Embedding at this level is described as

$$I_{lwt}(2, j) = \begin{cases} S_{lwt}(2, h) \\ I_{lwt}(2, j) \end{cases} \tag{9}$$

(vii) Using the inverse wavelet transformation the final watermarked image I_{wm} will be constructed.

3.3 Watermark Extraction

Since the SVs of the original images are needed in the extraction phase, the proposed technique is non-blind as it uses the singular vector matrices of the original signature image as the keys. The extraction phase is explained by the following steps

- (i) Compute the Lifting Wavelet Transform of the watermarked image according to the selected decomposition level (L)
- (ii) Locate the embedded coefficients and extract the singular values of the corresponding sub-band of the signature image through Equation 10.

$$\sum S_{wm} = (I_{wm} - I_{original})/k \tag{10}$$

- (iii) Combine the SVs thus obtained to recover the 2nd level approximation coefficient.
- (iv) Perform 2 –level Inverse LWT to obtain the watermark.

3.4 Template Matching Based Authentication

This extracted watermark is fed as an input to the biometric feature processing algorithm for template matching. The database contains 250 offline handwritten signatures collected from 50 users at different times to capture the intrapersonal differences in signing by a single user. Initially all the steps mentioned in biometric feature processing are applied to the entire signature database to generate a feature vector comprising the feature vectors corresponding to each signature image. These steps are applied to the recovered signature image to extract its features. The Euclidean distance between the feature vector of the recovered signature and the feature vectors of all the signatures in the database is calculated according to the formula as given by Equation 11.

$$dist(x, y)(a, b) = \sqrt{(x - a)^2 + (y - b)^2} \tag{11}$$

The database image with the least Euclidean distance with the extracted image is the corresponding template and hence the verification of the signature of the user.

4 Significance Measures

4.1 Peak Signal to Noise Ratio PSNR)

The proposed algorithm has been tested for various signal processing attacks like median filtering, salt and pepper noise addition, histogram equalization, Gaussian

noise and JPEG compression. The experimental results have been gauged using Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) which have been given below.

$$MSE = \frac{1}{mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (I(x, y) - W(x, y))^2. \quad (12)$$

where I and W are the original and the watermarked images having a resolution of m*n.

$$PSNR = 10 \log_{10} \frac{\max^2}{MSE}. \quad (13)$$

4.2 Structural Similarity Index Measure (SSIM)

SSIM [14] is a new paradigm metric designed to improve on traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE) for quality assessment. It is based on the hypothesis that the HVS is highly adapted for extracting structural information. The measure of structural similarity compares local patterns of pixel intensities that have been normalized for luminance and contrast. In practice, a single overall index is sufficient enough to evaluate the overall image quality; hence a mean SSIM (MSSIM) index is used as the quality measurement metric.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_1)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_1)}. \quad (14)$$

$$MSSIM(x, y) = \frac{1}{M} \sum_{m=1}^M SSIM(x_m, y_m). \quad (15)$$

5 Results and Discussions








The proposed algorithm based on Lifting wavelet transform and singular value decomposition generates results that are superior to the existing methods for offline handwritten signature watermarking [4,5]. The algorithm embeds the signature as a whole thus providing better authentication than the previous methods. Table 1 shows the PSNR between the original and the recovered watermark varies between 53 dB to 60 dB for the signatures of 30 users when the watermarked image is not subjected to any attack, while in Table 2, it can be seen that the PSNR value of recovered watermarks (for 3 users) after the watermarked image is subjected to various attacks varies between 50 dB to 55 dB.

Even after subjecting the watermarked image to a JPEG compression ratio ranging between 90% to 30%, the watermark recovery is pretty good. For various noise ratios between 10% to 20%, the PSNR varies between 30 dB to 40 dB as shown in Table 2. For implementing the algorithm, MATLAB 7 on a 1.73 GHz Pentium M Processor with minimum 256 MB of RAM has been used. The results have been verified on various standard images like Lena, Peppers, Baboon and Elaine. Figures 2 and 3 show the effect of varying the embedding factor while Figures 4 and 7 show the effect of varying the decomposition level on various host images.








Table 1. Extracted Watermarks

PSNR=54 MSSIM=0.99886 	PSNR=53 MSSIM=0.99826 	PSNR=54 MSSIM=0.99909 	PSNR=54 MSSIM=0.9983 	PSNR=55 MSSIM=0.99995 
PSNR=54 MSSIM=0.99922 	PSNR=54 MSSIM=0.99863 	PSNR=53 MSSIM=0.99929 	PSNR=53 MSSIM=0.99853 	PSNR=55 MSSIM=0.99949 
PSNR=59 MSSIM=0.99949 	PSNR=54 MSSIM=0.99969 	PSNR=54 MSSIM=0.9994 	PSNR=54 MSSIM=0.99914 	PSNR=57 MSSIM=0.99998 

Table 2. Extracted Watermarks after Simulation of Various Attacks

CROPPING	HISTOGRAM	MEDIAN	SALT & PEPPER	GAUSSIAN	SHARPENING	JPEG
PSNR=51 MSSIM=0.997 	PSNR=28 MSSIM=0.850 	PSNR=31 MSSIM=0.965 	PSNR=29 MSSIM=0.790 	PSNR=26 MSSIM=0.704 	PSNR=24 MSSIM=0.784 	PSNR=44 MSSIM=0.993 

(a) Signature 1

CROPPING	HISTOGRAM	MEDIAN	GAUSSIAN	SHARPENING	SALT & PEPPER	JPEG
PSNR=50 MSSIM=0.999 	PSNR=30 MSSIM=0.979 	PSNR=30 MSSIM=0.990 	PSNR=28 MSSIM=0.956 	PSNR=26 MSSIM=0.963 	PSNR=30 MSSIM=0.974 	PSNR=44 MSSIM=0.997 

(b) Signature 21

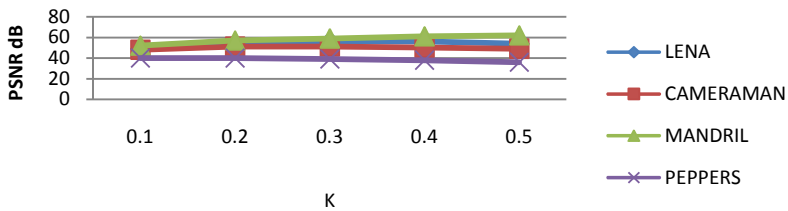


Fig. 2. PSNR versus embedding factor for various cover images

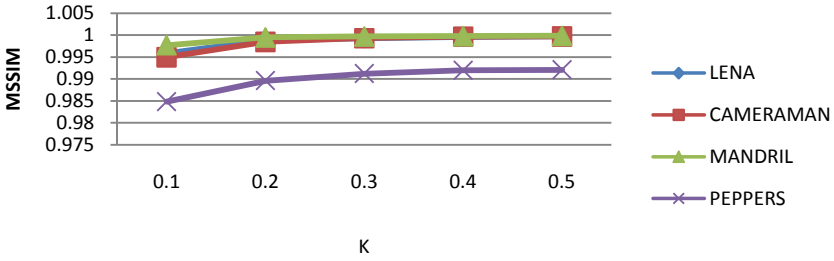


Fig. 3. MSSIM versus embedding factor for various cover images

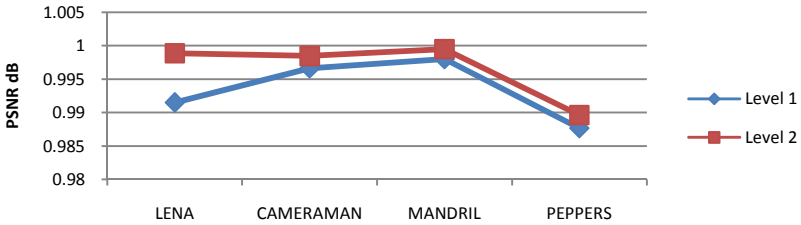


Fig. 4. PSNR versus decomposition level for various cover images

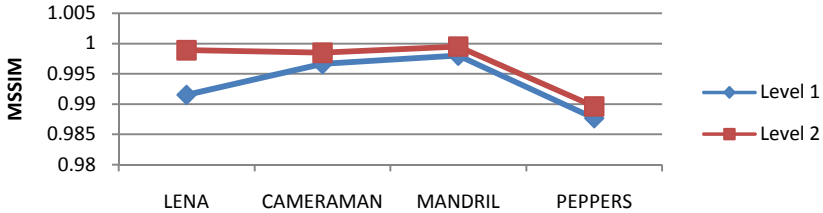


Fig. 5. MSSIM versus decomposition level for various cover images

Furthermore, the performance of the proposed algorithm has been tested for various values of embedding factor. The effect of embedding factor on PSNR and MSSIM has been presented in Tables 3 and 4 while Tables 5 and 6 show the effect of varying the decomposition level.

6 Conclusion

In this paper, a novel biometric watermarking scheme using LWT-SVD for offline handwritten signature has been proposed. The proposed technique shows superior results as compared to the existing technique. The work can be further expanded by incorporating the latest signature verification techniques so as to reduce the FAR or FRR of the proposed system and also amalgamate the two areas of biometric watermarking and signature authentication/ verification.

References

1. Jain, A.K., Hong, L., Bolle, R.: On-line Fingerprint Verification. *IEEE Trans. PAMI* 19(4), 302–314 (1997)
2. Pang, Y.H., Teoh, A.B.J., David Ngo, C.L.: Enhanced Pseudo Zernike Moments in Face Recognition. *IEICE Electron, Express* 2(3), 70–75 (2005)
3. Low, C.Y., Teoh, A.B.-J., Tea, C.: A Preliminary Study on Biometric Watermarking for Offline Handwritten Signature. In: *Proc. Of 2007 IEEE International Conference on Telecommunications, Malaysia* (2007)
4. Low, C.Y., Teoh, A.B.-J., Tea, C.: Fusion of LSB and LWT Biometric Watermarking for Offline Handwritten Signature. In: *2008 Congress on Image and signal processing*, pp. 702–708. *IEEE Computer Society, Los Alamitos* (2008)
5. Low, C.Y., Teoh, A.B.-J., Tea, C.: Support Vector Machines (SVM) based Biometric Watermarking for Offline Handwritten Signature. In: *2008 Congress on Image and signal processing*, pp. 702–708. *IEEE Computer Society, Los Alamitos* (2008)
6. Vasta, M., Singh, R., Noore, A.: Improving Biometric recognition accuracy and robustness using DWT and SVM watermarking. *IEICE Electronics Express* 2(12), 362–367 (2005)
7. Vasta, M., Singh, R., Noore, A.: Robust Biometric Image Watermarking for Fingerprint and Face Template Protection. *IEICE Tran. On Fundamentals of Electronics* (2006)
8. Sweldens, W.: The lifting scheme: A New Philosophy in Biorthogonal Wavelet Constructions. In: *Proceedings of SPIE*, pp. 68–79 (1995)
9. Daubechies, I., Sweldens, W.: Factoring Wavelet Transforms into Lifting Schemes. *The Journal of Fourier Analysis and Applications* 4, 247–269 (1998)
10. Chandra, D.V.S.: Digital Image Watermarking using Singular Value Decomposition. In: *Proc. of 45th IEEE Midwest Symposium on circuits and Systems, Tulsa, OK*, pp. 264–267 (2002)
11. Liu, R., Tan, T.: A SVD based Watermarking scheme for protecting rightful ownership. *IEEE Transactions on Multimedia* 4(1), 121–128 (2002)
12. Zhou, B., Chen, J.: A Geometric Distortion Resilient Image Watermarking Algorithm Based on SVD. *Chinese Journal of Image and Graphics* 9, 506–512 (2004)
13. Bao, P., Ma, X.: Image Adaptive Watermarking Using Wavelet Domain Singular Value Decomposition. *IEEE Transactions on Circuits and Systems for Video Technology* 15(1), 96–102 (2005)
14. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions On Image Processing* 13(4) (2004)
15. Ooi, S. Y., Teoh, A.B.J., David Ngo, C.L.: Offline Signature Verification through Discrete Radon Transform and Principal Component Analysis. In: *Proc. of International Conference on Computer and Communication Engineering (ICCCCE)*, (2006).
16. Tuyls, P., Akkermans, A.H.M., Kevenaar, T.A.M.: Face Recognition with Renewable and Privacy Preserving Binary Templates. In: *4th IEEE Workshop on Automatic Identification Advanced Technologies* (2005)

Detection and Prevention of Phishing Attack Using Dynamic Watermarking

Akhilendra Pratap Singh, Vimal Kumar, Sandeep Singh Sengar, and Manoj Wairiya

Department of Computer Science and Engineering

MNNIT, Allahabad

{akhil121282,vimaliitr10,sansen0911,wairiya}@gmail.com

Abstract. Nowadays phishing attacks are increasing with burgeoning rate which is highly problematic for social and financial websites. Many anti-phishing mechanisms currently focused to verify whether a web site is genuine or not. This paper proposes a novel anti-phishing approach based on Dynamic watermarking technique. According to this approach user will be asked for some additional information like watermark image, its fixing position and secret key at the time of user's registration and these credentials of particular user will be changed at per login. During each login phase a user will verify the authentic watermark with its position and decide the legitimacy of website.

Keywords: Phishing Attack, Watermarking, Website, Authentication.

1 Introduction

Today all organizations are using the internet for sharing the message so the communication must be secure. Few unethical hackers are doing the cyber criminal activity by committing fraud. Attacker sends the phishing message by the fake website that looks like a original site. This attack treated as a deceptive phishing attack which target to the financial organization. Criminals complete their life cycle in very short period by the login and personal detail of the people. Phishing attack has various types as deceptive, malware, keyloggers, data theft, search engine, content injection and web trojan. Many anti phishing tools exist to protect the attack. Juan Chen et.al. have proposed a very standard approach in [1]. According to their idea they have propose a new end-host based anti-phishing algorithm, which is called Link Guard, which works by utilizing the generic characteristics of the hyperlinks in phishing attacks. These characteristics are derived by analyzing the phishing data archive provided by the Anti-Phishing Working Group (APWG). Because it is based on the generic characteristics of phishing attacks, Link Guard can detect not only known but also unknown phishing attacks. Michael Atighetchi et .al. have given the idea in [2]. They suggested a framework based on attribute based checks for defending against phishing attacks. According to Amir Herzberg et.al.s approach presented in [3] they present an improved security and identification indicators, as author implemented, a browser extension in Trust Bar. Dmytro Iliev et.al.s proposed idea in [7] phishing prevention approach based on mutual authentication is provided. Mohsen Sharifi et.al.

have suggest an idea in [4] AntiPhishing Authentication (APA) technique to detect and prevent real-time phishing attacks. Daisuke Miyamoto et.al.s gave an approach in [5] on performance of machine learning-based methods such as Regression Trees and Classification, Naive Bays, Additive Regression Trees , Logistic Regression for detection of phishing sites. Daisuke Miyamoto et.al. Proposed approach in [6] that is study of users past trust decisions (PTDs) for improving the accuracy of detecting phishing sites. Mercan Topkara et.al. have suggested an approach using watermark in [8] which is called ViWiD. ViWiD is an integrity check system based on visible watermarking of logo images. Various webpage related security using watermarking technique is discussed in [9][10][11][13][14]. Section 2 describes the proposed algorithm. Section 3 contains all experimental results and section 4 concludes the paper.

2 Proposed Algorithm

There are many methods proposed earlier to detect and prevent phishing attack. Some of them use watermark technique and some them uses another approach. There is some limitation in this approach like many checks and enforcements which are used by the client-side defense tools can be tricked by attackers after getting a reasonable knowing of web site construction [4]. For example, using mosaic attack, an attacker can make fool the image check system of SpoofGaurd by partitioning the logo image into small parts and show it in such way that it looks like legitimate one. Some time due to irritation of some protection means like antivirus, user turnoff the protection mechanism of client side. Hence all client side schemes will live all defensive actions. Similarly in case of tools which are based on cryptography, they need individual downloaded software in each client side machine. In some cases in spite of using SSL secure connection, if client authentication tool will be turned off, it may suffers with Phishing attack. Few earlier proposed algorithms also use watermarking technique but visible and stationary nature of their watermark may also suffer with phishing attack. In this paper we are proposing an approach for prevention of phishing attack based on dynamic position watermarking technique. This approach is divided in to three modules viz. Registration process, Login verification process and Web site closing process. Different position for watermark image can be top left, bottom left, top right, bottom right, center.

2.1 Registration Process

Registration process is the first phase whenever we open the website and trying to become a member of the website. Hence first communication between client and server is done in this phase. Most of the financial, social networking web site gives the opportunity to make an account in their server by uploading user credentials like user name and password etc. These credentials play vital role for further communication with those web site. There are three phase in registration process. You have more than one assuming, please make sure that the Volume Editor knows how you are to be listed in the author index.

- Step 1. Client will open the web page which starts with registration phase. There will be five most essential credentials of user viz. Username, Password, Secret key, Watermark Image and position of watermark. Here Secret key will be unique for each user and hence it is called primary key. Since we are assuming a secure channel between client and server, hence all credentials will be in encrypted form and there will be no attacks like man in the middle.
- Step 2. In this phase web site hosting server will store all data related to particular user. Here Secret key will be primary key for the database.
- Step 3. User is acknowledged with a proper web page having desired watermark at predefined location. Now the user will get login page for entering his username and password.

2.2 Login Verification Phase

Once user has created his account he needs to log in the web site. This is the very critical phase which mostly suffers with phishing attack because most of the attackers create a fake website similar to the legitimate one. This phase consist of eight steps.

- Step 1. Whenever user will open the website he will be asked for proper Secret key which is unique for each user. In this phase there will not be login window.
- Step 2. A query will be forwarded to database of server for retrieving all credentials of user associated with that particular secret key.
- Step 3. After getting a proper match from database, two important credentials (Watermark image and its position) will be returned to website hosting server.
- Step 4. Login page will be displayed to user with proper watermark image and its location as set by user at the time of registration.
- Step 5. After ensuring the correctness of watermark image and its position, a legitimate user will verify the authenticity of the website and then only he will enter his login id and password.
- Step 6. A query will be passed to database for entered username and password.
- Step 7. Now according to username and password, all information related to that particular user will be retrieved from database to the server.
- Step 8. A proper account will be shown to user which is of legitimate website. Now we can trust on that website.

2.3 Website Closing Phase

Since proposed algorithm is based on dynamic watermarking, hence at per login the position and nature of watermark must be changed which is only known to legitimate user.

- Step 1. User starts the closing process by clicking on the close button.
- Step 2. During logout, user will be prompt for reentering the new watermark image and its location. This step is very essential and can not be ignored.
- Step 3. After resetting the watermark image and position, user must have to wait for acknowledgment from server.
- Step 4. These all new information will be stored in database and old information will be invalidated.
- Step 5. After successful updating user will be acknowledged.

2.4 Determination of Phishing Website

Suppose attacker has created a phishing website which looks similar to the original one. As soon as user will click on suspicious link, fake website will be open which is ask for secret key.

- Step 1. In this step user will enter the secret key and wait for the desired watermark image at particular position.
- Step 2. A fake website will not have the database related to watermark information and its position. That's why it will be very difficult to determine or guess the correct watermark and position by attacker.
- Step 3. Due to absence of proper watermark and its location user can determine that it is not authorized one and he is going to be suffered with phishing attack.

In this case the user must open the proper website by verifying the URL and then must change his current secret key by newer one. Because attackers now aware with the old secret key. Hence we will invalidate that secret key. When user enters correct username and password, then only he can see his account of website, so here our main aim is to protect username and password. Suppose if an attacker knows the secret key, which is not changed by newer one till now. At that condition the watermark image and its position doesn't matter for the attacker. Now he will get login page but still unaware of username and password so he will not be able to see the account information as well as he can not change the secret key because during changing it, a attacker must know all credentials of user. Hence the flowchart given in figure 1 will demonstrate the procedure to determine whether the website is phishing one or original one.

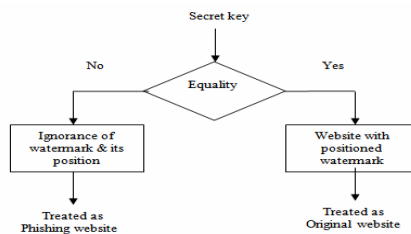


Fig. 1. Flow chart for phishing Identification

3 Experimental Results

Proposed algorithm is implemented and verified over the Local area network. In our experiment we have created a website for a bank named ABC Bank. First of all users has created his account and stored his credentials over the server database. Here we have taken a passport size photograph as a visible watermark and chosen the top left position as a location for watermark. After creating an account whenever we open the web site it will be looked like figure and it will only prompt for secret key as set by



Fig. 2. Webpage for secret key insertion



Fig. 3. Web page for login and password insertion

legitimate user at the time of registration. Once we have entered correct secret key it will show the web page look like figure 2. by which a user can verify the watermark and its location and according to this information he can assure that opened website is authentic one.

Now the user will be prompted for username and password. After entering correct user name and password user will able to see his actual account information from a legitimate and desired website. This is the exact way to prevent from Phishing attack as proposed in this paper.

4 Conclusion

In this paper, we propose a new anti-phishing approach based on Dynamic watermarking technique. This scheme neither requires online interactions with a Detection and Prevention of Phishing Attack using Dynamic Watermarking third party, nor requires any plug in or online tool hence this approach is more user friendly than the previous approaches. Experimental results show the working of our approach in which a user requires only different watermark at the time of per login which is more tolerable than being hacked by attacker. It is clearly revealed that by determining the main differences between the legitimate website and the phishing, one can reduce the risk of this type of attack. According to experimental results it is clear that proposed approach is more applicable for social and financial websites than others.

References

1. Chen, J., Guo, C.: Online Detection and Prevention of Phishing Attacks. IEEE, Los Alamitos (2006)
2. Atighetchi, M., Pal, P.: Attribute-based Prevention of Phishing Attacks. In: IEEE Int. Symposium on Network Computing and Applications (2009)
3. Herzberg, A., Jbara, A.: Security and Identification dicators for Browsers against Spoofing and Phishing Attacks, ACM Mohsen Sharif, A Zero Knowledge Password Proof Mutual Authentication Technique Against Real-Time Phishing Attacks. Springer, Heidelberg (2007)
4. Sharif, M., Saberi, A., Vahidi, M., Zorufi, M.: A Zero Knowledge Password Proof Mutual Authentication Technique Against Real-Time Phishing Attacks. Springer, Heidelberg (2007)
5. Miyamoto, D., Hazeyama, H., Kadobayashi, Y.: An Evaluation of Machine Learning-Based Methods for Detection of Phishing Site. Springer, Heidelberg (2009)
6. Miyamoto, D., Hazeyama, H., Kadobayashi, Y.: Human Boost: Utilization of Users past Trust Decision for Identifying Fraudulent Websites. Springer, Heidelberg (2009)
7. Iliev, D., Sun, Y.B.: Website forgery prevention IEEE (2010)
8. Topkara, M., Kamra, A., Atallah, M.J., Rotaru, C.N.: ViWiD: Visible Watermarking Based Defense Against Phishing. Springer, Heidelberg (2005)
9. Frattolillo, F.: Watermarking Protocol for Web Context. IEEE Transactions on Information Forensics and Security 2(3) (September 2009)
10. Jin, C., Xu, H., Zhang, X.: Web Pages Tamper-Proof Method Using Virus-Based Watermarking. IEEE, Los Alamitos (2008)
11. Sun, P., Lu, H.: An efficient web page watermarking scheme. IEEE, Los Alamitos (2009)
12. Sun, P., Lu, H.: Two efficient fragile web page watermarking schemes. In: Fifth International Conference on Information Assurance and Security (2009)
13. Long, X., Peng, H., Zhang, C.: A Fragile Watermarking Scheme Based On SVD for Web Pages. IEEE, Los Alamitos (2009)
14. Okada, M., Okabe, Y., Uehara, T.: A Web-based Privacy-Secure Content Trading System for Small Content Providers Using Semi-Blind Digital Watermarking. In: IEEE Communications Society subject matter experts for publication in the IEEE CCNC 2010 (2010)

A Search Tool Using Genetic Algorithm

M.K. Thanuja and C. Mala

Department of Computer science and Engineering, National Institute of technology,
Thiruchrapalli, India

thanujaprakash@yahoo.in, mala@nitt.edu

Abstract. In the current business scenario, with explosive growth of the amount of information resources available over the Internet and Intranets in any organization, the retrieval of the required information at the spot and time of requirement is very essential, for the effective performance of the organization, by reducing the process time and meeting customer satisfaction of meeting delivery schedule. When the number of document collection and users of the document go beyond an extent, an efficient search tool for the retrieving the information from the collection of documents becomes vital. Actual information retrieval means searching for keyword within documents. This study investigates the various stages of information retrieval, selection of best method for implementation at each stage and optimizing the solution using of genetic algorithm with different parameters. The method is tested with a training data of document collections, where more relevant documents are presented to users in the genetic modification. In this paper a new fitness function is presented in the genetic algorithm for appropriate information retrieval which is found to be efficient than other fitness functions.

Keywords: *Cosine* similarity, Fitness function, Genetic Algorithm, Information Retrieval.

1 Introduction

With large amount of information resources available over the Internet and Intranets in organizations, the information overload for the user has become overwhelming. With the corresponding dramatic increase of the number of users in finding the best and the newest information has increased exponentially. The absence of suitable alternatives in the functionality of most of current information systems are, the user looking for some topic and the Information Retrieval System(IRS) retrieves too much information. The systems provide no qualitative distinction between the relevant and irrelevant documents [1]. Genetic algorithms (GAs) are not new to information retrieval. Gordon suggested representing a posting as a chromosome and using genetic algorithms to select good indexes [2]. Yang *et al.* suggested using GAs with user feedback to choose weights for search terms in a query [3]. Morgan and Kilgour suggested an intermediary between the user and IR system employing GAs to choose search terms from a thesaurus and dictionary [4]. Boughanem *et al.* [5] examine GAs for information retrieval and they suggested new crossover and mutation operators.

2 Motivation

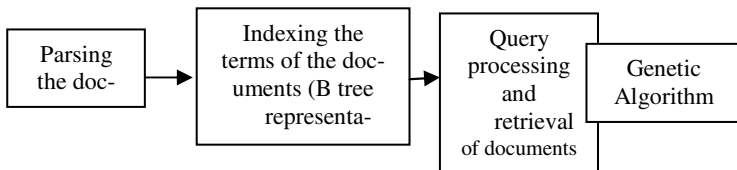
BHEL is a engineering and manufacturing organization in which a large volume of documents are stored into a vault through different process based applications. Due to lack of search tools, the past and the present knowledge in the documents are not readily available. This paper aims at the study the process of information retrieval System and the efficient methods at each stage and optimizing the solution for giving the best results to the users.

3 Proposed Method

3.1 Information Retrieval System

Information Retrieval System (IRS), a system used to extract and store information pertaining to a collection, in required form, that need to be processed, searched and retrieved based on the user's query. The methods for information retrieval that have been proposed in the literature are full text scanning, indexing, signature files and clustering[6].When a comparison is made among the above methods indexing method is found to be advantageous in terms of access time as well as storage space. So it is proposed to use indexing method of data retrieval, for the proposed search tool.

3.2 Frame Work



3.3 Indexing

Most IRSs use keywords to retrieve documents. The complete indexing process involves the steps namely markup and format removal of tags, tokenization, filtration and indexing. Various index techniques are available in the field of IR, but the common and more efficient indexing technique found in the literature is inverted index. It is proposed to use inverted index technique for the proposed search tool. An inverted index is an index data structure storing a mapping from content, such as words, to its location in a document or a set of documents. It is the most popular data structure used in document retrieval systems [6].A general inverted file index consists of two parts: a set of inverted file entries, being lists of identifiers of the records containing each indexed word, and a search structure for identifying the location of the inverted file entry for each term. The search structure may contain the information like the document id in which the term is located. A typical inverted index appears as follows

term1	(1,2) (3,6) (3,3)
term2	(3,4) (2,1)
term3	(1,3) (2,3) (3,1)
term4	(2,4)

Fig. 1. An index structure

It is proposed to implement the search structure using a B tree. In this method, for every word(Term) that occurs in a document, the inverted list contains an entry that includes a document number and the weight of the term in the document. This research presents the following weighting scheme for the weight of the terms[6].

$$W(t_i,d_j)=f(t_i,d_j)/\text{sqrt}(|d_j|) \times \ln(D/f(t_i)) \tag{1}$$

where $f(t_i,d_j)$ is the no. of times term t_i appears in document d_j , $|d_j|$ is the no. of terms in d_j , $f(t_i)$ is the no. of docs. Containing t_i and D is the no. of docs. in the collection.

3.4 Document Retrieval and Relevancy

In this paper it is proposed to use the vector model of the IRS, in which a document is viewed as a vector in n-dimensional document space (where n is the number of distinguishing terms used to describe contents of the documents in a collection). A query is also treated in the same way and constructed from the terms and weights provided in the user request. Document retrieval is based on the measurement of the similarity between the query and the documents. In this method, the retrieved documents can be presented with respect to their relevance to the query [7].

It is proposed to use Cosine similarity function in this work

$$\text{Sim}(Q,D_i) = \frac{\sum_j w_{Q,j} w_{i,j}}{\sqrt{\sum_j w_{Q,j}^2} \sqrt{\sum_j w_{i,j}^2}}$$

3.5 Optimization Using Genetic Algorithm

In GA, the search space is composed of solutions to the problem each represented by a string is termed as a chromosome. Each chromosome has a function value, called fitness. A set of chromosomes with their fitness is called the population. Population, at a given iteration of the GA, is called a generation. Pathak *et al.* used a genetic algorithm to choose weights for such a combination [8].It is found the time required for the retrieval of the relevant documents or the time required for downloading the relevant document are not considered in any of the fitness functions of GAs of IRS. In this paper a new fitness function is introduced involving the time factors.

3.5.1 The Genetic Approach

Once significant keywords are extracted from training data (relevant and irrelevant documents) including weights are assigned to the keywords. The binary weights of the keywords are formed as a query vector. The GA have been applied for a fitness function to get an optimal or near optimal query vector, also the results of the GA approach have been compared with the IR Systems without using GA.

3.5.1.1 Representation of the Chromosomes. These chromosomes use a binary representation, based on the relevancy of the documents with the given query. The no. of genes will be equal to the no. of documents in the collections. In this genetic approach the documents are represented as follows

$$\text{Doc}_1 = \{\text{term}_1, \text{term}_2, \text{term}_3, \dots, \text{term}_n\}$$

$$\text{Query} = \{q\text{term}_1, q\text{term}_2, q\text{term}_3, \dots, q\text{term}_m\}$$

The relevancy of the Doc1 with respect to the query is given by the cosine similarity function mentioned in section 3.4 and is either 1 or 0. The relevancy of the document with respect to the given query is encoded as chromosome as follows

Chromosome $C_1 = \{0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 10\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\}$ means the documents 2,3 & 10 are relevant to the query for a collection of 20 documents. The no. of genes in a chromosome is the total no. of documents in the collection.

3.5.1.2 Fitness Function. Fitness function is a performance measure or reward function, which evaluates how each solution, is good. In this work, the GA with following fitness function is used

$$F = \text{rel}(d) + 1/T_r + 1/T_d + S \quad (2)$$

where $\text{rel}(d)$ is the total no. of relevant documents retrieved with respect to the given query i.e $\sum_{k=1}^n x_k$ where x_k is k^{th} gene of the i^{th} chromosome, T_r is the time taken to select the relevant documents, T_d is the time taken to download the relevant documents, S total size of the relevant documents.

3.5.1.3 Selection. As the selection mechanism, the best chromosomes will on average achieve more copies, and the worst fewer copies. The algorithm stops when the fitness value of all the chromosomes of a generation is equal and are equal to the maximum fitness of the previous generations. In the GA approaches, two GA operators are used to produce offspring chromosomes, namely Crossover which occurs with crossover probability P_c . GAs construct a better solution by mixture of good characteristic of chromosome together. Higher fitness chromosome has an opportunity to be selected more, so good solution always alive to the next generation. In this work the crossover is done between the alternate chromosomes with maximum fitness. Mutation involves the modification of the gene values of a solution with some probability P_m . The mutation in this work is with probability of 0.001 and occurs at random.

4 Experimental Results

The test databases for the GA approach are documents in a vault. The experiment was applied on 10 queries. Initial generation and subsequent generation of a sample of 4 chromosomes are shown in Table 1. In the cross over process the chromosomes with maximum fitness are selected.

Table 1. Chromosomes and the fitness for initial generation and subsequent generation

Initial Generation		Subsequent generation	
Chromosomes	Fitness	Chromosomes	Fitness
01101111101100111010	63.40	11110011111010111000	63.40
11110011111010111000	63.42	01101111101100111010	63.42
01010101101001101000	59.69	010111000000110100	57.27
0101110000001000000	55.38	01010101101001000000	57.34

The maximum fitness values of 10 queries are listed in Table 2.

Table 2. Queries and the maximum fitness values

Query	Maximum fitness value
1-5	58.617 to 62.226
6-10	57.456 to 60.31

Precision= Number of documents retrieved and relevant/Total Retrieved.

From the experimental observation, the best values for this test documents collections at crossover probability $P_c = 0.8$ and mutation rate is $P_m = 0.0001$ for the GA follows

Table 3. Comparison of the precision with the study GA and without GA for first 10 queries

Precision for 50 queries			Precision for 50 queries		
Query	Precision		Query	Precision	
	With GA	Without GA		With GA	Without GA
1	0.44	0.5	6	0.63	0.7
2	0.55	0.55	7	0.625	0.75
3	0.65	0.66	8	0.65	0.66
4	0.78	0.555	9	0.8	0.85
5	0.5	0.5	10	0.56	0.56

The precision of 10 queries listed in the above table is plotted in the Fig 2.

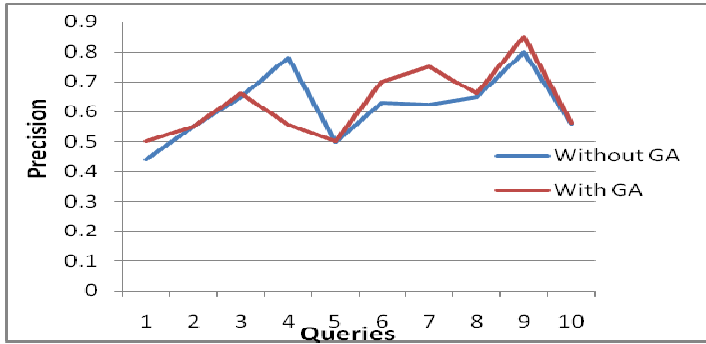


Fig. 2. Comparison of the precision with the study GA and without GA for 20 queries

5 Conclusion

From pervious results, it is noted that the new fitness function which is represented in the paper gives more sophisticated results than other fitness functions, in the test collections. Also, from the results, it is noted that our new fitness function has a precision value better than the other fitness functions.

References

- [1] Martin-Bautista, M.J., Vila, M.-A., Larsen, H.L.: A Fuzzy Genetic Algorithm Approach to an Adaptive Information Retrieval Agent
- [2] Gordon, M.: Probabilistic and genetic algorithms in document retrieval. *Communications of the ACM* 31(10), 1208–1218 (1988)
- [3] Yang, J., Korfhage, R., Rasmussen, E.: Query improvement in information retrieval using genetic algorithms—a report on the experiments of the TREC project. In: *Proceedings of the 1st text retrieval conference (TREC-1)*, pp. 31–58 (1992)
- [4] Morgan, J., Kilgour, A.: Personalising on-line information retrieval support with a genetic algorithm. In: Moscardini, A., Smith, P. (eds.) *PolyModel 16: Applications of Artificial Intelligence*, pp. 142–149 (1996)
- [5] Boughanem, M., Chrisment, C., Tamine, L.: On using genetic algorithms for multimodal relevance optimization in information retrieval. *Journal of the American Society for Information Science and Technology* 53(11), 934–942 (2002)
- [6] Faloutsos, C., Christodoulakis, S.: An access Method for Documents and its Analytical Performance Evaluation
- [7] Salton, G., McGill, M.H.: *Introduction to Modern Information Retrieval*. McGraw-Hill, New York (1983)
- [8] Pathak, P., Gordon, M., Fan, W.: Effective information retrieval using genetic algorithms based matching functions adaption. In: *Proc. 33rd Hawaii International Conference on Science (HICS)*, Hawaii, USA (2000)

Heterogeneous Data Mining Environment Based on DAM for Mobile Computing Environments

Ashutosh K. Dubey¹, Ganesh Raj Kushwaha², and Nishant Shrivastava³

¹ Dept. of Computer Science & Engineering
Trinity Institute of Technology and Research
Bhopal, India

ashutoshdubey123@gmail.com

² Dept. of Computer Science & Engineering
Trinity Institute of Technology and Research
Bhopal, India

ganeshraj Kushwaha@gmail.com

³ Dept. of Computer Science & Engineering
JNCT, Bhopal, India

nishantuit@gmail.com

Abstract. Today the concept of Data Mining services is not alone sufficient. Data mining services play an important role in the field of Communication industry. Data mining is also called knowledge discovery in several database including mobile databases and for heterogeneous environment. In this paper, we discuss and analyze the consumptive behavior based on data mining technology. We discuss and analyze different aspects of data mining techniques and their behavior in mobile devices. We also analyze the better method or rule of data mining services which is more suitable for mobile devices. In this paper, we propose a novel DAM (Define Analyze Miner) Based data mining approach for mobile computing environments. In DAM approach, we first propose about the environment according to the requirement and need of the user where we define several different data sets, then DAM analyzer accept and analyze the data set and finally apply the appropriate mining by the DAM miner on the accepted dataset. It is achieved by CLDC and MIDP component of J2ME.

Keywords: Data Mining, DAM, CLDC, MIDP.

1 Introduction

The rapidly expanding demand for digital mobile communication services, along with the wide popularity of mobile devices, has given rise to the development of new dimensions and requirements for future mobile communication systems. But today is also the requirement of heterogeneous environment where we accept different data sets according to the user and apply data mining on them. It is also use in mobile devices with the use of MIDLET and CLDC component of J2ME. In few years back, mobile extensions to Grid systems have been increasingly proposed in order to support ubiquitous access and selection to the Grid and to include mobile devices as additional

Grid resources [1, 2]. In today’s scenario mobile devices, such as mobile phones, PDAs, notebook and others, provide a basic building block [3][4][5][6].

Finding prevalent mobile user patterns and behavior in a heterogeneous environment has been one of the major problems in the area of mobile data mining. Particularly, the algorithms of discovering frequent user’s behavior patterns in the mobile agent system have been studied extensively in recent years. The key feature in most of these algorithms is that they use a dataset and frequent Item-Sets visited by the customers. In this case, some problems occur because they do not consider that mobile user’s behavior patterns are dynamically variable as time passes. In this paper we discuss some of the data mining service which are use in different areas and then apply those services to mobile devices and then apply those DMS services in mobile computing and exploiting the need of DMS in mobile computing environments using CLDC and MIDP components. The Connected, Limited Device Configuration (CLDC) and the Mobile Information Device Profile (MIDP) have emerged as J2ME standards for mobile phone applications development which are used with DMS services. The role of CLDC and MIDP component is to apply Data Mining Services in mobile. The remaining of this paper is organized as follows. We discuss CLDC and J2Me in Section 2. In Section 3 we discuss about MIDP. The proposed heterogeneous data mining algorithm, namely DAM in section 4. In section 5 we discuss about the Challenges. The conclusions and future directions are given in Section 6. Finally references are given.

2 CLDC and J2ME

The J2ME architecture is described in general before the components in the J2ME technology are introduced. J2ME applications are also discussed in general, and it is explained how they are made available to end users. J2ME is a highly optimized Java runtime environment. Fig 1 shows the J2ME architecture.

The fundamental branches of the J2ME platform are configurations. A configuration is a specification that describes a Java Virtual Machine and some set of APIs that are targeted at a specific class of device.

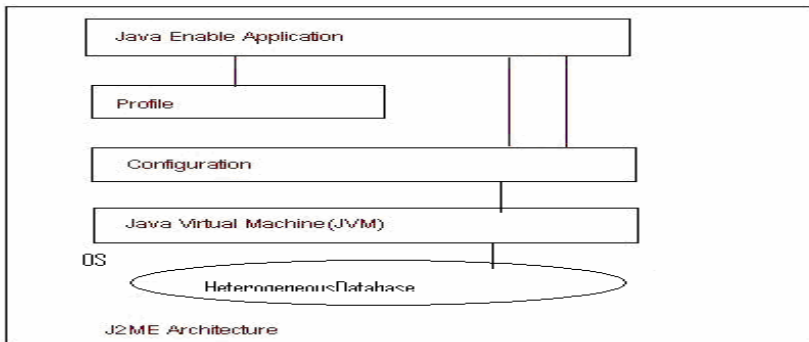


Fig. 1. J2ME Architecture

The Connected, Limited Device Configuration is one such specification. The CLDC specifies the APIs for devices with less than 512 KB of RAM available for the Java system and an intermittent (limited) network connection. It specifies a stripped-down Java virtual machine, called the KVM, as well as several APIs for fundamental application services. Three packages are minimalist versions of the J2SE java.lang, java.io, and java.util packages. A fourth package, javax.microedition.io, implements the Generic Connection Framework, a generalized API for making network connections.

Many J2ME games already exist and enjoy great popularity especially among young generation. Java comes with the immense requirement of the object-oriented programming language for developers to implement new mobile applications [7]. Configurations provide core functionality and a way to provide greater flexibility but no services for managing the application life-cycle, for driving the user interface, for maintaining and updating persistent data on the device or for secure access to information stored on a network server [8]. Fig 1 shows the CLDC position in J2ME Architecture in configuration part.

Several networks have conducted a survey on users’ watching behavior [9] which reflects that user behavior pattern recognition is not so easy task, we can achieve this by CLDC and MIDP component. Instead of replacing existing TV service, mobile services should be complementary [10], and offer more interactive means for users to watch their chosen content.

3 MIDP

The Mobile Information Device Profile is a specification for a J2ME profile. It is layered on top of CLDC and adds APIs for application life cycle, user interface, networking, and persistent storage [11]. An application written for MIDP is called a MIDlet. MIDlet applications are subclasses of the javax.microedition.midlet.MIDlet class that is defined by MIDP. MIDlets are packaged and distributed as MIDlet suites. A MIDlet suite can contain one or more MIDlets. A MIDlet is a J2ME application designed to operate on small computing device. A MIDlet is defined with at least a single class that is derived from the javax.microedition.midlet.MIDlet abstract class. The Position of MIDP is shown in Fig 2.

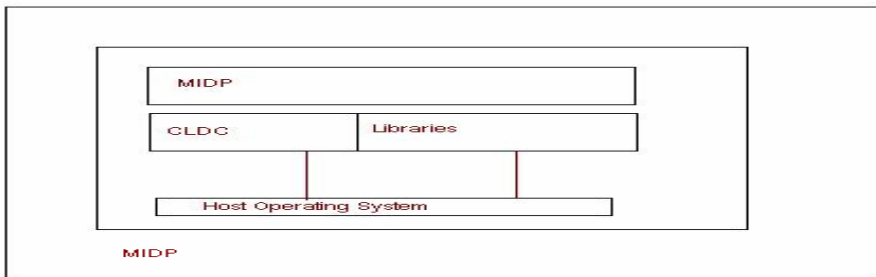


Fig. 2. MIDP in J2ME Architecture

4 Proposed Method: DAM

This method actually deal with the Heterogeneous Data Mining Environment Based on DAM for mobile computing environments, where we first design a framework where several data set activity is stored and their patterns are recognize based on the environment. When the related data set is input from any other external environment then our DAM environment recognize the particular data set if it is present in the DAM environment otherwise reject the data set for further processing.

In the second phase we select the data set if the DAM analyzer analyzes the data set from the database for further operation. Finally apply the DAM miner Strategy on the data set for finding frequent pattern and further pruning. We also consider those mining techniques which provide better performance and speed in terms of computation. In this section, we describe the proposed method. The entire concept is divided into four phases: 1) Define the data set for our database .2) Analyze and Select the data set when ever needed. 3) Apply the mining technique which is best suited. 4) Mobile computing for the data set.

1) Algorithm for Creating the Data Set

For accepting the data set we consider the following things:

- a) Table size
- b) No of Columns
- c) No of rows

Assumptions:

Col: Column, Dt: Data Type , v1, v2....vn: Values

Algorithm Define Datasets (DAM Definer)

1. Define the Data set from different data Source (WWW, XML, Data Warehouse etc.).
2. Apply the Create table statement when defining the data source Data Source.
Create table tablename as select * / Col1, Col2, Col from data source.
[We use Oracle as a Database Management Software]
3. Apply the Create table statement[Stand alone]

Create table tablename (col1 Dt1, Col2 Dt2.....Col n Dtn) with condition clause.

4. Insert the values considering the domain.
5. For Standalone [Insert into table name values (v1, v2.....vn).]
6. For step no2 when using universal false condition insert into table (Select query).
7. Finish.

2) Algorithm for Analyzing and selecting the Data Set(DAM Analyzer)

Algorithm Select Datasets (SDS)

1. Select the Data Source (WWW, XML, Data Warehouse etc.).
2. check in the database
3. if(value!=NULL) then
4. For evaluating all columns with rows [Select * from datasource.]
5. For selected set of columns [Select col1, col2coln from datasource.]
6. For some specified set of conditions

Select * / col1, col2...coln from datasource where condition.

5. Finish.

3) Algorithm for Alter the Data Set and Mining(DAM Miner)

By alter statement we can add a column, delete a column and increase and decrease the size of the column in the database.

1. Select the Data Source from (WWW, XML, Data Warehouse etc.)
2. Add new column to the data source.

Alter table tablename add (Col1 Dt1, Col2 Dt2.....Coln Dtn).

3. Modify the size of the column

If(table is empty)

{3a. Increase the size

Alter table tablename modify (Col1 Dt1, Col2 Dt2.....Coln Dtn).

3b. Decrease the size

Alter table tablename modify (Col1 Dt1, Col2 Dt2.....Coln Dtn).}

If(table is not empty)

{3c. Increase the size

Alter table tablename modify (Col1 Dt1, Col2 Dt2.....Coln Dtn).}

Else

{Exit (0) ;}

4) Algorithm for Mobile Communication the Data Set (MCDS)

1. Select the Data Source (WWW, XML, Data Warehouse etc.) .
2. Apply the MIDP Profile and using MIDlet establish the connection
3. Apply the packages of J2ME according to the need.
4. Using WTK (Wireless ToolKit) access the result on the simulator.
5. Apply data Mining techniques.
6. Finish.

After analyzing the several aspects of DAM method the picture is clear for any databases it is easy to manage and whenever necessary we can update the repository system. We also apply several data mining techniques very smoothly because our data base is consistent because of limiting redundancy in the database. Finally apply the J2ME for mobile devices so that we can coherent the entire above scenario for mobile computing environments.

5 Challenges

A number of constraints and technical difficulties faced by researchers, which are discussed in this section. These general problems must be considered for further research in this area to propose new technologies for making mobile computing easier. Some of these are:

- The screen size of the mobile is a big limitation. The screen size can affect the approximate visualization of complex results representing the discovered model.
- Mobile navigation facility is also a big task to achieve and implement.
- The overhead due to the communication between MIDLET and Data Mining service should not affect the execution time.
- The experiments on system performance depend almost entirely on the computing power of the server on which data mining task is executed.

6 Conclusions and Future Work

In this paper, we propose a novel DAM (Define Analyze Miner) Based data mining approach for mobile computing environments. In DAM approach, we first propose about the environment according to the requirement and need of the user where we define several different data sets, then DAM analyzer accept and analyze the data set and finally apply the appropriate mining by the DAM miner on the accepted dataset.

In future we also work on the limitations that were faced by the researchers.

References

- [1] Migliardi, M., Maheswaran, M., Maniymaran, B., Card, P.: Mobile Interfaces to Computational, Data, and Service Grid Systems. ACM, New York (2004)
- [2] Wesner, S., Dimitrakos, T., Jeffrey, K.: Akogrimo – The Grid goes Mobile. ERCIM (59) (2004)
- [3] Arcelus, A., Jones, M.H., Goubran, R., Knoefel, F.: Workshops, AINAW (2007)
- [4] Bahati, R.M., Bauer, M.A.: Adapting to runtime changes in policies driving autonomic management. In: ICAS 2008 : Proceedings of the Fourth International Conference on Autonomic and Autonomous Systems (ICAS 2008). IEEE Computer Society, USA (2008)
- [5] Beetz, M., Bandouch, J., Kirsch, A., Maldonado, A., Müller, A., Rusu, R.B.: Proceedings of the 4th COE Workshop on HAM (2007)
- [6] Bergmann, R.: Ambient intelligence for decision making in fire service organizations. In: AmI, pp. 73–90 (2007)
- [7] Isakow, A., Shi, H.: Review of J2ME and J2MEbased Mobile Applications. International Journal of Communication and Network Security, 189–198
- [8] Ortiz, A Survey of J2ME Today, Sun Developer Network (SDN) (2004a)
<http://developers.sun.com/mobility/getstart/articles/survey/>
(viewed August 13, 2007)
- [9] <http://www.3g.co.uk/PR/Sept2005/1943.htm>
- [10] <http://www.cellular-news.com/story/18707.php>
- [11] <http://docs.sun.com>

Selection of Views for Materialization Using Size and Query Frequency

T.V. Vijay Kumar and Mohammad Haider

School of Computer and Systems Sciences,
Jawaharlal Nehru University,
New Delhi-110067, India

Abstract. View selection is concerned with selecting a set of views that improves the query response time while fitting within the available space for materialization. The most fundamental view selection algorithm HRUA uses the view size, and ignores the query answering ability of the view, while selecting views for materialization. As a consequence, the view selected may not account for large numbers of queries. This problem is addressed by the proposed algorithm, which aims to select views by considering query frequency along with the size of the view. The proposed algorithm, in each iteration, computes the profit of each view, using the query frequency and size of views, and then selects from amongst them, the most profitable view for materialization. The views so selected would be able to answer a greater number of queries resulting in improvement in the average query response time. Further, experimental based comparison of the proposed algorithm with HRUA showed that the proposed algorithm was able to select views capable of answering significantly greater number of queries at the cost of a slight increase in the total cost of evaluating all the views.

Keywords: Materialized Views, View Selection, Greedy Algorithm.

1 Introduction

Data warehouse stores subject oriented, integrated, time variant and non-volatile data to support processing of analytical queries [7]. These analytical queries, which are long and complex, consume a lot of time when processed against a large data warehouse. Further, the exploratory nature of these analytical queries contributes to high average query response time. This query response time can be reduced by materializing views over a data warehouse[9]. Materialized views contain pre-computed and summarized information, computed from the information stored in the data warehouse. They are significantly smaller in size, when compared with the data warehouse, and can significantly reduce the response time if they contain relevant and required information for answering analytical queries. The selection of such information and storing them as materialized view is referred to as the view selection problem[4]. View selection deals with selecting appropriate set of views that provide answers to most of the future queries. View selection is formally defined in [4] as “Given a database schema D , storage space B , Resource R and a workload of queries Q , choose a set of views V over D to

materialize, whose combined size is at most B and resource requirement is at most R". The number of possible views is exponential in the number of dimensions and for higher dimensions it would become infeasible to materialize all views due to space constraints[6]. Further, the space and resource constraint translates the views selection problem into an optimization problem that is NP-Complete[6]. Alternatively, views can be selected empirically, based on past query patterns[12], or heuristically using algorithms that are greedy, evolutionary etc. This paper focuses on greedy based view selection.

Greedy based view selection, in each iteration, select the most beneficial view for materialization. Among the several greedy based algorithms presented in literature [1, 2, 3, 5, 6, 8, 10, 11, 13, 14], the algorithm in [6] is considered the most fundamental one. This algorithm, which hereafter in this paper would be referred to as HRUA, selects the top-T beneficial views, from amongst all possible views, in a multidimensional lattice. HRUA computes the benefit of a view in terms of its cost, which is defined in terms of the size of the view. HRUA computes benefit as given below:

$$\text{Benefit}_V = \sum \{ (\text{Size}(\text{SMA}(W)) - \text{Size}(V)) \mid V \text{ is an ancestor of view } W \text{ in the lattice and } (\text{Size}(\text{SMA}(W)) - \text{Size}(V)) > 0 \}$$

where $\text{Size}(V)$ = Size of view V

$\text{Size}(\text{SMA}(V))$ = Size of Smallest Materialized Ancestor of view V

HRUA uses size of the view to compute the benefit, It does not consider the number of queries that can be answered by a view, referred to as its query frequency. As a consequence, the views selected using HRUA may not be beneficial with respect to answering most of the future queries. As an example, consider a three dimensional lattice shown in Fig. 1(a). The size of the view in million (M) rows, and the query frequency (QF) of each view, is given alongside the view. Selection of Top-3 views using HRUA is shown in Fig. 1(b).

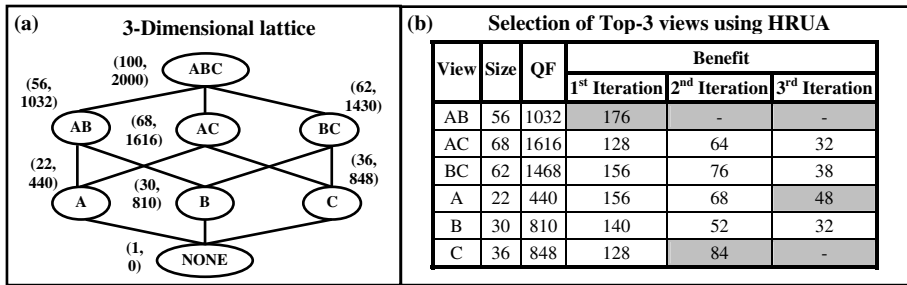


Fig. 1. Selection of Top-3 views using HRUA

HRUA selects AB, C and A as the Top-3 views. These selected views result in a Total View Evaluation Cost (TVEC) of 492. Considering the query frequency along with the size of each view, the Total Queries Answered (TQA) by the selected views AB, C and A is 3130, from among 6214 queries. An increase in this TQA value would result in more queries being answered by the selected views. The proposed algorithm aims to select Top-T profitable views for materialization that improves the TQA value by considering query frequency along with the size of each view. As a

result, the views selected would be able to answer a greater number of queries thereby improving the average query response time. The paper is organized as follows: The proposed algorithm is given in section 2 followed by experimental results in section 3. Section 4 is the conclusion.

2 Proposed Algorithm

Unlike HRUA, the proposed algorithm aims to select views that are not only profitable with respect to size but are also capable of answering greater number of queries. The proposed algorithm, in each iteration, considers the query frequency, along with the size of each view, to select the most profitable view for materialization. The query frequency of each view reflects past trends in querying and is computed as the number of queries, posed in the past, that can be answered by the view. The proposed algorithm, as given in Fig. 2, takes the lattice of views along with the size and query frequency of each view as input and produces the Top-T views as output.

```

Input: lattice of views L along with size and query frequency of each view
Output: Top-T views
Method:
  Let
   $V_R$  be the root view in the lattice,  $S(V)$  be the size of view  $V$ ,  $QF(V)$  be the query frequency of  $V$  in the lattice,
   $SMA(V)$  be the smallest materialized ancestor of  $V$ ,  $D(V)$  be the set of all descendent views of  $V$ ,  $MV$  be the set
  of materialized views,  $P(V) = \text{Profit of view } V$ ,  $P_M = \text{Maximum Profit}$ ,  $V_P = \text{View with maximum profit}$ 
  FOR  $V \in L$ 
     $SMA(V) = \text{RootView}$ 
  END FOR
  REPEAT
     $P_M = 0$ 
    FOR each view  $V \in (L - V_R \cup MV)$ 
       $V_P = V$ 
       $P(V) = 0$ 
      FOR each view  $W \in D(V)$  and  $(S(SMA(W)) - S(V)) > 0$ 
        
$$P(V) = P(V) + \left| \frac{QF(SMA(W)) - QF(V)}{S(SMA(W)) - S(V)} \right|$$

        END FOR
        IF  $P_M < P(V)$ 
           $P_M = P(V)$ 
           $V_P = V$ 
        END IF
      END FOR
       $MV = MV \cup \{V_P\}$ 
      FOR  $W \in D(V_P)$ 
        IF  $S(SMA(W)) > S(V_P)$ 
           $SMA(W) = V_P$ 
        END IF
      END FOR
    UNTIL  $|MV| < T$ 
  Return  $MV$ 

```

Fig. 2. Proposed Algorithm

The proposed algorithm, in each iteration, computes the profit of each view $P(V)$ as given below:

$$P(V) = \sum \left\{ \left| \frac{QF(SMA(W)) - QF(V)}{S(SMA(W)) - S(V)} \right| \mid V \text{ is an ancestor of view } W \text{ in the lattice and } (S(SMA(W)) - S(V)) > 0 \right\}$$

The profit of a view V is computed as the product of the number of dependents of V and the ratio of frequency difference between V and its smallest materialized ancestor and the size difference between V and its smallest materialized ancestor. The proposed algorithm (PA), in each iteration, computes profit of the as yet unselected views and selects, from amongst them, the most profitable view for materialization. The selection continues in this manner until T views are selected.

Let us consider the selection of the Top-3 views from the multidimensional lattice in Fig. 1(a) using PA. The selection of Top-3 views is given in Fig. 3.

View	Size	QF	Profit		
			1 st Iteration	2 nd Iteration	3 rd Iteration
AB	56	1032	88	-	-
AC	68	1616	48	24	24
BC	62	1468	56	28	28
A	22	440	40	35	-
B	30	810	34	17	9
C	36	848	36	27	18

Fig. 3. Selection of Top-3 views using PA

PA selects AB, A and BC as the Top-3 views. The views selected using PA has a TVEC of 480, which is less than TVEC of 492 due to views selected using HRUA. Also, the views selected using PA have a comparatively higher value of TQA of 4598 against the TQA of 3130 due to views selected using HRUA. Thus, it can be said that PA, in comparison to HRUA, is capable of selecting views that account for a greater number of queries at a lower total cost of evaluating all the views.

In order to compare the performance of PA with respect to HRUA, both the algorithms were implemented and run on data sets with varying dimensions. The experimental based comparisons of PA and HRUA are given next.

3 Experimental Results

The PA and HRUA algorithms were implemented using JDK 1.6 in Windows-XP environment. The two algorithms were experimentally compared on an Intel based 2 GHz PC having 1 GB RAM. The comparisons were carried out on parameters like TQA and TVEC for selecting the Top-10 views for materialization. The experiments were conducted by varying the number of dimensions of the data set from 5 to 10.

First, graphs were plotted to compare PA and HRUA algorithms on TQA versus number of dimensions. The graphs are shown in Fig. 4(a). It is observed from the graph that the increase in TQA, with respect to number of dimensions, is higher for PA vis-à-vis HRUA.

In order to ascertain the impact of higher TQA on TVEC due to views selected using PA, graphs for TVEC against number of dimensions were plotted and are shown in Fig. 4(b). It is evident from the graph that the TVEC of PA is slightly more than that of HRUA. This small difference shows that the PA selects views which are almost similar in quality to those selected by HRUA.

It can be reasonably inferred from the above that PA trades significant improvement in TQA with a slight increase in TVEC of views selected for materialization.

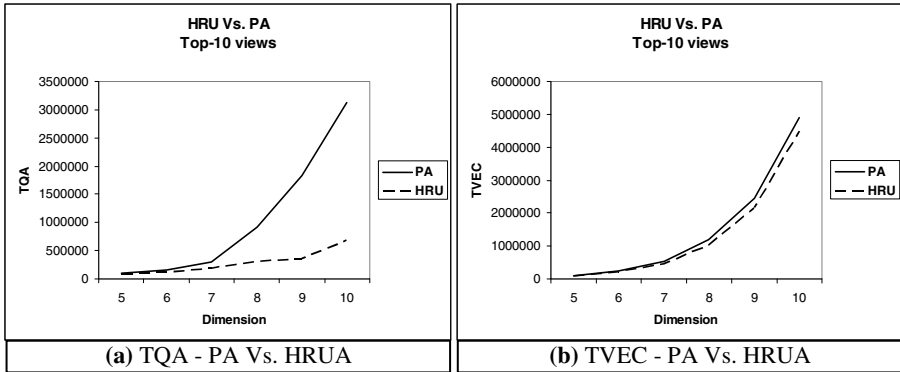


Fig. 4. PA Vs. HRUA – (TQA, TVEC) Vs. Dimensions

4 Conclusion

In this paper, an algorithm is proposed that selects Top-T views from a multidimensional lattice using both the size and the query frequency of each view. The proposed algorithm, in each iteration, computes the profit of each view using the size and query frequency of the views and then selects, from amongst them, the most profitable view for materialization. Unlike HRUA, the proposed algorithm is able to select fairly good quality views that are able to account for large number of queries. This would result in improvement in the average query response time.

The experiment based comparison of PA with HRUA on parameters TQA and TVEC showed that PA was found to achieve higher TQA at the cost of a slight increase in the TVEC in respect of views selected for materialization. That is, PA is able to select views capable of answering significantly greater number of queries at the cost of a slight drop in the quality of views selected for materialization.

References

1. Agrawal, S., Chaudhuri, S., Narasayya, V.: Automated Selection of Materialized Views and Indexes in SQL Databases. In: Proceedings of VLDB 2000, pp. 496–505. Morgan Kaufmann Publishers, San Francisco (2000)
2. Aouiche, K., Darmoni, J.: Data mining-based materialized view and index selection in data warehouse. *Journal of Intelligent Information Systems*, Pages, 65–93 (2009)
3. Baralis, E., Paraboschi, S., Teniente, E.: Materialized View Selection in a Multidimensional Database. In: Proceedings of VLDB 1997, pp. 156–165. Morgan Kaufmann Publishers, San Francisco (1997)
4. Chirkova, R., Halevy, A., Suciu, D.: A Formal Perspective on the View Selection Problem. *The VLDB Journal* 11(3), 216–237 (2002)
5. Gupta, H., Mumick, I.: Selection of Views to Materialize in a Data Warehouse. *IEEE Transactions on Knowledge and Data Engineering* 17(1), 24–43 (2005)

6. Harinarayan, V., Rajaraman, A., Ullman, J.: Implementing Data Cubes Efficiently. In: Proceedings of SIGMOD 1996, pp. 205–216. ACM Press, New York (1996)
7. Inmon, W.H.: Building the Data Warehouse, 3rd edn. Wiley Dreamtech, Chichester (2003)
8. Nadeau, T.P., Teorey, T.J.: Achieving scalability in OLAP materialized view selection. In: Proceedings of DOLAP 2002, pp. 28–34. ACM, New York (2002)
9. Roussopoulos, N.: Materialized Views and Data Warehouse. In: 4th Workshop KRDB 1997, Athens, Greece (August 1997)
10. Serna-Encinas, M.T., Hoya-Montano, J.A.: Algorithm for selection of materialized views: based on a costs model. In: Proceeding of Eighth International Conference on Current Trends in Computer Science, pp. 18–24 (2007)
11. Shah, A., Ramachandran, K., Raghavan, V.: A Hybrid Approach for Data Warehouse View Selection. *Int. Journal of Data Warehousing and Mining* 2(2), 1–37 (2006)
12. Teschke, M., Ulbrich, A.: Using Materialized Views to Speed Up Data Warehousing. Technical Report, IMMD 6, Universität Erlangen-Nürnberg (1997)
13. Vijay Kumar, T.V., Ghoshal, A.: A Reduced Lattice Greedy Algorithm for Selecting Materialized Views. *CCIS*, vol. 31, pp. 6–18. Springer, Heidelberg
14. Vijay Kumar, T.V., Haider, M., Kumar, S.: Proposing Candidate Views for Materialization. *CCIS*, vol. 54, pp. 89–98. Springer, Heidelberg (2010)

Key Validation Using Weighted-Edge Web of Trust Model

Sumit Kumar¹, Nahar Singh², and Ashok Singh Sairam¹

¹ Department of Computer Science & Engineering
Indian Institute of Technology Patna
Patna, India

sumit.itech@gmail.com, ashok@iitp.ac.in

² Department of Image Engineering
Chung-Ang University
Seoul, Korea Republic of
nahar.iitg@gmail.com

Abstract. Public key cryptography is widely used in the establishment of secure communication. An important issue in the use of public key cryptography is to ensure that a public key actually belongs to its owner. This problem is referred to as the key validation problem. One solution to this key validation problem is the centralized public key infrastructure (PKI) model ([1]). A second alternative is the decentralized Web of Trust model ([2]). In this model a user depends on digital certificate introduced by other users, who may be trustworthy or otherwise, to establish the binding between a public key and its owner. In this work we propose to model the trust relationship between users as weighted edges. We propose an algorithm to find trusted paths and thus validate public keys. The proposed algorithm is simple and has a low complexity both in terms of time and space.

Keywords: Web of Trust, Trust Hierarchy, Public key infrastructure, Key Validation Problem, PGP.

1 Introduction

The need to protect e-mail and data files has become very essential with the widespread availability of attack tools. Public key encryption allows users to securely communicate with each other over a non-secure channel without having to agree upon a shared secret key beforehand. In public-key encryption a user has two keys - a public key and a private key pair. The public key is made publicly available. Other users wanting to securely communicate with a user encrypt their data with the public key of the user. The encrypted data is decrypted by the user by using his private key which is only known to him. In public key cryptography an important problem is to establish that a public key actually belongs to its supposed owner. We refer to it as the key validation problem. A solution to the key validation problem is the use of public-key infrastructures (PKI) where one or several central authorities are responsible for issuing digital certificates for public keys. Such a certificate is a trusted warranty for the

binding between the involved public key and its owner. Another solution to the key validation problem is a distributed trust model called Web of Trust used in PGP, GnuPG and other OpenPGP compatible systems.

In the Web of Trust model the users themselves (also called clients) are able to issue public keys as introducers. The decision of whether to follow a trust endorsement depends on the individual users. Given a user and a key endorsement by a client, it is often possible for the user to find a chain of clients such that there exist bidirectional trust relationship between adjacent clients. In this paper we propose to design the trust relationship between users as weighted edges depending on the trust level of users. Our proposed algorithm facilitates to find a trusted path. The paper is organized as follows. Section 2 reviews the web of trust model and essential trust levels to validate destination. Section 3 is describing related work and their defects while the section 4 provides the details of proposed key validation using weighted edge Web of Trust model. Section 5 presents the theoretical analysis, while section 6 concludes.

2 The Web of Trust Model

The PGP trust model has some particular characteristics. First of all, (only) four levels of trust are supported: complete trust, marginal trust, no trust and legitimate. The owner of the key ring (Someone's personal collection of certificates is called key ring); who needs to manually assign these trust values for all other users, automatically receives full trust (also called implicit or Ultimate Trust). When a user places trust in an introducer, implicitly it means that the user possesses a certain amount of confidence in the introducer's capability to issue valid certificates, i.e. correct bindings between users and public keys. Therefore, Trust levels can be one of these:

1. **Undefined:** we cannot say whether this public key is valid or not.
2. **Marginal:** This public key may be valid be we cannot be too sure.
3. **Complete:** we can be wholly confident that this public key is valid.
4. **Legitimate:** Deemed legitimate by you.

Based on such trust values, the PGP trust model suggests accepting a given public key in the key ring as completely valid, if either

- (a) The public key belongs to the owner of the key ring,
- (b) The key ring contains at least C certificates from completely trusted introducers with valid public keys,
- (c) The key ring contains at least M certificates from marginally trusted introducers with valid public keys.

As we discussed before PGP Trust Models vary in two categories, one of them is Web of Trust. Web of trust is the encroachment of classical trust hierarchy model (HPKI). Web of trust supports graph based structure consequently it is a de-centralized approach. In Web of Trust each client able to issue the authentication certificate as introducer. The issuers of such certificates are called introducers, who can make them publicly available. All clients placed in the graph as nodes which able to authorize another node. There may be multiple paths between two entities & bi-directional trust

relationship flow of trust in a two directions (node to node). Web of trust is not more scalable & give difficulties in path discovery due to graph based structure.

3 Related Work

- Proposed model solves major part of key validation problem like limited trust level, limited validity levels, counter-intuitive key validation and hidden key dependencies. Proposed model gives strong assurance to trust someone if it covered all necessary conditions of approach.
- Probabilistic key validation approach [3, 4] has reliability problems such as how would you map trust with continuously varying probability values beside of this there may be complex mathematical computation to compute key validity but proposed weighted graph model is much reliable to compute trust with simple calculations.
- Self organized key Based Trust [5, 6] is superior than other existing approaches but its main concerns is only to have updated relationship status, but proposed model able to adopt the problem start point to end.

4 The Proposed Weighted Edge Trust Model

In order to resolve the defects of existing methods, we put forward a new proposal as “Weighted Edge Trust Model for PGP”.

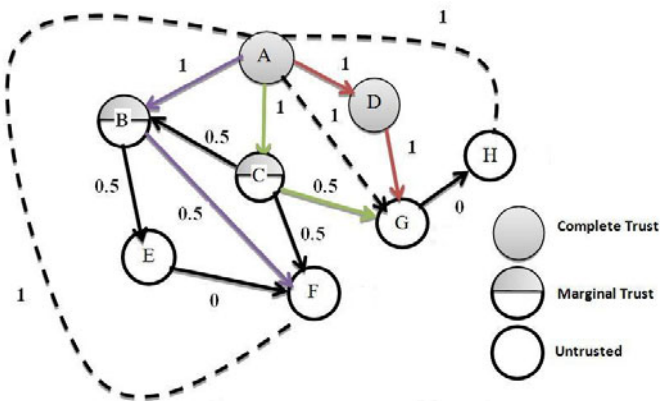


Fig. 1. Weighted Web of Trust Model Scenario

In general setting, the starting point of scenario as a “weighted edge graph”, where graph represent the key ring, vertices are introducers, edges are certificates. If an edge comes from a complete trust node, its weight will be ‘1’ and the key introduced is said to be complete trust. If edge comes from a marginal trust node, its weight will be ‘0.5’ (or less than 1 greater than 0) and the key introduced is said to be

marginal trust. If an edge comes from a un-trusted node, its weight will be '0' and the key introduced is said to be not-trusted, as shown in Figure 1.

In this approach when a user D wants to validate the public key of S it will execute the following algorithm:

Algorithm:

```

1.   For each existing path  $S$  to  $D$  (prefer shortest path first).
2.   {
3.   if (weight=  $N$ ) then
4.   {
5.       key trust is complete;
6.       CreateNewEdge( $S$ ,  $D$ );
7.   }
8.   if(weight<  $N$ )
9.   {
10.      if(CheckEdgeWeight( $E$ )!=0 && CheckWeight( $D$ )>=1)
11.          {
12.              if ((exist alternate path  $S$  and  $D$ ) && CheckEdgeWeight( $E$ )!=0)
13.                  {
14.                      key trust is complete;
15.                      CreateNewEdge( $S$ ,  $D$ );
16.                  }
17.          }
18.      Else
19.          {
20.              key trust is incomplete or un-trusted;
21.          }
22.      }
23.  }
```

- N = no. of edges between source & destination.
- Weight = total weight from source to destination (by adding weight of each existing edge in path).
- CheckWeight () = a function which calculate the weight of each incoming edges at the each node in path (also consists destination node).
- CheckEdgeWeight () =checks the weight of individual edges.
- CreateNewEdge (S , D) = create new edge between source(S) to destination (D) with weight = 1: where trust is complete.
- Source always will be complete trust.
- E =set of edges in path
- V =set of nodes in path
- Dotted edges created by CreateNewEdge (S , D) with weight '1'.

The input to the algorithm will be the weighted edge graph as shown in fig 1. In order for a node (referred to as source) wants to validate the public key of another node (referred to as destination) it will first try to find a path between them. There are three possibilities. The first case can be if total weight of the edges between the two is equal to no. of edges then it means all existing node in the path must be completely trusting

except the last node. Then key introduced by this path will be complete trust. As shown in fig.1, source A to destination G through D have all edges 1. Hence we can create a new edge (*CreateNewEdge()*) between them. This new edge will decrease the time complexity for future communication between the nodes.

The second case can be if total weight between sources to destination path is less than no. of edges & the weight is an integral value then it means there must be at least one un-trusted node in the path. As shown in fig.1, the path from source A to destination H through D , G has a total weight of 2 but the number of edges is 3. Such a path is an incomplete trust path. The third possibility is that the total weight between sources to destination path is less than no. of edges but none of the edges have a weight 0. The weight of the path will not be an integer value. Such a path may be valid or invalid. If we can find another path between the source and destination with exactly identical conditions than the source can be validated else it will be invalid. For example in fig.1, the path from A to F through B has a weigh of 1.5 but the total number of edges is 2 and there are no edges with 0 weight. However there is another path from A to F through C with weight of 1.5 and no edges with 0 weight. Thus F can validate A .

5 Analysis of Proposed Algorithm

To analyze the efficiency of proposed algorithm, we apply the web of trust probabilistic key validation method [7] in the scenario shown in figure 1. Assume client F wants to validate A 's public key. The key validation in the probabilistic model consists of two phases:

Phase 1: Determine all certificate paths leading from A to F .

Path1:[A, B, E, F], Path2:[A, B, F], Path3:[A, C, F]

Phase 2: Compute the probability that at least one operating path exists.

Let presume probabilities of nodes are as follow:

$A=1.0, B=0.4, C=0.5, D=0.9, E=0.1, F=0.2, G=0.1, H=0.2$.

The probability of a single path is simply the product of its (stochastically independent) trust values and for the overall probability of the set minpath (A, F), we can apply the so-called inclusion-exclusion formula, so probability of Path1, Path2, and Path3 are 0.04, 0.4, and 0.5 respectively. Exact result (hard to compute): key validity of $F = P(\text{path1, path2, path3, path4}) = 0.488$ (approx).

The result of whether F validates A certificate will depend on F own validation policy. For example if F specifies the validity threshold as 0.5 then the key of A will not be accepted.

On the other hand our proposed deterministic proposal will give a uniform result and will not depend on the client's validation policy. Moreover, the proposed algorithm has a lower complexity as can be seen below.

- **Time Complexity:** Probabilistic key validation models calculate the trust at each intermediary node so if a scenario has n number of nodes than the worst case complexity would be $O(n)$. In proposed model time complexity will be minimum after creating the all possible edges it will be $O(1)$.

- **Space Complexity:** Probabilistic web of trust models need to store information about intermediary nodes but in proposed model, information of them would be useless after creating direct edge consequently space complexity will decrease. And formation of direct edges gives better path discovery.
- **Ambiguity:** Ambiguity is a major problem of probabilistic models. Probabilistic model has to decide probability ranges to make trust. It is quite ambiguous that what probabilities make sure complete trust but in proposed model, there is no ambiguity trust in path due to frequent calculation.

6 Conclusion

Establishment of a hierarchical PKI is costly and lacks dynamicity. Further PKI being a centralized approach there is the possibility of a single point of failure and the system doesn't allow users to make their own decisions. On the other hand, the decentralized Web of Trust model is more scalable and do not suffer from the limitations of PKI. In the Web of Trust model any client can act as an *introducer* and issue digital certificates. However, this introducers or clients may not be completely trustworthy. In this paper we propose to model the trust relationship between clients deterministically using a weighted edge graph. We propose a simple algorithm whereby a user can validate digital certificates issued by the introducers.

References

1. Abdul-Rahman, A.: The PGP Trust Model: EDI-Forum (April 1997)
2. Zimmermann, P.R.: The Official PGP User's Guide. MIT Press, Cambridge (1994)
3. Cristina, S., Rafael, P., Jordi, F.: PROSEARCH: A Protocol to Simplify Path Discovery in Critical Scenarios. In: López, J. (ed.) CRITIS 2006. LNCS, vol. 4347, pp. 151–165. Springer, Heidelberg (2006)
4. Rolf, H., Jacek, J.: A New Approach to PGP's Web of Trust: ENISA European eIdentity conference, Paris (2007)
5. Hideaki, K., Osamu, M., Hiroaki, N., Hiroshi, I.: Self-Organized Key Management based on Trust Relationship List: Course of Computer and Communications. Graduate School of Engineering, Tokai University
6. Ross, D.: Pretty Good Privacy (1998-2008)
7. Jacek, J., Markus, W., Rolf, H.: A Probabilistic Trust Model for GnuPG. In: 23C3, 23rd Chaos Communication Congress, pp. 61–66 (2006)

A Novel Reconfigurable Architecture for Enhancing Color Image Based on Adaptive Saturation Feedback

M.C. Hanumantharaju¹, M. Ravishankar¹, D.R. Rameshbabu², and S. Ramachandran³

¹ Department of Information Science & Engineering,
{mchanumantharaju, ravishankarmcn}@gmail.com

² Department of Computer Science & Engineering,
Dayananda Sagar College of Engineering, Bangalore, India
bobramysore@gmail.com

³ Department of Electrical & Electronics, SJB Institute of Technology, Bangalore, India
ramachandar@gmail.com

Abstract. In this paper, a novel architecture suitable for Field Programmable Gate Array (FPGA) implementation of an adaptive color image enhancement based on Hue-Saturation-Value (HSV) color space is presented. The saturation feedback is used in order to enhance contrast and luminance of the color image. The saturation component is enhanced by stretching its dynamic range. Hue is preserved in order to avoid color distortion. The adaptive luminance enhancement is achieved by using a simple arithmetic mean filter. An efficient architecture for histogram equalization is also developed in order to evaluate the performance of proposed algorithm. The algorithm is implemented on Xilinx Vertex II XC2V2000-4ff896 FPGA device. The pipelining and parallel processing techniques have been adapted in order to speed up the enhancement process. The experimental results show that the color images enhanced by the proposed algorithm are clearer, vivid and efficient.

Keywords: FPGA, Adaptive Color Image Enhancement, Saturation Feedback, HSV, Arithmetic Mean Filter.

1 Introduction

Digital Image Enhancement [1] refers to accentuation, sharpening of image features such as edges, boundaries, or contrast to make a graphic display more useful for display and analysis. The color image enhancement can be classified into two categories according to the color space: (i) Color Image Enhancement in RGB color Space and (ii) Color Image Enhancement based on Transformed space [2].

HSV color space discriminates between color and intensity and hence these spaces reconstruct better images than is possible with the RGB space. In this work, HSV color space is chosen since it offers good image enhancement. Further, arithmetic mean filter [3] has been adopted that achieves very good quality reconstructed images. The algorithm is implemented on Xilinx Vertex II XC2V2000-4ff896 FPGA device. The pipelining and parallel processing techniques have been adapted in order

to speed up the enhancement process. It is the first of its kind in the literature, which FPGA implementation of adaptive color image enhancement [4] based on HSV color space has been proposed.

2 Proposed Method

This paper proposes a new adaptive color image enhancement based on arithmetic mean filter [5] in order to improve the quality of an image. The arithmetic means for luminance and saturation may be expressed by Eqns. (1) and (2):

$$\bar{V}_w = \sum_{(i,j) \in w} V(i,j) \quad (1)$$

$$\bar{S}_w = \sum_{(i,j) \in w} S(i,j) \quad (2)$$

where $m = 3$ and $n = 3$ for 3×3 window. The local variance for luminance and saturation are now expressed as: (3) and (4)

$$\sigma_v^2(x,y) = \sum_{(i,j) \in w} [V(i,j) - \bar{V}_w]^2 \quad (3)$$

$$\sigma_s^2(x,y) = \sum_{(i,j) \in w} [S(i,j) - \bar{S}_w]^2 \quad (4)$$

Further, the local correlation coefficient of luminance and saturation follows the Eqn. (5).

$$\rho(x,y) = \frac{\sum_{(i,j) \in w} [V(x,y) - \bar{V}_w(x,y)][S(x,y) - \bar{S}_w(x,y)]}{\sqrt{\sigma_v^2(x,y)\sigma_s^2(x,y)}} \quad (5)$$

The new luminance enhancement with saturation feedback is given by the Eqn. (6).

$$V_{enh}(x,y) = V(x,y) + k_1[V(x,y) - \bar{V}(x,y)] - k_2[S(x,y) - \bar{S}(x,y)] \times \rho(x,y) \quad (6)$$

The value of k_1 and k_2 was arrived at a value of 2 after conducting elaborate experiments. In order to improve the whole effect of color image with brighter and richer color the saturation component is enhanced by stretching its dynamic range. The mathematical model for saturation component enhancement is given by Eqn. (7).

$$S_{enh} = S^\gamma \quad (7)$$

where S represents original saturation component and S_{enh} is the enhanced saturation component. The value of gamma was arrived at a value of 0.77 after experiment analysis. In this work, hue is preserved in order to avoid color distortion.

3 Hardware Implementation

The Fig. 1 shows the block diagram for FPGA implementation adaptive color image enhancement system [6]. The proposed system has been realized using Register Transfer Level (RTL) compliant, Verilog Hardware Description Language (HDL). In the proposed work, the image size is chosen as 256×256 pixels.



Fig. 1. Block diagram of the Proposed Adaptive Image Enhancement System

The first step in the proposed scheme is RGB to HSV conversion. This module transforms pixels from RGB space to HSV. RGB to HSV conversion needs a total of 18 clock cycles to complete the conversion process. The digital hardware for RGB to HSV conversion is developed using Eqns. (8) to (10).

$$H = \begin{cases} 0 + \frac{43 \times |G - B|}{\text{Max}(R, G, B) - \text{Min}(R, G, B)}, \text{Max}(R, G, B) = R \\ 85 + \frac{43 \times |B - R|}{\text{Max}(R, G, B) - \text{Min}(R, G, B)}, \text{Max}(R, G, B) = G \\ 171 + \frac{43 \times |R - G|}{\text{Max}(R, G, B) - \text{Min}(R, G, B)}, \text{Max}(R, G, B) = B \end{cases} \quad (8)$$

$$S = 255 \times \frac{\{\text{Max}(R, G, B) - \text{Min}(R, G, B)\}}{\text{Max}(R, G, B)} \quad (9)$$

$$V = \text{Max}(R, G, B) \quad (10)$$

The architecture for adaptive color image enhancement shown in Fig. 2 is developed using Eqns. (6) and (7) described in the previous section. The enhanced pixels are transformed back to RGB space using HSV to RGB converter module. The HSV to RGB converter module takes 23 clock cycles to complete the conversion process. In order to speed up the complete system pipelining and parallel processing technique has been adapted. The Fig. 3 shows the hardware architecture for histogram equalization. This module performs contrast enhancement [7] of an image. Histogram equalization [8] module includes special decoder (8 to 256), 256 counters (16-bit), and a memory 256 words (each word of 8-bits) followed by a combinational logic for pixel mapping. In order to realize the digital hardware for histogram equalization we use Eqn. (11)

$$S_k = \frac{(L-1)}{MN} \sum_{j=0}^k n_j \tag{11}$$

where S_k is the discrete transformation, L represents intensity levels in the image ($L=256$ for an 8-bit image), The product MN is the total number of pixels in the image.

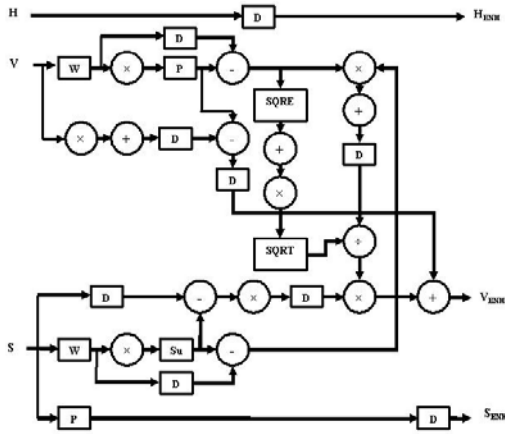


Fig. 2. Proposed Architecture for Adaptive Color Image Enhancement

H = Hue; S = Saturation; V = Value (Intensity); D = Delay; W = Window
 Su = Sum; SQRE = Square; SQRT = Square Root; H_{ENH} = Hue Enhanced
 S_{ENH} = Saturation Enhanced; V_{ENH} = Value (Intensity) Enhanced

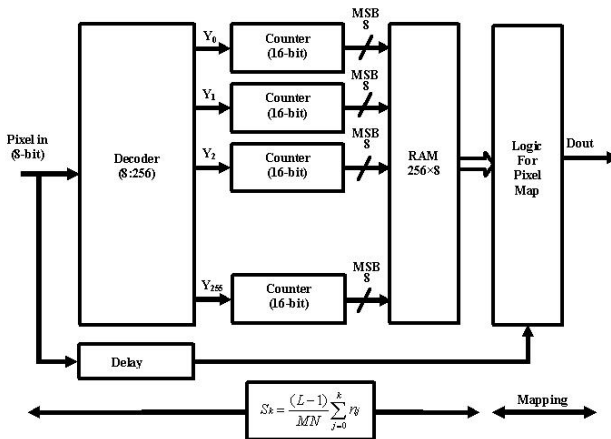


Fig. 3. Proposed Architecture for Histogram Equalization

4 Experimental Results and Comparative Study

The proposed FPGA implementation of adaptive color image enhancement [9] based on HSV color space has been coded and tested in Matlab (Version 8.1) first in order to ensure the correct working of the algorithm. Subsequently, the complete system has been coded in Verilog HDL so that it may be implemented on an FPGA or Application Specific Integrated Circuit (ASIC). The system simulation is done using ModelSim (Version SE 6.4) and synthesized using Xilinx ISE 9.2i. The algorithm is implemented on Xilinx Vertex II XC2V2000-4ff896 FPGA device. In this work, window size was chosen as 3×3 since the image looks better and more colorful than the original image. The experiment was conducted and presented in Fig. 4 by considering three poor quality test images.



Fig. 4. Experimental Results of Tree, Tractor and Squares

First Column: Original Image; **Second Column:** Histogram Equalized Image obtained from the Proposed Architecture shown in Fig. (3); **Third Column:** Enhanced Images obtained from proposed hardware architecture shown in Fig. (2).

The first column of Fig. 4 shows three different original color images. The second column of Fig. 4 shows the images enhanced by proposed histogram equalization of R, G and B channels in RGB color space. The result shows that histogram equalized images suffer from color distortion. The last column of Fig. 4 shows the images enhanced by proposed luminance enhancement algorithm based on saturation feedback.

Table 1. Summary of the FPGA Device Utilization XC2V2000-4ff896

Logic Utilization	Used	Available	Utilization
No. of Slice Flip Flops	3128	21504	14%
No. of 4 input LUTs	1436	21504	6%
Logic Distribution			
No. of Occupied Slices	3427	10752	31%
No. of Slices containing only related logic	3427	4792	71%
No. of Slices containing unrelated logic	0	4792	0%
Total No. of 4 input LUTs	7483	21504	34%
Total Equivalent Gate Count for Design	64312		

Table 1. shows the device utilization summary of FPGA implementation [10] adaptive color image enhancement based on adaptive saturation feedback. Experimental results show that the color image enhancement based on the proposed adaptive luminance and saturation feedback offers much better enhanced images than that of other methods. The reconstructed images offer richer color, clearer details and higher contrast. In order to evaluate the performance of the proposed method, we present contrast enhancement performance, luminance enhancement performance Peak Signal to Noise Ratio (PSNR) and histogram plots. The contrast enhancement performance C , Luminance enhancement performance L and PSNR is evaluated using the following Equations.

$$C = \frac{\sigma_{out} - \sigma_{in}}{\sigma_{in}} \quad (12)$$

$$L = \frac{I_{out} - I_{in}}{I_{in}} \quad (13)$$

$$PSNR = 10 \log_{10} \left[\frac{255^2}{MSE} \right] \quad (14)$$

The Mean Square Error (MSE) is given by

$$MSE = \sum_{x=1}^p \sum_{y=1}^q \frac{(E(x, y) - I(x, y))^2}{pq} \quad (15)$$

where σ_{out} and I_{out} are luminance variance and luminance mean values of the output image, σ_{in} and I_{in} are luminance variance and luminance mean values the input image, $E(x, y)$ and $I(x, y)$ are the enhanced and original gray pixel at position (x, y) , p and q denote the size of the gray image.

Table 2. Performance Comparison

Parameter	Histogram Equalization Method		
	Tree	Tractor	Squares
Contrast	0.45	0.48	0.63
Luminance	0.53	0.61	0.77
PSNR	34.7	36.7	39.1

Parameter	Proposed Method		
	Tree	Tractor	Squares
Contrast	0.41	0.44	0.64
Luminance	0.44	0.65	0.72
PSNR	32.5	33.6	34.2

Table 2. shows the performance comparison. It is clear that, the contrast enhancement performance of the proposed method is good as compared to histogram equalization. The luminance of the histogram equalization has increased; hence the image looks over enhanced leading to color distortion. The proposed method has better PSNR as compared to other methods.

5 Conclusion

This paper presented the adaptive color image enhancement and its FPGA implementation based on arithmetic mean filter. Arithmetic mean filter provides better visual effects as compared with other filters. The proposed scheme uses HSV color space since it discriminates between color and intensity. In addition, it offers good image enhancement. In this work, efficient hardware architectures for RGB to HSV, HSV to RGB, Adaptive image enhancement and histogram equalization has been developed. In order to increase the processing speed the techniques such as pipeline and parallel processing schemes has been adapted. The Verilog code developed for the complete system is RTL compliant and works for ASIC design. The picture resolution of 256×256 pixels is chosen for this scheme. The implementation presented in this paper has been realized on an FPGA. The experimental results show that the enhanced images are better in visual quality compared to other methods.

References

1. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Addison- Wesley, Reading (1992)
2. Jain, A.K.: Fundamentals of Digital Image Processing. Prentice-Hall, Englewood Cliffs (1989)
3. Thomas, B.A., Strickland, R.N., Heffrey, J.: Color Image Enhancement using Spatially Adaptive Saturation Feedback. In: IEEE International Conference on Image Processing, vol. 3, pp. 30–33 (1997)
4. Song, G., Qiao, X.-L.: Adaptive Color Image Enhancement based on Human Visual Properties. In: Proceedings of International Congress on Image and Signal Processing (2008)

5. Song, G., Qiao, X.-L.: Color Image Enhancement based on Luminance and Saturation Components. In: Proceedings of International Congress on Image and Signal Processing (2008)
6. Kokufuta, K., Maruyama, T.: Real-time Processing of Contrast Limited Adaptive Histogram Equalization on FPGA. In: Proceedings of IEEE International Conference on Field Programmable Logic and Applications, pp. 155–158 (2010)
7. He, K.-J., Chen, C.-C., Lu, C.-H., Wang, L.: Implementation of a New Contrast Enhancement Method for Video Images, pp. 1982–1987 (2010)
8. Xie, X., Shi, Z., Guo, W., Yao, S.: An Adaptive Image Enhancement Technique based on Image Characteristics. In: Proceedings of IEEE International Conference on Image and Signal Processing, pp. 1–5 (2009)
9. Zhang, M.Z., Seow, M.-J., Asari, V.K.: A Hardware Architecture for Color Image Enhancement Using a Machine Learning Approach with Adaptive Parameterization. In: Proceedings of IEEE International Joint Conference on Neural Network, pp. 35–40 (2006)
10. Ngo, H.T., Zhang, M.Z., Tao, L., Asari, V.K.: Design of a Digital Architecture for Real-Time Video Enhancement Based on Illuminance- Reflectance Model. In: Proceedings of 49th IEEE International Midwest Symposium on Circuits and Systems, pp. 286–290 (2006)

Signal Processing Approach for Prediction of Kink in Transmembrane α -Helices

Jayakishan K. Meher¹, Nibedita Mishra², Pranab Kishor Mohapatra³, Mukesh Kumar Raval⁴, Pramod Kumar Meher⁵, and Gananath Dash⁶

¹ Department of Electronics & Telecom Engineering, SITE, Balangir, Odisha, India -767002

jk_meher@yahoo.co.in

² Department of Chemistry, Rajendra College, Balangir, Odisha, India -767002

nibedita1976@yahoo.com

³ Department of Chemistry, CV Raman College of Engineering, Bhubaneswar, Odisha, India – 752054

pkmohapatra@yahoo.co.in

⁴ Department of Chemistry, Gangadhar Meher College, Sambalpur, Odisha, India – 768004

mraval@yahoo.com

⁵ Institute for Infocomm Research, Singapore – 138632

pkmeher@i2r.astar.edu.sg

⁶ School of Physical Sciences, Sambalpur University, Jyoti Vihar, Odisha, India – 768019

gndash@ieee.org

Abstract. The functions of transmembrane proteins are attributed by kinks (bends) in helices. Kinked helices are believed to be required for appropriate helix-helix and protein-protein interaction in membrane protein complexes. Therefore, knowledge of kink and its prediction from amino acid sequences is of great help in understanding the function of proteins. However, determination of kink in transmembrane α -helices is a computationally intensive task. In this paper we have developed signal processing algorithms based on discrete Fourier transform and wavelet transform for prediction of kink in the helices with a prediction efficiency of ~80%. The numerical representation of the protein in terms of probability of occurrence of amino acids constituted in kinked helices contains most of the necessary information in determining the kink location, and the signal processing methods capture this information more effectively than existing statistical and machine learning methods.

Keywords: DFT, Kink, Transmembrane α -Helices, Wavelet.

1 Introduction

Knowledge of segments of transmembrane proteins and the bends in helices help in the study of tertiary structure and hence understanding the role played by that protein. 20-30% of all the proteins in any organism are membrane proteins. These are of particular importance because they form targets for over 60% of drugs on the market. Transmembrane α -helix bundle is a common structural feature of membrane proteins

except porins, which contains β -barrels. Membrane spanning α -helices differ from their globular counterpart by the presence of helix breakers, Pro and Gly, in the middle of helices. Pro is known to induce a kink in the helix [1, 2]. A hypothesis suggests that Pro is introduced by natural mutation to have a bend and later further mutated leaving the bend intact for required function during the course of evolution [3]. The role of Pro and kinks in transmembrane helices were extensively investigated both experimentally and theoretically to unravel the nature's architectural principles [2,4]. Another observation suggest induction of kink at the juncture of α -helical and 3_{10} helical structure in a transmembrane helix [2-6]. Mismatch of hydrophobicity of lipid bilayer and peptide may also result in distortion of α -helical structure [7]. Sequences of straight and kinked helices were further subjected to machine learning to develop a classifier for prediction of kink in a helix from amino acid sequences. Support vector machine (SVM) method [8] projects that helix breaking propensity of amino acid sequence determines kink in a helix. Kinked and straight helix of protein Type-4 Pilin and Chlorophyll A-B binding protein respectively are shown in Fig.1.

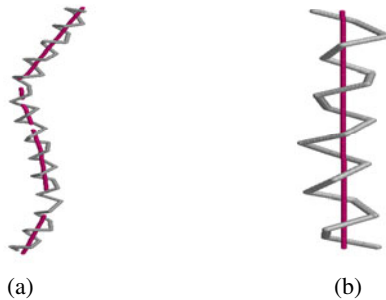


Fig. 1. Backbone representation of (a) Kinked helix of Type-4 Pilin protein (2pil) (b) Straight helix (second helix) of Chlorophyll A-B binding protein (1rwt)

A kink in a helix may be formed by helix-helix interaction. In such cases the intrinsic kink forming or helix breaking tendency may not be required. Even a helix forming tendency may be overridden. This possibility clamps a theoretical limit to predict a kink with high accuracy. Hence there is a need to develop advanced algorithm for faster and accurate prediction of kink in transmembrane helices. This motivates to develop novel approach based on signal processing methods such as DFT and wavelet transform to effectively predict kink in transmembrane α -helices.

2 Materials and Methods

Database. List of transmembrane proteins and their coordinate files were obtained from the Orientation of Proteins in Membranes (OPM) database at College of Pharmacy, University of Michigan (<http://www.phar.umich.edu>).

Determination of α -helical regions. Dihedral angles were computed using MAP-MAK from coordinate files and listed for each residue along with assignment of conformational status of the residue namely right or left helical, β -strand. Molecular

visual tools RasMol were used to visually confirm the transmembrane α -helical regions.

Computation of helix axis. Helix axis was computed from the approximate local centroids $\theta_i^0(x_i^0, y_i^0, z_i^0)$ of the helix by taking a frame of tetrapeptide unit [9].

$$x_i^0 = \frac{1}{4} \sum_i^{i+3} x_i, \quad y_i^0 = \frac{1}{4} \sum_i^{i+3} y_i, \quad z_i^0 = \frac{1}{4} \sum_i^{i+3} z_i \quad (1)$$

where x_i , y_i , and z_i are the coordinates of $C\alpha$ atoms of the tetrapeptide frame. Unit vector in the direction of resultant of vectors $\theta_i^0, \theta_{i+1}^0$ yields direction cosines (l, m, n) of axis of helix (A). The axis pass through the centroid of the helix $\theta^0 = (X^0, Y^0, Z^0)$.

$$X^0 = \frac{1}{n} \sum_{i=1}^n x_i, \quad Y^0 = \frac{1}{n} \sum_{i=1}^n y_i, \quad Z^0 = \frac{1}{n} \sum_{i=1}^n z_i \quad (2)$$

where n is the number of residues in a helix. Refined local centers θ_i of helix are then calculated for each $C\alpha$ by computing the foot of perpendicular drawn from $C\alpha_i$ to A .

Location of hinges. Hinges were located in a helix by a distance parameter $d(C_i N_{i+4})$, where C_i is the backbone carbonyl carbon of i^{th} residue and N_{i+4} is backbone peptide nitrogen of $i+4^{\text{th}}$ residue [9]. Value of $d(C_i N_{i+4})$ beyond the range $4.227 \pm 0.35 \text{ \AA}$ reflects a hinge at the i^{th} residue in the helix. Hinge was quantified by two parameters kink and swivel [3].

3 Proposed Signal Processing Methods for Kink Prediction

Signal processing methods such as Fourier transform and wavelet transform can identify periodicities and variations in signals from a background noise. This property of signal processing approach plays a major role in prediction of helix kink in amino acid sequence. In this paper the presence of kink in amino acid sequence is determined effectively in transform domain analysis.

In our approach, helix forming propensity of amino acid residues in a sequence window of nine residues is taken as an input vector represented as $x_i \in R^9$, where $i = 1, 2, \dots, N$ and N represents the number of samples. The fifth residue of the sequence exhibits kink in the sequence of nine in case of kinked helix. A sequence of $i-4$ to $i+4$ is selected because to determine a kink at i^{th} residue axes of $i-4$ to i and i to $i+4$ is necessary and minimum five residues are required to determine axis of a helix. Hence, a sequence of nine residues is minimum requirement for determination of kink. The sequence x_i is converted to numerical representation. The window of 9 residues slides across the sequence and at each position an average value in the transform domain is calculated and assigned to the middle residue. The plot of the average value of each window against the relative residue location is found to exhibit peaks at kink locations which indicate the presence of a kink region in the amino acid sequence and no such peak at straight helices. Based on this method two transformed domains such as DFT and wavelet transform are demonstrated in this paper.

DFT Based Approach. Discrete Fourier transform plays a role in the implementation of digital signal processing algorithm for kink prediction. It is found that DFT of a given input amino acid sequence having kink helices exhibits peaks at the kink locations thus detecting periodicity in the sequence. The amino acid sequence of 9 characters is converted into numerical sequence by substituting its probability of occurrence of residues (Table 1) from a known 500 dataset of OPM database. Alternately physicochemical property such as polarizability as shown in Table 1 that has been computed from Hyperchempro 8.0 software play also similar role.

Table 1. Physico-chemical properties of amino acids

Residues	Probability of occurrence	Polarizability	Residues	Probability of occurrence	Polarizability
A	0.8492	4.44	L	1.0908	9.95
R	0.6979	14.16	K	1.4422	10.72
N	1.1960	7.72	M	1.0261	11.11
D	0.8988	6.55	F	1.2654	14.10
C	1.2513	7.44	P	10.000	8.79
Q	0.8762	11.39	S	0.8676	5.08
E	1.6414	8.38	T	0.7235	6.92
G	0.9859	2.61	W	1.0967	19.37
H	0.8137	11.84	Y	0.898	14.74
I	0.6388	9.95	V	0.8624	8.11

For example, the amino acid sequence $x(n) = [\text{SWWNFGSLL}]$, the corresponding numerical sequence $x_i(n)$ substituting probability of occurrence of residues is $x_i(n) = [0.8676 \ 1.0967 \ 1.0967 \ 1.1960 \ 1.2654 \ 0.9859 \ 0.8676 \ 1.0908 \ 1.0908]$. The DFT of an N -point of sequence $x_i(n)$ is defined as

$$X(k) = \sum_{n=0}^{N-1} x_i(n) \cdot e^{-j2\pi kn/N}, \quad 0 \leq k \leq N-1 \quad (3)$$

where N is the length of the segment of amino acid sequence. The window slides across the sequence and at each position an average DFT value for the nine residues is calculated and assigned to the middle residue. The frequency spectrum of $X(k)$ is found to exhibit peaks at kink locations which indicate the presence of a kink region in the amino acid sequence. The plot shows that there are remarkable peaks in kinks (bend) helices and no such peak in straight (nonbend) helices as shown in Fig.2 and Fig.3. This unique property is used to predict the kink in amino acid sequence.

Wavelet transform based approach. Recently, the use of wavelet transform, both continuous and discrete in the Bioinformatics field is promising. Continuous Wavelet Transform (CWT) allows one-dimensional signal to be viewed in a more discriminative two-dimensional time-scale representation. CWT is calculated by the continuous shifting of the continuously scalable wavelet over the signal. In discrete wavelet transform (DWT) a subset of scales and positions are chosen, in which the correlation between the signal and the shifted and dilated waveforms are calculated. Consequently, the signal is decomposed into several groups of coefficients, each containing signal features corresponding to a group of frequencies. Small scales refer to compressed

wavelets, depicted by rapid variations appropriate for extracting high frequency features of the signal. An important attribute of wavelet methods is that, due to the limited duration of every wavelet, local variations of the signal are better extracted and information on the location of these local features is retained in the constituent waveforms.

DWT has been applied on hydrophobicity signals in order to predict hydrophobic cores in proteins [10]. Protein sequence similarity has also been studied using DWT of a signal associated with the average energy states of all valence electrons of each amino acid [11]. Wavelet transform has been applied for transmembrane structure prediction [12]. In this work, the wavelet transform is used to determine kink in segments of amino acid sequences of α -helical membrane proteins.

A wavelet is a waveform that is localised in both time and frequency domains. This wavelet is dilated and translated along the signal to perform the analyses. The commonly used wavelets in practice are Haar, Daubechies, Gaussian wave, Mexican hat and Morlet wavelets. The selection of particular wavelet for any analysis depends on the kind of signal being studied and kind of signal variation to be captured. In case of analysis of protein sequence signal the Mexican hat wavelet seemed to be choice. The mother wavelet of Mexican hat wavelet is defined as

$$\psi(t) = (1-t^2)e^{-t^2/2} \quad (4)$$

The wavelet transform gives rise to patterns that are distinct between the kink regions from straight helix regions. To analyze the protein sequence for prediction of kink, it is first transformed into a numerical signal based on probability of occurrence of residues along a protein sequence.

A sliding window of 9 residues has been used. Wavelet coefficients are computed by translating the wavelet along the signal. The window slides across the sequence shifted by one residue with the window size of nine residues. At each position an average coefficients for the nine residues is calculated and assigned to the middle residue. The length of the signal produced across the protein sequence is equal to the number of residues of the protein. The plot of average value against relative residues is found to exhibit peaks at kink locations that indicate the presence of kink regions in the amino acid sequence of transmembrane α -helices as shown in Fig. 2 and Fig.3.

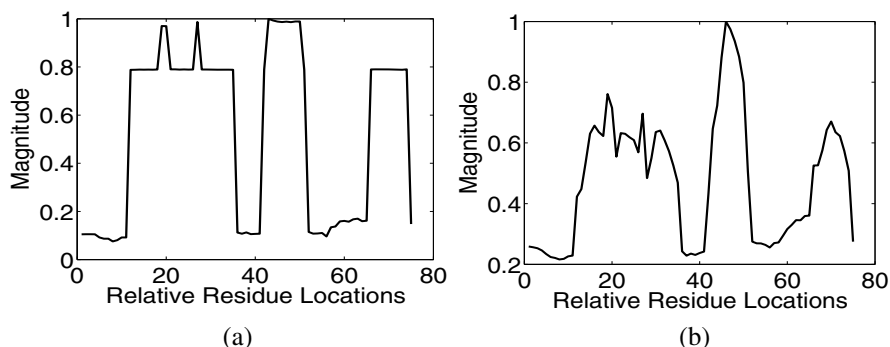


Fig. 2. Cytochrome BC1 Complex Protein (pdb id: 1BGY) showing four kinked helices including two overlapped kinks by (a) DFT, (b) Wavelet transform

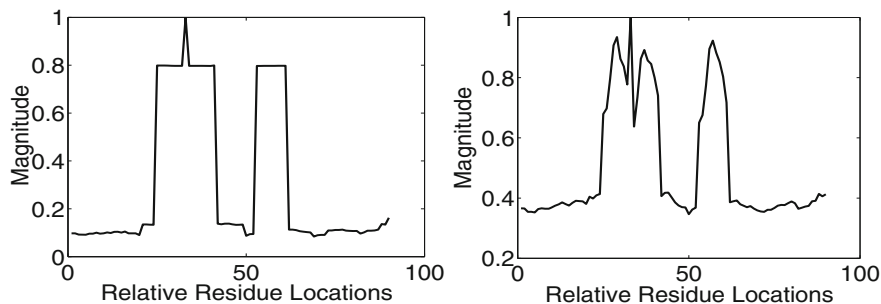


Fig. 3. Potassium Channel protein (pdb id: 1BL8) showing three kink helices including two overlapped kinks by (a) DFT, (b) Wavelet transform.

4 Result and Discussion

We have used the DFT and wavelet transform techniques to detect the kink locations using numerical representation based on probability of occurrence of amino acid residues and polarizability, physico-chemical properties obtained from Hyperchempro 8.0 software of HyperCubeInc, USA. 200 proteins data sets used as bench mark for this purpose. In a good number of cases all the proposed methods performed well. Helix forming propensity of amino acid residues in a sequence window of nine residues is taken as an input vector of dimension 9. The fifth residue of the sequence exhibits kink in the sequence of nine in case of kinked helix. For example in protein Cytochrome BC1 Complex with accession no: 1JBY, proposed method using DFT and wavelet transform have detected all four kinked helices including two overlapped kinks as shown in Fig. 2. Again in Potassium Channel protein with Accession No: 1BL8, three kink helices including two overlapped kinks are identified by the proposed methods as shown in Fig. 3.

The list of proteins under study and the corresponding results obtained using proposed methods are shown in Table 2. It is found that the proposed methods show better accuracy in predicting kink helices. The performance analysis of various methods can be made by prediction measures such as accuracy (A), precision (P) and recall (R) which are defined in terms of four parameters true positive (t_p), false positive (f_p), true negative (t_n) and false negative (f_n). t_p denotes the number of actual kinks and are also predicted as kinks, f_p denotes the number of actually straight helices but are predicted to be kinks, t_n is the number of actually straight helices and also predicted to be straight helices, and f_n is the number of actually kinks and predicted to be straight helices.

Accuracy. The accuracy of prediction kink in amino acid sequence is defined as the percentage of helices correctly predicted of the total helices present. It is computed as:

$$A = \frac{\text{Number of correct predictions}}{\text{Total number of helices}} ; \quad A = \frac{t_p + t_n}{t_p + f_p + t_n + f_n} \quad (5)$$

Precision. Precision is computed for kink and straight helix classes separately. It is defined as the percentage of kinks correctly predicted to be one class of the total kink predicted to be of that class. Precision is computed as:

$$P = \frac{\text{Number of correctly predicted kinks}}{\text{Total number of kinks predicted}} ; P = \frac{t_p}{t_p + f_p} \quad (6)$$

Recall. Recall is also computed separately for kink and straight helix classes. It is defined as the percentage of the kinks that belong to a class that are predicted to be that class. Recall is computed as:

$$R = \frac{\text{Number of correctly predicted kinks}}{\text{Total number of actual kinks}} ; A = \frac{t_p}{t_p + f_n} \quad (7)$$

Table 2. Accuracy comparisons of various techniques of OPM protein data set

PDB Id., Protein Name	Methods	Prediction measures		
		A	P	R
1BL8, Potassium channel protein	SVM	0.75	1	0.75
	DFT	.8	1	0.8
	DWT	.8	1	0.8
2PIL, Type-4 Pilin,	SVM	0.75	0.75	1
	DFT	0.8	0.8	1
	DWT	0.8	0.8	1
1BGY, Cytochrome BC1 complex	SVM	0.8	1	0.8
	DFT	0.88	1	0.88
	DWT	0.88	1	0.88
1JBO, C-Phycocyanin Alpha chain	SVM	0.66	1	0.66
	DFT	0.75	1	0.75
	DWT	0.75	1	0.75
1NEK, Succinate Dehydrogenase Flavoprotein	SVM	0.75	1	0.75
	DFT	0.75	0.75	1
	DWT	0.75	0.75	1
2BL2, V-Type Sodium ATP synthase SUB-UNIT K	SVM	0.8	1	0.8
	DFT	88	1	0.88
	DWT	88	1	0.88

The result suggests that non-helix former amino acids may induce kink. But no such correlation with single amino acid residues at kink is observed, rather helix formers namely Ala, Leu, Ile, Phe, more frequently occur at kink position. However, at the $i+4^{th}$ position Pro and Gly occur with higher frequency. Again no correlation is observed between propensity for helix formation of $i+4^{th}$ and frequency of its occurrence in kinks at $i+4^{th}$ position. Combinations of residues in a definite sequence may be responsible for non-Pro kinks in helices [13]. It appears that there may be a non-linear correlation between helix forming tendency of amino acid residues and formation of kink.

5 Conclusion

Signal processing approach plays a vital role in the prediction of kink in transmembrane α -helix. The proposed method is not only fast but also has improved accuracy (more than 80%) as compared to SVM learning system reported by us earlier [8]. However prediction of kink in a helix depends on the features of amino acid sequence. Feature vector with probability of occurrence of residues and polarizability are only used for numerical representation in the present study. Although kink prediction has its own limitations, the present work is the first report in the area of helix kink prediction from amino acid sequence based on signal processing algorithms.

References

1. Ramachandran, G., Ramakrishnan, C., Sasisekharan, V.: Stereochemistry of Polypeptide Chain Configuration. *J. Mol. Biol.* 7, 95–97 (1963)
2. Sankararamkrishnan, R., Vishveshwara, S.: Conformational Studies on Peptides with Proline in the Right-Handed-Helical Region. *Biopolymers* 30, 287–298 (1990)
3. Cordes, F., Bright, J., Sansom, M.P.: Proline Induced Distortions of Transmembrane Helices. *J. Mol. Biol.* 323, 951–960 (2002)
4. von Heijne, G.: Proline Kinks in Transmembrane-Helices. *J. Mol. Biol.* 218, 499–503 (1991)
5. Yohannan, S., Faham, S., Whitelegge, J., Bowie, J.: The Evolution of Transmembrane Helix Kinks and the Structural Diversity of G-protein Coupled Receptors. *Proc. Natl. Acad. Sci. U.S.A.* 101, 959–963 (2004)
6. Pal, L., Dasgupta, B., Chakrabarti, P.: 3(10)-Helix Adjoining Alpha-helix and Beta-strand: Sequence and Structural Features and Their Conservation. *Biopolymers* 78, 147–162 (2005)
7. Daily, A., Greathouse, D., van der Wel, P., Koeppe, R.: Helical Distortion in Tryptophan- and Lysine-Anchored Membrane-Spanning Alpha-Helices as a Function of Hydrophobic Mismatch: A Solid-State Deuterium NMR Investigation using the Geometric Analysis of Labeled Alanines Method. *Biophys. J.* 94, 480–491 (2008)
8. Mishra, N., Khamari, A., Mohapatra, P.K., Meher, J.K., Raval, M.K.: Support Vector Machine Method to Predict Kinks in Transmembrane Helices, pp. 399–404. Excel India Publishers, India (2010)
9. Mohapatra, P.K., Khamari, A., Raval, M.K.: A Method for Structural Analysis of Helices of Membrane Proteins. *J. Mol. Model.* 10, 393–398 (2004)
10. Hirakawa, H., Muta, S., Kuhara, S.: The Hydrophobic Cores of Proteins Predicted by Wavelet Analysis. *Bioinformatics* 15, 141–148 (1999)
11. de Trad, C., Fang, Q., Cosic, I.: Protein Sequence Comparison Based on the Wavelet Transform Approach. *Protein Eng.* 15, 193–203 (2002)
12. Murray, K.B., Gorse, D., Thornton, J.: Wavelet Transforms for the Characterization and Detection of Repeating Motifs. *J. Mol. Biol.* 316, 341–363 (2002)
13. Chou, K., Nemethy, G., Scheraga, H.: Energetic Approach to the Packing of Helices. 2. General treatment of Nonequivalent and Nonregular Helices. *J. Am. Chem. Soc.* 106, 3161–3170 (1984)

Cascaded H-Bridge Multilevel Boost Inverter without Inductors for Electric/Hybrid Electric Vehicle Applications

S. Dhayanandh¹, A.P. Ramya Sri¹, S. Rajkumar¹, and N. Lavanya²

¹ Assistant Professor, Department of ECE,
Kathir College of Engineering
Coimbatore, India

dhais@sify.com, ramyaashree123@gmail.com, rajped22@gmail.com

² Assistant Professor, Department of ECE,
Maharaja Engineering College
Coimbatore, India
lavanya.nataraj87@gmail.com

Abstract. This paper presents a cascaded H-bridge multilevel boost inverter for electric vehicle (EV) and hybrid EV (HEV) applications implemented without the use of inductors. Currently available power inverter systems for HEVs use a dc–dc boost converter to boost the battery voltage for a traditional three-phase inverter. A cascaded H-bridge multilevel boost inverter design for EV and HEV applications implemented without the use of inductors is proposed in this paper. The proposed design uses a standard three-leg inverter (one leg for each phase) and an H-bridge in series with each inverter leg which uses a capacitor as the dc power source. Experiments show that the proposed dc–ac cascaded H-bridge multilevel boost inverter can output a boosted ac voltage without the use of inductors.

Keywords: Cascaded H-bridge multilevel boost inverter, Three-leg inverter, electric vehicle (EV)/hybrid electric vehicle (HEV).

1 Introduction

Now a day's increasing of oil prices and environmental issues, hybrid electric vehicles (HEVs) and electric vehicles (EVs) are gaining increased attention due to their higher efficiencies and lower emissions associated with the development of improved power electronics and motor Technologies. HEV typically combines a smaller internal combustion engine of a conventional vehicle with a battery pack and an electric motor to drive the vehicle. This combination having advantage of lower emissions. And convenient fueling of conventional (gasoline and diesel) vehicles. An EV typically uses rechargeable batteries and an electric motor. Here batteries want to be charged regularly. Both HEVs and EVs need a traction motor and a power inverter to drive the traction motor. The requirements for the power inverter include high peak power and low continuous Power rating. Currently available power inverter systems for HEVs

use a dc–dc boost converter to boost the battery voltage for a traditional three-phase inverter. If the motor is running at low to medium power, in this case dc–dc boost converter is not necessary; here battery voltage will be directly applied to the inverter to drive the traction motor. If the motor is running in a high power mode, the dc–dc boost converter will boost the battery voltage to a higher voltage, so that the inverter can provide higher power to the motor. Present HEV traction drive inverters have low power density, are expensive, and have low efficiency because they need bulky inductors for the dc–dc boost converters. To achieve a boosted output ac voltage from the traditional inverters for HEV and EV applications, the Z-source inverter is proposed, which also requires an inductor. A cascaded H-bridge multilevel boost inverter shown in Fig. 1 for EV and HEV applications is described in this paper. Traditionally, each H-bridge of a cascaded multilevel inverter needs a dc power supply. The proposed cascaded H-bridge multilevel boost inverter uses a standard three-leg inverter (one leg for each phase) and an H-bridge in series with each inverter leg which uses a capacitor as the dc power source. In this topology, the need for large inductors is eliminated. A fundamental switching scheme is used to do modulation control and to output five-level phase voltages. Experiments show that the proposed dc–ac cascaded H-bridge multilevel boost inverter without inductors can output a boosted ac voltage.

2 Working Principle of Cascaded H-Bridge Multilevel Boost Inverter without Inductors

The topology of the proposed dc–ac cascaded H-bridge multilevel boost inverter is shown in Fig. 1. The inverter uses a standard three-leg inverter and an H-bridge with a capacitor as its dc source in series with each phase leg.

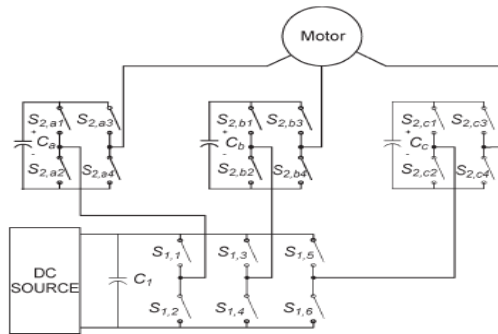


Fig. 1. Topology of the proposed dc-ac cascaded H-bridge multilevel boost inverter

2.1 Switching Technique of Cascaded H-Bridge Multilevel Boost Inverter without Inductors

There are several kinds of modulation control methods such as traditional sinusoidal pulse width modulation (SPWM), space vector PWM, harmonic optimization or selective harmonic elimination, and active harmonic elimination, and they all can be

used for inverter modulation control. For the proposed dc–ac boost inverter control, a practical modulation control method is the fundamental frequency switching control for high output voltage and SPWM control for low output voltage, which only uses the bottom inverter. In this paper, the fundamental frequency switching control is used.

$$V(\omega t) = \frac{4V_{dc}}{\pi} \times \sum_{n=1,3,5,\dots}^{\infty} \frac{1}{n} (\cos(n\theta_1) + \cos(n\theta_2)) \sin(n\omega t). \tag{1}$$

The key issue of fundamental frequency modulation control is choice of the two switching angles θ_1 and θ_2 . In this paper, the goal is to output the desired fundamental frequency voltage and to eliminate the fifth harmonic. Mathematically, this can be formulated as the solution to the following:

$$\begin{aligned} \cos(\theta_1) + \cos(\theta_2) &= ma \\ \cos(5\theta_1) + \cos(5\theta_2) &= 0. \end{aligned} \tag{2}$$

This is a system of two transcendental equations with two unknowns θ_1 and θ_2 , and ma is the output voltage index. Traditionally, the modulation index is defined as

$$m = \frac{V_1}{V_{dc}/2}. \tag{3}$$

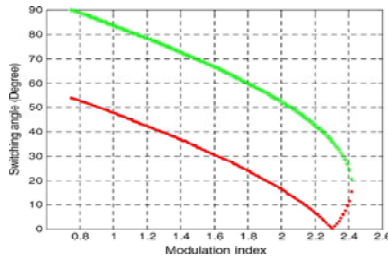


Fig. 2. Switching angle solutions for proposed dc–ac cascaded H-bridge multilevel boost inverter control

Therefore, the relationship between the modulation index m and the output voltage index ma is

$$m = \frac{4}{\pi} m_a. \tag{4}$$

There are many ways one can solve (2) for the angles. Here, the resultant method is used to find the switching angles. A practical solution set is shown in Fig. 2, which is continuous from modulation index 0.75 to 2.42. Although it can be seen from Fig. 2 that the modulation index range for the five-level fundamental frequency switching control method can reach 2.42, which is double that of the traditional power inverter, it requires the capacitors’ voltage to be kept constant at $V_{dc}/2$. Traditionally, the maximum modulation index for the linear operation of a traditional full-bridge bi-level inverter using SPWM control method is 1 (without third harmonic compensation) and

1.15 (with third harmonic compensation, and the inverter output voltage waveform is an SPWM waveform, not a square waveform). With the cascaded H-bridge multilevel inverter, the maximum modulation index for linear operation can be as high as 2.42; however, the maximum modulation index depends on the displacement power factor, as will be shown in the next section.

3 Output Voltage Boost

As previously mentioned, the cascaded H-bridge multilevel inverter can output a boosted ac voltage to increase the output power, and the output ac voltage depends on the displacement power factor of the load. Here, the relationship of the boosted ac voltage and the displacement power factor is discussed. It is assumed that the load current displacement angle is ϕ . To balance the capacitor voltage, the net capacitor charging amount needs to be greater than the pure discharging amount. The inverter can regulate the capacitor's voltage with a displacement power factor of one if the modulation index is below 1.27; if the modulation index is above 1.27, the displacement power factor must be less than a specified amount. For practical applications, the highest output voltage is determined when the load is determined. As mentioned previously, there are many methods to do modulation control for the proposed dc-ac cascaded H-bridge multilevel boost inverter without inductors. The fundamental frequency method with regulated $V_{dc}/2$ capacitor voltage is only one of the possible methods to output continuous power. The traditional SPWM method can also be applied to this inverter to boost the output voltage with a lower maximum continuous output power and high switching loss but better THD for a lower output frequency range. It is also possible to use SPWM for low output frequency low output voltage conditions and staircase waveform for high output frequency high output voltage range to achieve optimal performances with maximum continuous output power, lower switching loss, and lower THD. It can also be seen that accurate load inductance is not required for controller design, and the controller is robust independent of the leakage inductance of stator windings. For HEV and EV applications, sometimes, only short period peak Power is required. The modulation control can store energy to the capacitors by boosting the capacitor voltage to a higher voltage, which could be higher than V_{dc} when the vehicle is working in a low power mode. When the vehicle is working in high power modes, the capacitors will deliver much higher power than the continuous power to the motor load combined with the battery, fuel cell, or generator. This feature will greatly improve the vehicle's dynamic (acceleration) performance.

4 Experimental Implementation and Validation

A real-time variable output-voltage variable-frequency three-phase motor drive controller based on an Altera FLEX 10 K field programmable gate array (FPGA) is used to implement the control algorithm. For convenience of operation, the FPGA controller is designed as a card to be plugged into a personal computer, which uses a peripheral component interconnect bus to communicate with the microcomputer. To maintain the capacitors' voltage balance, a voltage sensor is used to detect the capacitors voltage

and feed the voltage signal into the FPGA controller. The FPGA controller will output the corresponding switching signals according to the capacitor's voltage. To further test the cascaded multilevel boost inverter, experiments with load current versus modulation indexes with different fundamental frequencies were performed to achieve the highest output voltages. These were implemented by using an $R-L$ load bank and compared to a traditional inverter. For these experiments, the $R-L$ load was fixed, the modulation index was changed with different fundamental frequencies, and the load currents were recorded. In this experiment, to achieve the highest output voltages for the cascaded multilevel boost inverter without inductors and the traditional inverter, two steps were involved. First, the load was connected to the bottom traditional inverter to output its highest voltage; second, the load was connected to the cascaded H-bridge multilevel inverter with the same dc power supply voltage. The output voltages for the two cases are shown in Table 1. Table 1 shows that the highest output voltage of the cascaded H-bridge multilevel inverter is much higher than that of the Traditional inverter. The voltage boost ratio is higher than 1.4 for the whole testing frequency range.

Table 1. Highest Output Voltage For Traditional Inverter And Cascaded H-Bridge Multilevel Inverter (Dc Bus Is 40 V)

Test frequency (Hz)	Traditional inverter output voltage (V)	Cascaded H-bridge multilevel inverter output voltage (V)	Boost ratio
200	23.1	42.8	1.85
150	23.1	42.2	1.82
100	23.1	41.2	1.78
60	23.1	37.7	1.63
40	23.1	33.1	1.43

Table 1 also shows that the highest output voltage of the inverter is decreasing when the frequency is decreasing; this is because the impedance of the inductor is decreasing. Another issue is that the boost voltage ratio is decreasing when the frequency is decreasing; this is because the power factor is increasing for the fixed $R-L$ load.

5 Conclusion

The proposed cascaded H-bridge multilevel boost inverter without inductors uses a standard three-leg inverter and an H-bridge in series with each inverter leg. The application of this dc-ac boost inverter on HEV and EV can result in the elimination of the bulky inductor of present dc-dc boost converters, thereby increasing the power density.

References

1. Du, Z., Tolbert, L.M., Chiasson, J.N.: DC-AC Cascaded H-Bridge Multilevel Boost Inverter With No Inductors for Electric/Hybrid Electric Vehicle Applications. IEEE Transactions On Industry Applications 45(3) (May/June 2009)

2. Rahman, K.M., Patel, N.R., Ward, T.G., Nagashima, J.M., Caricchi, F., Crescimbeni, F.: Application of direct-drive wheel motor for fuel cell electric and hybrid electric vehicle propulsion system. *IEEE Trans. Ind. Appl.* 42(5), 1185–1192 (2006)
3. Hinkkanen, M., Luomi, J.: Braking scheme for vector-controlled induction motor drives equipped with diode rectifier without braking resistor. *IEEE Trans. Ind. Appl.* 42(5), 1257–1263 (2006)
4. Rivetta, C.H., Emadi, A., Williamson, G.A., Jayabalan, R., Fahimi, B.: Analysis and control of a buck dc–dc converter operating with constant power load in sea and undersea vehicles. *IEEE Trans. Ind. Appl.* 42(2), 559–572 (2006)
5. Lai, J.S., Peng, F.Z.: Multilevel converters—A new breed of power converters. *IEEE Trans. Ind. Appl.* 32(3), 36–44 (1996)
6. Lai, J.S., Rodriguez, J., Lai, J., Peng, F.: Multilevel inverters: A survey of topologies, controls and applications. *IEEE Trans. Ind. Appl.* 49(4), 724–738 (2002)
7. Tolbert, L.M., Peng, F.Z., Habetler, T.G.: Multilevel converters for large electric drives. *IEEE Trans. Ind. Appl.* 35(1), 36–44 (1999)
8. Jacobina, C.B., dos Santos, E.C., de Rossiter Correa, M.B., da Silva, E.R.C.: AC motor drives with a reduced number of switches and boost inductors. *IEEE Trans. Ind. Appl.* 43(1), 30–39 (2007)
9. Ben-Brahim, L., Tadakuma, S.: A novel multilevel carrier-based PWM-control method for GTO inverter in low index modulation region. *IEEE Trans. Ind. Appl.* 42(1), 121–127 (2006)
10. Schuch, L., Rech, C., Hey, H.L., Grundling, H.A., Pinheiro, H., Pinheiro, J.R.: Analysis and design of a new high-efficiency bidirectional integrated ZVT PWM converter for dc-bus and battery-bank interface. *IEEE Trans. Ind. Appl.* 42(5), 1321–1332 (2006)
11. Shen, M., Wang, J., Joseph, A., Peng, F.Z., Tolbert, L.M., Adams, D.J.: Constant boost control of the Z-source inverter to minimize current ripple and voltage stress. *IEEE Trans. Ind. Appl.* 42(3), 770–778 (2006)

Design of Microstrip Meandered Patch Antenna for Mobile Communication

Shobhit Patel^{1,*}, Jaymin Bhalani², Yogesh Kosta³, and Sanket Patel⁴

^{1,4} Assistant Professor, ² Associate Professor, ³ Professor
Electronics & Communication Department, Charotar University of Science and Technology,
Changa-388 221, Gujarat, India
shobhit_65@yahoo.com

Abstract. The enhancing bandwidth and size reduction mechanism that improves the performance of a conventional micro strip patch antenna on a relatively thin substrate (about 0.006λ), is presented in this research. The design adopts meandered patch structure. Introducing the novel meandered square patch, offer a low profile, broadband, high gain, and compact antenna element. The proposed patch has a compact dimension of $0.384 \lambda \times 0.384 \lambda$ (where λ is the guided wavelength of the centre operating frequency). The design is suitable for applications with respect to a given frequency of 750-1100 MHz. The simulated bandwidth of the proposed antenna is about 39%.

Keywords: Microstrip Patch Antenna, Meandered Patch.

1 Introduction

With the ever-increasing need for mobile communication and the emergence of many systems, it is important to design broadband antennas to cover a wide frequency range. The design of an efficient wide band small size antenna, for recent wireless applications, is a major challenge. Microstrip patch antennas have found extensive application in wireless communication system owing to their advantages such as low-profile, conformability, low-cost fabrication and ease of integration with feednetworks [1]. However, conventional microstrip patch antenna suffers from very narrow bandwidth, typically about 5% bandwidth with respect to the center frequency. This poses a design challenge for the microstrip antenna designer to meet the broadband techniques [2][3].

There are numerous and well-known methods to increase the bandwidth of antennas, including increase of the substrate thickness, the use of a low dielectric substrate, the use of various impedance matching and feeding techniques, the use of multiple resonators, and the use of slot antenna geometry [4][5][6][7]. However, the bandwidth and the size of an antenna are generally mutually conflicting properties, that is, improvement of one of the characteristics normally results in degradation of the other.

Recently, several techniques have been proposed to enhance the bandwidth. A novel meandered patch antenna with achievable impedance bandwidth of greater than 39% has been demonstrated [9].

* Corresponding author.

In this paper, a novel meandered shape patch is investigated for enhancing the impedance bandwidth on a thin substrate (about 0.006λ). In this paper, the design and simulations results of the novel wideband micro strip patch antenna, is described.

Many techniques have been reported to reduce the size of microstrip antennas at a fixed operating frequency. In general, microstrip antennas are half-wavelength structures and are operated at the fundamental resonant mode TM₀₁ or TM₁₀, with a resonant frequency given by (valid for a rectangular microstrip antenna with a thin microwave substrate)

$$f \cong \frac{c}{2L\sqrt{\epsilon_r}} \quad (1)$$

where c is the speed of light, L is the patch length of the rectangular micro strip antenna, and ϵ_r is the relative permittivity of the grounded microwave substrate.

Meandering the excited patch surface current paths in the antenna's radiating patch is also an effective method for achieving a lowered fundamental resonant frequency for the micro strip antenna [9][10][11]. For the case of a rectangular radiating patch, the meandering can be achieved by inserting several narrow slits at the patch's nonradiating edges. It can be seen in Fig.1 that the excited patch's surface currents are effectively meandered, leading to a greatly lengthened current path for a fixed patch linear dimension. This behavior results in a greatly lowered antenna fundamental resonant frequency, and thus a large antenna size reduction at a fixed operating frequency can be obtained.

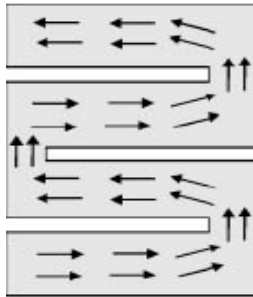


Fig. 1. Surface current distributions for meandered rectangular microstrip Patch with meandering slits

2 Designing and Modeling

It is known that increasing the thickness of the patch antenna will increase the impedance bandwidth. However, the thicker the substrate of the antenna, the longer the coaxial probe will be used and, thus, more probe inductance will be introduced [8], which limits the impedance bandwidth. Consequently, a patch antenna design that can counteract or reduce the probe inductance will enlarge the impedance bandwidth.

Fig.2 depicts the geometry of the proposed patch antenna. The square patch, with dimensions as given in Table 1. is supported by a low dielectric substrate with

dielectric permittivity ϵ (2.2). it is in between patch and ground plane. Here the patch contains two slits on both the sides of centerline of the x-axis as shown in Fig.2 making total four slits. The dimensions of slits are given in Table 1.

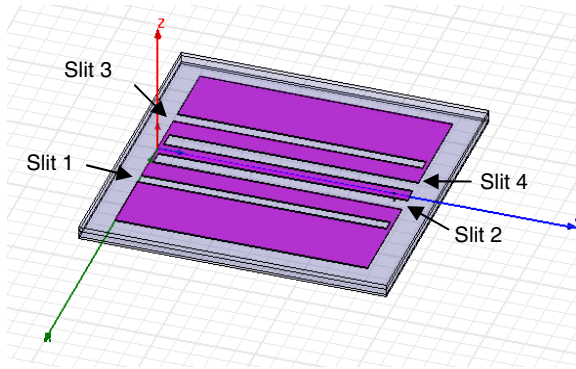


Fig. 2(a). Actual HFSS Model (top view) of meandered microstrip patch

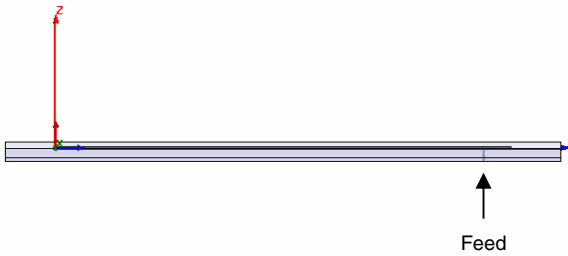


Fig. 2(b). Actual HFSS Model (side view) of meandered microstrip patch

Table 1. Dimentions of Patch Antenna

	Size in m.m
Patch	128×128×0.5
Substrate	135×135×2
Slit1	6×126×0.5
Slit2	4×126×0.5
Slit3	6×126×0.5
Slit4	4×126×0.5

3 Simulation Results and Discussions

For simulation we used High Frequency Structure Simulator (HFSS) version 11 of ansoft, which is very good simulator for RF antennas. After simulating the design the results we got is as follows. Fig. 3 shows the Return Loss(S_{11}) plot of the design.

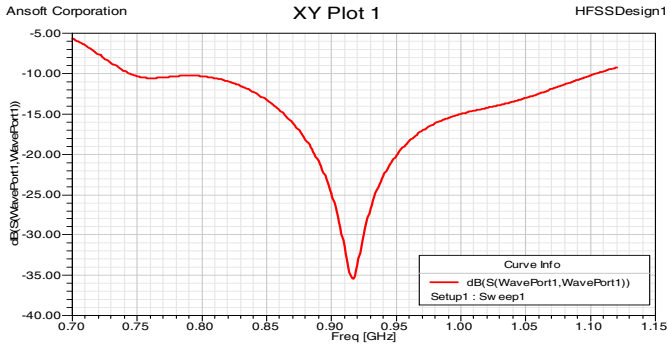


Fig. 3. S_{11} paramater of the antenna

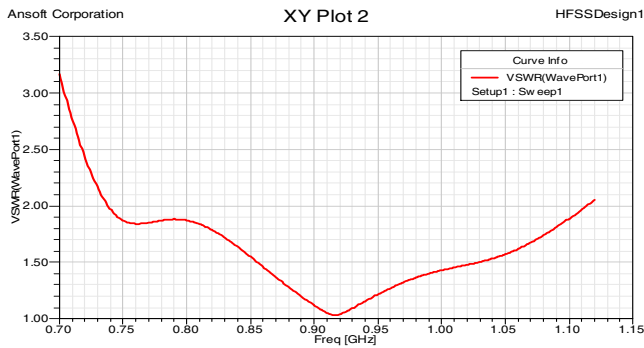


Fig. 4. VSWR of the antenna

Table 2. S_{11} & VSWR values

Frequency MHz	VSWR	Return loss(S_{11}) dB	Frequency MHz	VSWR	Return loss(S_{11}) dB
750	1.86	-10	930	1.09	-27
760	1.84	-11	940	1.15	-23
770	1.85	-11	950	1.23	-20
780	1.87	-10	960	1.26	-18
790	1.88	-10	970	1.32	-17
800	1.87	-10	980	1.36	-16
810	1.83	-11	990	1.40	-16
820	1.78	-11	1000	1.43	-15
830	1.71	-12	1010	1.46	-15
840	1.63	-12	1020	1.48	-14
850	1.55	-13	1030	1.49	-14
860	1.46	-15	1040	1.53	-13
870	1.37	-16	1050	1.56	-13
880	1.28	-18	1060	1.62	-12
890	1.20	-21	1070	1.68	-12
900	1.10	-25	1080	1.72	-11
910	1.05	-31	1090	1.81	-11
916	1.03	-36	1100	1.87	-10

Fig. 4 shows the VSWR plot of the design and Table 2 shows values of S_{11} & VSWR for different frequencies. For the whole range VSWR less than 2 and S_{11} is less than -10 dB and at frequency 916 MHz both are minimum

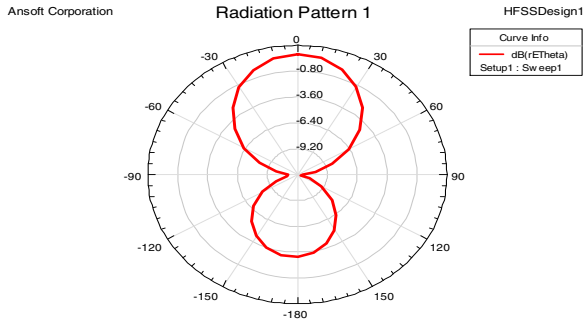


Fig. 5. Radiation Pattern of the antenna

Fig.5 shows the radiation pattern for frequency 916 MHz. Same way patterns for other frequencies can be generated.

The driven model simulations are carried out for frequency sweep of 700 MHz to 1150 MHz as shown in Fig. 3-5 under first order fast mode time domain analysis-set-up. The solution frequency is 900 MHz with linear step size of 0.0004 GHz, Maximum number of passes taken was 6, maximum delta S to be 0.02 and error tolerance was assumed to be 0.5%.

4 Conclusion and Future Scopes/Perspectives

The theoretical-geometric investigations and analysis, design, modeling and iterative simulations are carried out for center frequency of 900 MHz. The result indicates the total band of 750 MHz to 1100 MHz. So the antenna can be used for GSM 900, CDMA 800 and CDMA 850 bands. Further the band can be changed by modifying the design and further it can be designed for GSM 1800. The results are in very good agreement with the industry and standard published antenna-requirements with respect to ease of fabrication, compactness and volume miniaturization compared to other antennas so far designed for similar applications.

References

1. He, W., Jin, R., Geng, J.: E-Shape patch with wideband and circular polarization for millimeter-wave communication. *IEEE Transactions on Antennas and Propagation* 56(3), 893–895 (2008)
2. Lau, K.L., Luk, K.M., Lee, K.L.: Design of a circularly-polarized vertical patch antenna. *IEEE Transactions on Antennas and Propagation* 54(4), 1332–1335 (2006)
3. Zhang, Y.P., Wang, J.J.: Theory and analysis of differentially-driven microstrip antennas. *IEEE Transactions on Antennas and Propagation* 54(4), 1092–1099 (2006)

4. Matin, M.M., Sharif, B.S., Tsimenidis, C.C.: Probe fed stacked patch antenna for wide-band applications. *IEEE Transactions on Antennas and Propagation* 55(8), 2385–2388 (2007)
5. Pozar, D.M., Schaubert, D.H.: *Microstrip Antennas*. IEEE press, New York (1995)
6. Wi, S.H., Sun, Y.B., Song, I.S., Choa, S.H., Koh, I.S., Lee, Y.S., Yook, J.G.: Package-Level integrated antennas based on LTCC technology. *IEEE Transactions on Antennas and Propagation* 54(8), 2190–2197 (2006)
7. Wi, S.H., Kim, J.M., Yoo, T.H., Lee, H.J., Park, J.Y., Yook, J.G., Park, H.K.: Bow-tieshaped meander slot antenna for 5 GHz application. In: *Proc. IEEE Int. Symp. Antenna and Propagation*, vol. 2, pp. 456–459 (2002)
8. Yang, F., Zhang, X., Rahmat-Samii, Y.: Wide-band E-shaped patch antennas for wireless communications. *IEEE Transactions on Antennas and Propagation* 49(7), 1094–1100 (2001)
9. Dey, S., Mitra, R.: Compact microstrip patch antenna. *Microwave Opt. Technol. Lett.* 13, 12–14 (1996)
10. Wong, K.L., Tang, C.L., Chen, H.T.: Acompact meandered circular microstrip antenna with a shorting pin. *Microwave Opt. Technol. Lett.* 15, 147–149 (1997)
11. George, J., Deepukumar, M., Aanandan, C.K., Mohanan, P., Nair, K.G.: New compact microstrip antenna. *Electron. Lett.* 32, 508–509 (1996)

Building Gaussian Mixture Shadow Model for Removing Shadows in Surveillance Videos

Archana Chougule and Pratap Halkarnikar

Shivaji University, Kolhapur
Maharashtra, India

chouguleab@gmail.com, pp_halkarnikar@rediffmail.com

Abstract. Moving cast shadows are a major concern for foreground detection algorithms. Processing of foreground images in surveillance applications typically requires that such shadows have been identified and removed from the detected foreground. It is critical for accurate object detection in video streams since shadow points are often misclassified as object points, causing errors in segmentation and tracking. Processing of foreground images in surveillance applications typically requires that such shadows have been identified and removed from the detected foreground. We have implemented a novel pixel based statistical approach to model moving cast shadows of non-uniform and varying intensity by modelling each pixel as a mixture of Gaussians and using an on-line approximation to update the model.

Keywords: Gaussian mixture model, cast shadow, chromaticity.

1 Introduction

Shadows result from the obstruction of light from the light source. Shadows cause serious problems while segmenting and extracting moving objects due to the misclassification of shadow points as foreground. To obtain a better segmentation quality, detection algorithms must correctly separate foreground objects from the shadows they cast. Model-based approaches have shown less robustness than property-based algorithms [1]. A review of shadow detection techniques can be found in [2]. In [3], Salvador et al. use the fact that a shadow darkens the surfaces on which it is cast to identify an initial set of shadowed pixels that is then pruned by using color invariance and geometric properties of shadows. In [4], Cucchiara et al. use the hypothesis that shadows reduce surface brightness and saturation while maintaining hue properties in the HSV color space. Schreer et al. [5] adopt the YUV color space to avoid using the time consuming HSV color transformation and segment shadows from foreground objects based on the observation that shadows reduce the YUV pixel value linearly. In [6], Horprasert et al. build a model in the RGB color space to express normalized luminance variation and chromaticity distortions. [7] uses edge width information to differentiate penumbra regions from the background. The algorithm in [8] combines luminance, chrominance, and gradient density scores in a shadow confidence score function for segmentation.

In this paper, we present an approach for detecting and modelling moving cast shadows using Gaussian mixture model (GMM) in surveillance video.

2 Gaussian Mixture Model

The foreground object is detected as described in [9] and modified for online implementation in [10]. For each pixel, a fixed number of states K , typically between 3 and 5, is defined. Some of these states will model the YUV values of background surfaces, and the others, foreground surfaces. Each pixel value X_t is a sample in a YUV color space of a random variable \mathbf{X} . The K states are ordered decreasingly by their $\omega_k/\|\sigma_k\|$ ratio and the first B states whose combined probability of appearing is greater than the threshold T , i.e.,

$$B = \arg \min \left(\sum_{k=1}^b W_k > T \right) \quad (1)$$

are labelled as background states, and the other states are labelled as foreground states.

2.1 Shadow Modelling

Our approach is based on the hypothesis that, for a given pixel, the shadow cast by different moving foreground objects is relatively similar. The states that capture cast shadows are identified, and their pdfs are used to build a second GMM, which we call the Gaussian mixture shadow model (GMSM).

When using a GMM to build a background model for foreground detection, it is possible to observe that some of the states model the shadows cast by persons moving across the scene. A first person crosses the scene, and on pixels where its shadow is cast, a new state representing the value of this shadow will be created with a small *a priori* probability ω_{init} . When the next person crosses the scene and casts a shadow on the same pixels, the algorithm will associate the pixel values to the same new state describing the shadow value. The *a priori* probability of this state will then increase as

$$\omega_{k,t+1} = \omega_{k,t} + \alpha_{s,t} M_{k,t} \quad (2)$$

where $\alpha_{s,t}$ is the learning parameter, $M_{k,t}$ is equal to 1 for the state that is associated to the pixel value, and zero for the other states. When the pixel value could describe a shadow over the surface background, we increase the learning rate of the pixel associated state parameter ω_k . This modification allows a state representing a shadow on a background surface to rapidly become a stable foreground state and this without preventing the creation of states describing the background. To do so, we define

$$\alpha_{s,t} = \alpha_t S \quad (3)$$

where the parameter $S > 1$ when X_t correspond to a shadowed background surface, and $S = 1$ in the other cases. When there are no persons or objects crossing the scene for a long time, the *a priori* probability w_k of the foreground states modelling the cast shadows will tend toward zero. Since w_k will become smaller than w_{init} , any future

detection of a foreground event will result in a new state and the destruction of foreground states capturing the values of shadow surfaces.

2.2 Properties of Shadowed Surfaces

As in [5], we adopted the YUV color space and our model is also based on the observation that a shadow cast on a surface will equally attenuate the value of its three components. We first estimate this attenuation ratio using the luminance component Y , and we then verify that both U and V components are also reduced by a similar ratio. More specifically, if color vector X represents the shadow cast on a surface whose average color value is μ with variance σ_μ , we have

$$\alpha_{min} < \alpha_Y < 1 \text{ with } \alpha_Y = X_Y/\mu Y \quad (4)$$

$$(1/\sigma_{\mu U})|X_U - \alpha_Y \mu U| < \Lambda_U \quad (5)$$

$$(1/\sigma_{\mu V})|X_V - \alpha_Y \mu V| < \Lambda_V \quad (6)$$

where α_{min} is a threshold on maximum luminance reduction. $\Lambda_{U,V}$ represent the tolerable chromaticity fluctuation around the surface value μ_U, V . If we match a pixel value X_t to a non-background state of the GMM, we then verify if that pixel value could be the shadowed surface of the most frequent background state $k = 1$:

$$\alpha_{min} < \alpha_Y < 1 \text{ with } \alpha_Y = X_{t,Y}/\mu_{1,t,Y} \quad (7)$$

$$(1/\sigma_{1,\mu,t,U})|X_{t,U} - \alpha_Y \mu_{1,t,U}| < \Lambda_U \quad (8)$$

$$(1/\sigma_{1,\mu,t,V})|X_{t,V} - \alpha_Y \mu_{1,t,V}| < \Lambda_V \quad (9)$$

If these three conditions are met, we increase the learning rate $\alpha_{s,t}$ of the *a priori* probability ω_k of that state.

2.3 Gaussian Mixture Shadow Models

Updating the Gaussian mixture shadow model is the next step. Unlike the GMM [9], the GSM filters the input values, and these input values are Gaussian probability density functions $f_{X|k}$ with parameters $\theta_k = \{\mu_k, \sigma_k\}$. These Gaussian pdf are those of the states which have been identified as describing cast shadows on background surfaces. We use the following conditions on $f_{X|B+1}$ to determine if this state describes a cast shadow on the background surface of state $k = 1$:

$$\alpha_{min} < \alpha_Y < 1 \text{ with } \alpha_Y = \mu_{B+1,t,Y}/\mu_{1,t,Y} \quad (10)$$

$$(1/\sigma_{1,\mu,t,U})|\mu_{B+1,t,U} - \alpha_Y \mu_{1,t,U}| < \Lambda_U \quad (11)$$

$$(1/\sigma_{1,\mu,t,V})|\mu_{B+1,t,V} - \alpha_Y \mu_{1,t,V}| < \Lambda_V \quad (12)$$

If these conditions are met, pdf $f_{X|B+1}$ is transferred to the GSM. It is then compared to the existing GSM pdfs:

$$d_{k,t}^T d_{k,t} < \lambda_{s,a}^2$$

where

$$d_{k,t} = (\sigma_{k,t}^s I) - 1(\mu_{B+1,t} - \mu_{k,t}^s) \quad (13)$$

and where we use the superscript s when referring to the Gaussian mixture shadow model. If there is a match, the parameters θ_{sk} are updated:

$$\mu_{k,t+1}^s = (1 - \alpha) \mu_{k,t}^s + \alpha \mu_{B+1,t} \quad (14)$$

$$\sigma_{k,t+1}^s = (1 - \alpha) \sigma_{k,t}^s + \alpha \sigma_{B+1,t} \quad (15)$$

where α is a constant. Also, we set $M_{sk,t} = 1$ in the following equation :

$$\omega_{k,t+1}^s = \omega_{k,t}^s + \alpha M_{sk,t} \quad (16)$$

If there is no match, a new state is added in the GSM, upto a maximum of K_s states. For this new state,

$$\omega_{k,t+1}^s = \omega_{init}^s \quad (17)$$

$$\mu_{k,t+1}^s = \mu_{B+1,t} \quad (18)$$

$$\sigma_{k,t+1}^s = \sigma_{B+1,t} \quad (19)$$

The *a priori* probabilities $\omega_{k,t}^s$ are then normalized, and the states sorted in decreasing order of $\omega_{sk,t}$. We do not use the ratio $\omega_{k,t}^s / \sigma_{k,t}^s$ since the variance is relatively constant for all states in the GSM. Within the GSM, the first B_s states, where

$$B_s = \arg \min_b \left(\sum_{k=1}^b \omega_{k,t}^s > T_s \right) \quad (20)$$

are used to model moving cast shadows on a background surface.

2.4 Removing Shadows

In a GMM, a pixel is labelled as foreground if its value X_t is matched to a foreground state. This detected pixel is then compared to the shadow states of the GSM

$$d_{k,t}^T d_{k,t} < \lambda_{s,b}^2$$

where

$$d_{k,t} = (\sigma_{k,t}^s I) - 1(X_t - \mu_{k,t}^s) \quad (21)$$

If this condition is met for a state ω_k^s with $k \leq B_s$, the pixel is labeled as representing a moving cast shadow. It is then simply removed from the detected pixels of the

GMM. We are then left with detected pixels representing foreground objects without their cast shadows.

3 Experimental Results

The method explained above has been tested on a number of image sequences. The experiments were performed on Intel Core2Duo CPU and the algorithm was implemented using Matlab 7.6. Results shown here are raw results, without any post treatment. For each environment, parameters were set once. Fig. 1 shows results for the human object scene when GMM is applied. Here moving object along with the shadow is extracted. Fig. 2 shows the result after applying GSM.

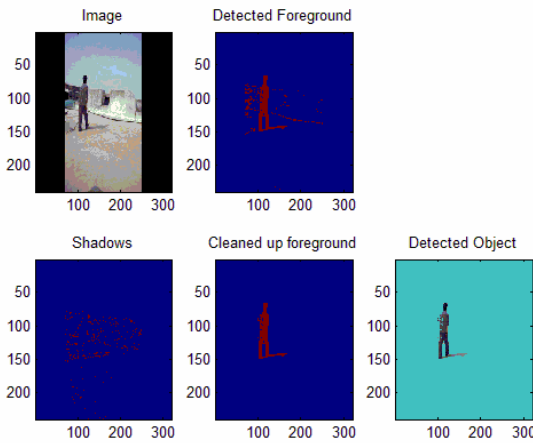


Fig. 1. Moving object detection

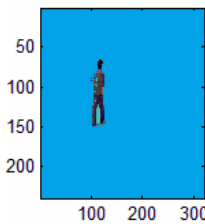


Fig. 2. Result of GSM

4 Conclusion

In this paper, we have described a system for removing shadows from a video, which is pixel-based statistical approach to model and detect moving cast shadows. The proposed approach uses a GMM to learn from repetition the properties of the

shadowed background surfaces. The algorithm identifies distributions that could represent shadowed surfaces, modify their learning rates to allow them to converge within the GMM, and then uses them to build a GMM for moving shadows on background surfaces. This approach can be used with various shadow models in different color spaces.

References

1. Prati, A., Mikic, I., Trivedi, M.M., Cucchiara, R.: Detecting Moving Shadows: Algorithms and Evaluation. *IEEE Trans. Pattern Analysis and Machine Intelligence* 25(7), 918–923 (2003)
2. Nadimi, S., Bhanu, B.: Physical Models for Moving Shadow and Object Detection in Video. *IEEE Trans. Pattern Analysis and Machine Intelligence* 26(8) (August 2004)
3. Salvador, E., Cavallaro, A., Ebrahimi, T.: Cast Shadow Segmentation Using Invariant Color Features. *Computer Vision and Image Understanding*, 238–259 (2004)
4. Cucchiara, R., Grana, C., Piccardi, M., Prati, A., Sirotti, S.: Improving Shadow Suppression in Moving Object Detection with HSV Color Information. In: *Proc. Intelligent Transportation Systems Conf.*, pp. 334–339 (2001)
5. Schreer, O., Feldmann, I., Goelz, U., Kauff, P.: Fast and Robust Shadow Detection in Videoconference Applications. In: *Proc. Fourth IEEE Int'l Symp. Video Proces. and Multimedia Comm.*, pp. 371–375 (2002)
6. Horprasert, T., Hardwood, D., Davis, L.S.: A Statistical Approach for Real-Time Robust Background Subtraction and Shadow Detection. In: *Proc. Int'l Conf. Computer Vision FRAMERATE Workshop* (1999)
7. Stauder, J., Mech, R., Ostermann, J.: Detection of Moving Cast Shadows for Object Segmentation. *IEEE Trans. Multimedia* 1(1), 65–76 (1999)
8. Fung, G.S.K., Yung, N.H.C., Pang, G.K.H., Lai, A.H.S.: Effective Moving Cast Shadows Detection for Monocular Image Sequences. In: *Proc. 11th Int'l Conf. Image Analysis and Processing (ICIAP)*, pp. 404–409 (2001)
9. Stauffer, C., Grimson, W.E.L.: Learning Patterns of Activity Using Real-Time Tracking. *IEEE Trans. Pattern Analysis and Machine Intelligence* 22(8), 747–757 (2000)
10. Power, P.W., Schoonees, J.A.: Understanding Background Mixture Models for Foreground Segmentation. *Proc. Image and Vision Computing*, 267–271 (2002)

FAutoREDWithRED: To Increase the Overall Performance of Internet Routers

K. Chitra¹ and G. Padmavathi²

¹ Associate Professor, Dept. of Computer Science,
D.J. Academy for Managerial Excellence, Coimbatore,
Tamil Nadu, India – 641032

chitrakandaswamy@yahoo.com

² Department of Computer Science,
Avinashilingam university for Women, Tamil Nadu, India

Abstract. Active Queue Management is a solution to the problem of congestion control in Internet routers. The problem of congestion degrades the overall performance of Internet routers. As data traffic is bursty in routers with varying flows, burstiness must be handled intelligently without comprising the overall performance. A congested link leads to many problems such as large delay, unfairness among flows, underutilization of the link and packet drops in burst. In this paper, we propose an AQM scheme that considers the characteristics of queue length based AQMs and uses the flow information to increase the overall performance and to satisfy the QOS requirements of the network.

Keywords: Congestion, Drop Probability, Fairness, Queue length, Misbehaving flows.

1 Introduction

A router in the Internet may receive heavy traffic at any given time due to high use of real-time applications. A Internet receives varying flows that should receive its fair share while sharing queue in the router. Another possibility is that the heavy load tends to vary in an Internet router resulting in a queue oscillation. Router must take care of the above problem. The robustness of today's Internet depends heavily on the congestion control mechanism. So the buffer in the routers is to be used effectively by using an efficient Active Queue Management Mechanisms. Active Queue Management prevents congestion and provides quality of service to all users. A router implementing RED AQM [1] maintains a single queue that drops an arriving packet at random during periods of congestion. RED suffers from lockout and global synchronization problems when parameters are not tuned properly. RED allows unfair bandwidth allocation when a mixture of traffic types traverses a link. Adaptive RED [2], AutoRED with RED [3] tries to improve the parameter tuning problem in RED. While AQMs like PD-RED [4], MRED [5], DS-RED [6] tries to improve the performance compared to RED.

Some AQMs [7] [8] techniques were introduced that used input rate as congestion indicators besides using average queue size. AQM [9] was proposed that used both congestion indicators queue size and input rate to detect congestion where as BLUE [10] uses link history and packet loss as congestion indicator to compute the packet drop probability.

The objective of this paper is to propose an algorithm that improves the overall performance of Internet routers. Because of the simplicity and easy implementation that exists in AutoREDwithRED, this paper takes this Queue based AQM as the base algorithm with certain modifications. The proposed AQM is implemented in this to bring in the advantages of Queue based algorithm and to remove the problems like parameter tuning problem. This algorithm is simple to implement, improves the overall performance of the routers. The rest of the paper is organized as follows: Section 2 explains the background study that includes the various AQM algorithms. In Section 3, the concepts regarding the proposed algorithm are discussed. In section 4, the simulation is carried out with discussion of results. Our conclusions are presented in section 5.

2 Background

The Internet routers face the problem of congestion from the birth of real-time applications in Internet. So research in this field of congestion has become a continuous process to bring out the best Active Queue Management algorithm to improve the performance of Internet even in case of heavy traffic. The various existing AQMs detect congestion based on different factors and calculate the packet dropping probability. Floyd et al proposed the first RED AQM in 1993 with the objective of preventing congestion with reduced packet loss. This AQM alleviates congestion by detecting incipient congestion early and delivering congestion notification to the source to reduce its transmission rates avoiding overflow from occurring. But the appropriate selection of the RED parameters defines the success of RED. In case of heavy traffic, RED AQM also leads to global synchronization, lock-out problem and unstable queue size if parameters not properly tuned. Further in improving RED and RED based AQMs, AutoRED technique was implemented in them that used the concept of dynamic w_q .

The YELLOW AQM proves that the packet drop probability just does not depend only on the queue length rather on input rate also that helps in identifying the real congestion in the queue. In case of the rate-based AQM AVQ, it maintains a virtual queue whose capacity is less than the actual capacity of the link. In improving the method for setting the value for γ , SAVQ [11] is proposed. SAVQ stabilizes the dynamics of queue maintaining high link utilization. In REM, both queue length and load is used as congestion indicators. The BLUE algorithm resolves some of the problems of RED by employing two factors: packet loss from queue congestion and link utilization. So BLUE performs queue management based on packet loss and link utilization. It maintains a single probability p_m to mark or drop packets. SRED in [12] pre-emptively discards packets with a load-dependent probability when a buffer in a router is congested. It stabilizes its buffer occupancy at a level independent of the number of the active connections. GREEN [13] algorithm uses flow parameters and the knowledge of TCP end-host behavior to intelligently mark packets to prevent queue build up, and prevent congestion from occurring.

3 Proposed Algorithm

The proposed algorithm is motivated by the need for a stable operating point for the queue size and fair bandwidth allocation irrespective of the dynamic traffic and congestion characteristics of the n flows. An unstable queue size due to the parameter tuning problem results in high queue oscillation in queue based AQMs. We are motivated to identify a scheme that penalizes the unresponsive flows with the stable queue size. The proposed algorithm - FAutoREDWithRED enforces the concept of queue-based and uses the flow information. We propose an algorithm that modifies the Queue based AutoREDwithRED algorithm to penalize the unresponsive flows. This algorithm calculates the average queue size of the buffer for every packet arrival. As in Fig. 1, average queue size is compared with the thresholds min_{th} , max_{th} , for every arriving packet. If average queue size is less than min_{th} , every arriving When the average queue size is greater than min_{th} , every arriving packet is compared with a randomly selected packet from the queue for their flow id. If they have the same flow id, both are dropped. In this proposed algorithm, the arriving packet is dropped with a probability depending on the average queue size. This algorithm uses AutoRED technique where w_q is dynamic in nature compared to a constant w_q in RED. This algorithm will work fine as the parameters are well tuned automatically and parameterized.

```

For every packet arrival {
  Calculate  $Q_{ave}$ 
  if ( $Q_{ave} < min_{th}$ )
    Forward the new packet
  Else {
    Select randomly a packet from the queue for their flow id
    Compare arriving packet with a randomly selected packet.
    If they have the same flow id
      Drop both the packets
    Else
      if ( $Q_{ave} \leq max_{th}$ )
        Calculate the dropping probability  $p_a$ 
        Drop the packet with probability  $p_a$ 
      Else
        Drop the new packet
  }
}
    
```

Fig. 1. Pseudocode of FAutoREDwithRED

The dynamic value of w_q adapts itself to the varying nature of the congestion and traffic. The w_q is redefined as in [4], and results in reduced instantaneous queue oscillation. The definition of weighting parameter w_q is written as a product of three characteristics as follows:

$$T_t = p_t (1 - p_t) \quad J_t = \frac{2(5.923 + (q_t - q_{avg,t-1}))}{\ln(5.923 + (q_t - q_{avg,t-1}))} \quad K_t = \frac{1}{bs}$$

$$w_{q,t} = T_t * J_t * K_t$$

The first characteristic T_t represents the dynamic status of the congestion in the network. It signifies the probability with which the system can lead to congestion with the information available at time t . The second characteristic J_t projects the current status of traffic in the network at time t . T_t and J_t is a time dependent function. The third characteristic K_t is time independent parameter and it allows normalization of instantaneous queue size changes with respect to the buffer size (bs). Therefore these three characteristics are used to incorporate the dynamic changes in the congestion and traffic in the calculation of average queue size.

Though the dynamic varying nature of w_q takes care of the network characteristics it keeps the average queue length high and in a unstable point in case of heavy traffic. This algorithm overcomes this problem with the help of the flow based information. So both w_q and the flow information take care of the unresponsive flows and brings in stable fair queuing. Packets of unresponsive flows are dropped more often than the adaptive flows and well behaved flows to bring in fairness.

4 Experimentation

In this section, we will use the packet-simulator ns-2 to simulate the FAutoRED-WithRED algorithm. In this simulation the network topology is as in Fig. 2. The congestion link is in between the two routers R1 and R2. The TCP flows are derived from FTP sessions which transmit large size files. The UDP hosts send packets at a constant bit rate of 2 Mbps. In the simulation setup we consider 32 TCP flows and 1 UDP flow in the network. The minimum threshold min_{th} in the FAutoREDWithRED scheme is set to 100 and the maximum threshold max_{th} to be twice the min_{th} and the physical queue size is fixed at 300 packets.

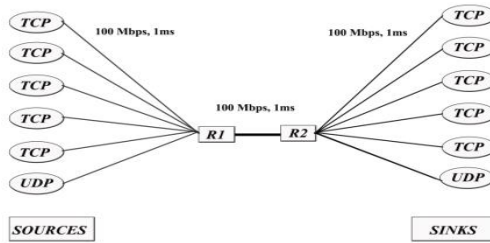


Fig. 2. Network Topology

In a dynamic varying mixture of traffic, the control parameter w_q alone does not help in achieving the stable operating point for the queue size. As shown in Fig. 3 dynamic varying parameter w_q maintains the average queue size and instability at a higher level in case of AutoREDwithRED for dynamic varying traffic. But this algorithm shows a stable and a low queue size compared to other AQMs as in Fig 3.

RED and other AQMs are unable to penalize unresponsive flows. The misbehaving traffic like UDP can take up a large % of the link bandwidth and starve out TCP

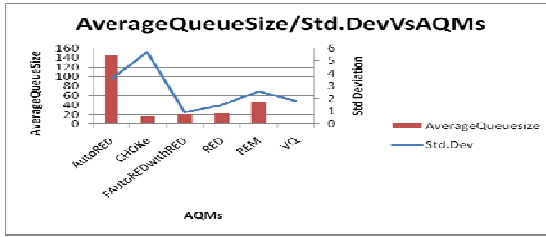


Fig. 3. Comparison of Average queue size/Queue Stability of AQMs

friendly flows in existing AQMs as in Fig 4. FAutoREDWithRED identifies and penalizes misbehaving flows effectively compared to the existing AQMs as in Fig 4. In existing RED based AQMs, the UDP flow uses 98% of the link capacity while TCP uses only 2% of the link capacity. This algorithm indicates the improvement of TCP throughput by limiting the UDP throughput to 212 Kbps which is around 21 % of the link capacity. The total TCP throughput is increased from 2 Kbps to 750 Kbps in FAutoREDwithRED.

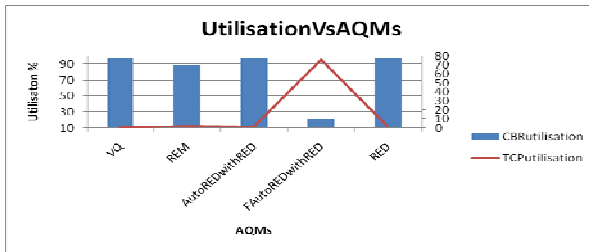


Fig. 4. CBR and TCP Utilisation of other AQMs with FAutoREDWithRED

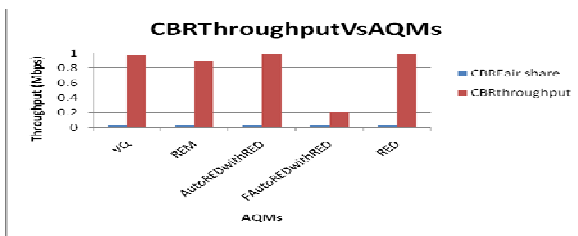


Fig. 5. Comparison of CBR Fair share and Throughput of other AQMs with FAutoRED-WithRED

To enforce the degree to which FAutoREDwithRED gains fair bandwidth allocation, CBR throughput of UDP connection along with their ideal fair share is shown in the Fig 5. A misbehaving flow with a high arrival rate and high buffer occupancy incurs high packet dropping mostly due to matches. The other AQMs almost take up the entire bandwidth for UDP flow showing unfairness though its actual CBR fair

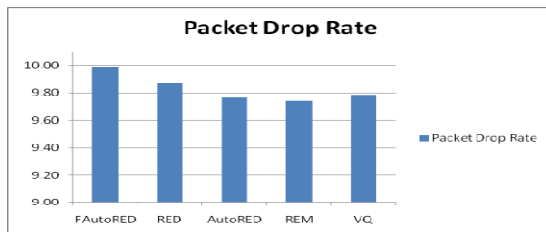


Fig. 6. Comparison of Packet Drop Rate of other AQMs with FAutoREDWithRED

share is very minimum. But in this algorithm, UDP is penalized and CBR throughput is only 21% of the bandwidth. In case of packet drop rate shown in Fig 6, the FAutoREDwithRED tries to give a fair share by dropping the UDP packets otherwise the UDP utilises the link to the maximum without allowing the TCP packets.

5 Conclusions

This paper proposes an AQM scheme called FAutoREDwithRED which aims to increase the overall performance of Internet routers. It reduces the queue oscillation in case of mixture of traffic at the link. It penalises the non-adaptive flows without comprising high utilisation, low queuing delay and controlled packet loss. It is achieved with well dynamically tuned parameters and flow information. This packet dropping scheme discriminates against the unresponsive flows resulting in fair congestion indication. This proposed AQM scheme inherits the advantages of these queue length based AQM and uses flow information to satisfy the QOS requirements of the network.

References

- [1] Floyd, S., Jacobson, V.: Random early detection gateways for congestion avoidance. *IEEE/ACM Trans. Networking* 1, 397–413 (1993)
- [2] Floyd, S., Gummadi, R., Shenkar, S., ICSI: Adaptive RED: An algorithm for Increasing the robustness of RED's active Queue Management, Berkely, CA <http://www.icir.org/floyd/red.html> (online)
- [3] Suthaharan, S.: Reduction of queue oscillation in the next generation Internet routers. Science Direct, *Computer Communication* (2007)
- [4] Sun, J., Ko., K.-T., Chen, G., Chan, S., sukerman, M.: PD –Red: To Improve Performance of Red. *IEEE Communications Letter* (August 2003)
- [5] Koo, J., Song, B., Chung, K., Lee, H., Kahng, H.: MRED: A New Approach To Random Early Detection. In: 15th International Conference on Information Networking (February 2001)
- [6] Zheng, B., Atiquzzaman, M.: DSRED: An Active Queue Management Scheme for Next Generation Networks. In: Proceedings of 25th IEEE conference on Local Computer Networks, LCN 2000 (November 2000)

- [7] Kunniyur, S., Srikant, R.: Analysis and design of an adaptive virtual queue (AVQ) algorithm for active queue management. In: Proceedings of ACM SIGCOMM, San Diego (2001)
- [8] Long, C., Zhao, B., Guan, X., Yang, J.: The Yellow active queue management algorithm. *Computer Networks* (November 2004)
- [9] Athuraliya, S., Li, V.H., Low, S.H., Yin, Q.: REM: Active queue management. *IEEE Network Mag* 15, 48–53 (2001)
- [10] Feng, W., Kandlur, D.D., Saha, D., Saha, D.: The Blue active queue management algorithms. *IEEE Transactions on Networking* (2002)
- [11] C.N., long, C.N., Zhao, B., Guan, X.-P.: SAVQ: Stabilized Adaptive Virtual Queue Management Algorithm. *IEEE Communications Letters* (January 2005)
- [12] Ott, T.J., Lakshman, T.V., Wong, L.: SRED: Stablised RED. *IEEE Infocomm* (March 1999)
- [13] Feng, W.-c., Kapadia, A., Thulasidasan, S.: GREEN: Proactive Queue Management over a Best-Effort Network. In: *IEEE GlobeCom*, Taipei, Taiwan (November 2002)

Scan Based Sequential Circuit Testing Using DFT Advisor

P. Reshma

M. Tech, VLSI design, Amrita Vishwa Vidyapeetham, Coimbatore
reshma_p2004@rediffmail.com

Abstract. This paper shows that not every scan cell contributes equally to the power consumption during scan based test. The transitions at some scan cells cause more toggles at the internal signal lines of a circuit, which have a larger impact on the power consumption during test application than the transitions at other scan cells. They are called power sensitive scan cells. A verilog based approach is proposed to identify a set of power sensitive scan cells. Additional hardware is added to freeze the outputs of power sensitive scan cells during scan shifting in order to reduce the shift power consumption.

Keywords: Scan, Scan DFlip-Flop, DFTAdvisor.

1 Introduction

Scan-based tests might cause excessive circuit switching activity compared to a circuit's normal operation. Higher switching activity causes higher peak supply currents and higher power dissipation. High power dissipation during test can cause many problems, which are generally addressed in terms of average power and peak power. Average power is the total distribution of power over a time period. Peak power is the highest power value at any given instant. When peak power is beyond the design limit, a chip cannot be guaranteed to function properly due to additional gate delays caused by the supply voltage drop. The power consumption within one clock cycle may not be large enough to elevate the temperature over the chip's thermal capacity limit. To damage the chip, high power consumption must last for an enough number of clock cycles. The test power consumed during scan shifting and capture cycles is referred to as shift power and capture power, respectively. Average power consumption is determined by the shift power. This paper, focus on reducing the average shift power in scan-based tests. An effective method to identify a set of power sensitive scan cells is used. By inserting additional hardware to freeze power sensitive scan cell outputs to pre-selected values, the average shift power consumption can be reduced dramatically.

2 The Proposed Method

A mentorgraphics tool called DFTAdvisor is used for Scan insertion. An S27 ISCAS benchmark circuit is taken and scan is inserted in that circuit using DFTAdvisor. In

that scan inserted circuit, number of toggling is calculated using verilog. The toggling between the first and second ScanD-FF is higher. By introducing logic gates at the output of first ScanD-FF we can block the complete toggling occurring at the combinational part, when only the Flip-Flops are selected. Implementation of a frozen scan cell is shown in Fig. 1. An additional AND gate with the second input inverted is inserted between the scan cell output and the function logics it drives. During scan shifting, the scan enable signal *se* is asserted to be 1. It makes additional gate output become constant. During normal operations, *se* is de-asserted to be 0. Vector re-ordering is also done, so as to reduce the toggling activity, thereby reducing the power dissipation to a larger extent.

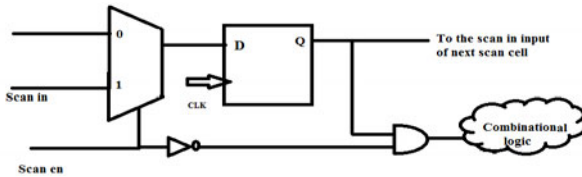


Fig. 1. Hardware implementation of frozen scan cell

3 Experimental Results

For testing a sequential circuit, both combinational cloud testing as well as memory element testing is needed. DFTAdvisor accepts gate level net list format and generates a new net list with scan cells inserted. The transitions at some scan cells cause more toggles at the internal signal lines of a circuit than the transitions at other scan cells. These scan cells are called power sensitive scan cells. Using verilog, power sensitive scan cells are identified. Additional hardware is added to freeze the outputs of power sensitive scan cells during scan shifting in order to reduce the shift power consumption. The number of togglings is higher in between first and second flipflop. By introducing the circuit in Fig.1b the combinational cloud in between *d_out_0* and *d_out_1* can be isolated,when scan enable becomes 1(only flip-flops are selected).Thereby power dissipation of the whole circuit can be reduced to a larger extent.

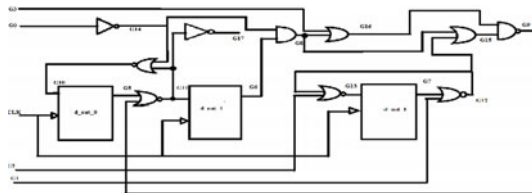


Fig. 2. ISCAS S27 Benchmark Circuit

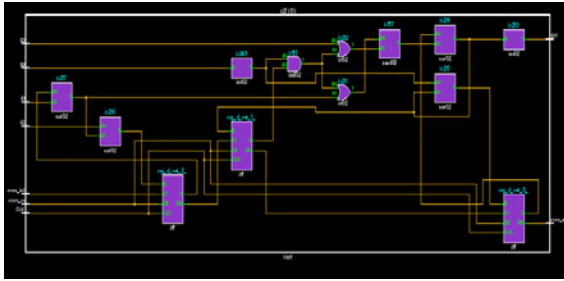


Fig. 3. Scan inserted circuit obtained from DFTAdvisor

```

/*
 * DESC: Generated by DFTAdvisor at Sat Dec 4 14:22:11 2010
 */
module s27 (CLK , G0 , G1 , G17 , G2 , G3 , scan_in1 , scan_out1 , scan_en);
input CLK , G0 , G1 , G2 , G3 , scan_in1 , scan_en;
output G17 , scan_out1;
wire G9 , G16 , G15 , G12 , G8 , G14 , G7 , G13 , G6 , G11 , G5 , G10;
wire [3:0] $dummy;
sff reg_d_out_0_ (D (G10) , .SI ($dummy [1]) ,
.SE (scan_en) , .CLK (CLK) , .Q (G5) , .QB ($dummy [0]));
sff reg_d_out_1_ (D (G11) , .SI ($dummy [2]) ,
.SE (scan_en) , .CLK (CLK) , .Q (G6) , .QB ($dummy [1]));
sff reg_d_out_2_ (D (G13) , .SI (scan_in1) ,
.SE (scan_en) , .CLK (CLK) , .Q (G7) , .QB ($dummy [2]));
inv02 ix249 (A (G0) , .Y (G14));
inv02 ix250 (A (G11) , .Y (G17));
and02 ix155 (A0 (G6) , .A1 (G14) , .Y (G8));
or02 ix211 (A0 (G8) , .A1 (G12) , .Y (G15));
or02 ix212 (A0 (G3) , .A1 (G8) , .Y (G16));
nand02 ix157 (A0 (G15) , .A1 (G16) , .Y (G9));
nor02 ix215 (A0 (G14) , .A1 (G11) , .Y (G10));
nor02 ix216 (A0 (G5) , .A1 (G9) , .Y (G11));
nor02 ix217 (A0 (G7) , .A1 (G1) , .Y (G12));
nor02 ix218 (A0 (G2) , .A1 (G12) , .Y (G13));
assign scan_out1 = $dummy [0];
endmodule

```

Fig. 4. Scan inserted module generated by DFTAdvisor

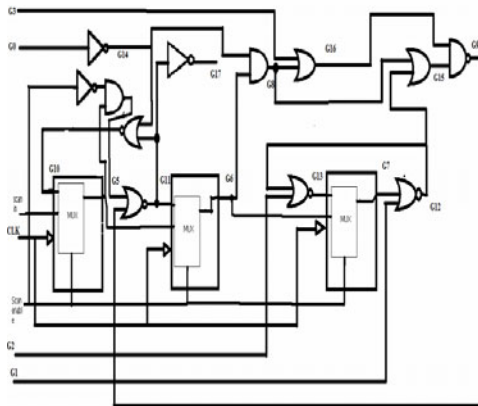


Fig. 5. Freezing Power sensitive scancells by logic gates

4 Conclusion

This paper proposed an effective and efficient method to reduce switching activity during scan shifting. The method identifies a set of power sensitive scan cells and an optimal input control vector through a fast signal probability analysis. By inserting additional hardware to freeze the outputs of the identified scan cells and applying the input control vector during scan shifting, reduction in scan shift power by more than 45% on average for the industrial circuits that are experimented, even if only 5% of scan cells are frozen, can be achieved. Compared with the previous work that freezes all the scan cells, this method achieves a good tradeoff between shift power reduction and area overhead.

References

- [1] Girard, P.: Survey of low-power testing of VLSI circuits. *IEEE Design Test Comput.* 19(3), 80–90 (2002)
- [2] Ravi, S.: Power-aware test: Challenges and solutions. In: *Proc. Int. Test Conf.*, pp. 1–10 (October 2007)
- [3] Butler, K.M., Saxena, J., Jain, A., Fryars, T., Lewis, J., Hetherington, G.: Minimizing power consumption in scan testing: Pattern generation and DFT techniques. In: *Proc. Int. Test Conf.*, pp. 355–364 (October 2004)
- [4] Wang, S., Gupta, S.: An automated test pattern generator for minimizing switching activity during scan testing activity. *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.* 21(8), 954–968 (2002)
- [5] Sankaralingam, R., Oruganti, R.R., Toubia, N.: Static compaction techniques to control scan vector power dissipation. In: *Proc. VLSI Test Symp.*, pp. 35–40 (2000); A Study of Wavelet Thresholding Denoising, *Proceedings of ICSP 2000*
- [6] Wen, X., Yamashita, Y., Kajihara, S., Wang, L.-T., Saluja, K., Kinoshita, K.: On low-capture power test generation for scan testing. In: *Proc. VLSI Test Symp.*, pp. 265–270 (2005)
- [7] Wen, X., Kajihara, S., Miyase, K., Suzuki, T., Saluja, K., Wang, L.-T., Abdel-Hafez, K., Kinoshita, K.: A new ATPG method for efficient capture power reduction during scan testing. In: *Proc. VLSI Test Symp.*, pp. 60–65 (2006)
- [8] Lin, Y.-T., Wu, M.-F., Huang, J.-L.: PHS-fill: A low power supply noise test pattern generation technique for at-speed scan testing
- [9] Kajihara, S., Ishida, K., Miyase, K.: Test vector modification for power reduction during scan testing. In: *Proc. VLSI Test Symp.*, pp. 160–165 (2002)
- [10] Saxena, J., Butler, K.M., Jayaram, V.B., Kundu, S., Arvind, N.V., Sreepakash, P., Hachinger, M.: A case study of IR-drop in structured at-speed testing. In: *Proc. Int. Test Conf.*, pp. 1098–1104 (2003)
- [11] Li, W., Reddy, S.M., Pomeranz, I.: —On reducing peak current and power during test. In: *Proc. IEEE Comp. Soc. Annu. Symp. VLSI*, May 2005, pp. 156–161 (2005)
- [12] Remersaro, S., Lin, X., Zhang, Z., Reddy, S.M., Pomeranz, I., Rajski, J.: Preferred fill: A scalable method to reduce capture power for scan based designs. In: *Proc. Int. Test Conf.*, October 2006, p. 110 (2006)
- [13] Badereddine, N., Girard, P., Pravossoudovitch, S., Landrault, C., Virazel, A., Wunderlich, H.-J.: Structural-based power-aware assignment of don't cares for peak power reduction during scan testing. In: *Proc. IFIP Int. Conf. VLSI*, October 2006, pp. 403–408 (2006)

Rate Adaptive Distributed Source-Channel Coding Using IRA Codes for Wireless Sensor Networks

Saikat Majumder and Shrish Verma

Department of Electronics and Telecommunication,
National Institute of Technology, Raipur, India
{smajumder.etc, shrishverma}@nitrr.ac.in

Abstract. In this paper we propose a scheme for rate adaptive lossless distributed source coding scheme for wireless sensor network. We investigate the distributed source-channel coding of correlated sources when correlation parameter is not fixed or may change during sensor network operation. For achieving rate adaptability we propose the puncturing and extension of IRA code depending on the value of correlation between two sources and the quality of channel. In our scheme we need to transmit only incremental redundancy for decreased correlation or fall in channel quality to meet energy constraints and reducing computation cost.

Keywords: Distributed source-channel coding, IRA code, sensor networks.

1 Introduction

Distributed lossless compression of information is required in many applications like wireless sensor networks. To achieve efficient transmission the sources may be compressed at the individual nodes independently and sent to the fusion center through wireless channel. Compression is required before transmission because wireless sensor nodes are characterized by low power constraint and limited computation and communication capabilities. If there is correlation between the data sent by sensors, then, though they do not communicate with each other, exploiting the correlation would lead to further reduction in number of transmitted bits. Slepian-Wolf theorem [1] shows that lossless compression of two separate correlated sources can be as efficient as if they were compressed together, as long as decoding is done jointly at the receiver. Such a scheme is known as distributed source coding (DSC). Practical schemes for exploiting the potential of Slepian-Wolf theorem were introduced based on channel codes [2], [3], [4], [5] and some of them used modern error correcting codes (e.g. LDPC, Turbo codes) to achieve performance very close to theoretical Slepian-Wolf limit. Schemes for distributed joint source-channel coding has been proposed in literature [6], [7], [8] and the references therein which use a 'good' channel code for jointly performing the operation of compression (source coding) and adding error correction capability (channel coding).

The basic idea of distributed source coding with side information at the decoder is shown in the Fig. 1. It is a communication system of two binary sources X and Y with

conditional probability mass function $P(X|Y) = p$. For many sensor network applications, this statistical dependency between X and Y may not be known before deployment. Instead of designing the code for lowest correlation probability p_{min} , the rate of the code should adapt to the correlation parameter. Designing code for lowest correlation requires larger generator and parity check matrices and it results in longer codes. Punctured low density parity check (LDPC) codes were used [10] for achieving the required rate. Though puncturing reduces the number of transmitted bits, it does not in any way ease the computational and memory requirements of the encoder and decoder. It is because puncturing channel code removes bits after encoding process and reinserts null bits before decoding. Other variations of LDPC code has been used for achieving rate-adaptive codes for distributed source coding [11].

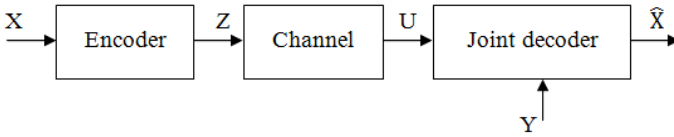


Fig. 1. Distributed source-channel coding with side information at the decoder

The primary contribution of this paper is the development of a scheme for rate-adaptive distributed source-channel coding that has lesser computational and memory requirements than the methods using puncturing only for achieving rate-compatibility. Though puncturing has been the accepted method for achieving rate adaptation, code extension has not been applied in the context of distributed source-channel coding. We apply efficient puncturing [12] and extending [13] methods for obtaining the desired code rate for given value of source correlation and channel quality. Our idea of applying code extension methods for distributed source-channel coding reduces the memory and computational requirements at both sensor nodes and fusion center, besides conserving battery power.

The paper is outlined as follows; Section 2 describes the overall system and explains the relation and tradeoffs between correlation parameter, channel quality and desired BER. Section 3 discusses the methods employed for puncturing and extending IRA codes. Finally in section 4 and 5 we present simulation results and conclusion.

2 System Model

Consider the system of Fig. 1 with the following notations used for rest of the paper. There are two correlated source vectors $X = [x_1, x_2, \dots, x_k]$ and $Y = [y_1, y_2, \dots, y_k]$ of length k . The dependency between these two sources is given by conditional probability mass function $P(x_i|y_i) = p$. Since the correlation parameter p may not be known before deployment, optimal rate of transmission has to be decided during run time. Y is available lossless and error free at the decoder. We try to encode X as efficiently as possible; encoding operation being joint source-channel coding. In this paper we apply systematic irregular repeat accumulate (IRA) codes as joint source-channel encoder using the method in [6].

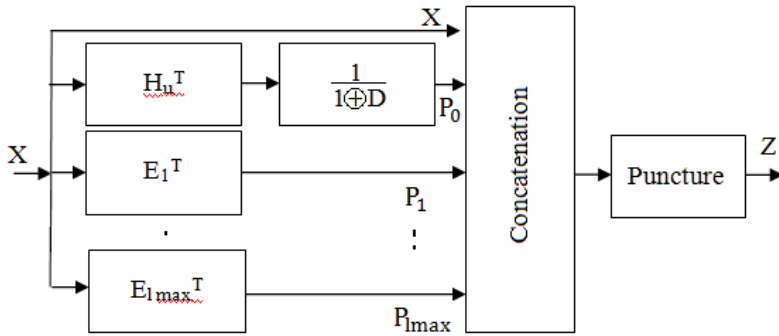


Fig. 2. Rate adaptive source-channel IRA encoder. H_u is the mother code matrix, whereas matrices E_i are obtained by code extension. All the generated parity bits are concatenated to produce the final code word Z . Puncture is applied when increase in code rate is required.

IRA codes, a special class of LDPC codes, are used here in the context of sensor network because they enjoy extremely simple encoding and low complexity decoding. Parity check matrix for systematic IRA code has the form $H = [H_u H_p]$, where H_u is a sparse matrix and H_p is an $m \times m$ dual diagonal matrix. The output of the systematic encoder is $Z = [X P]$, where P is the parity vector. We design the IRA encoder in Fig. 2 which provides incremental redundancy and is based on such an LDPC encoder in [14]. In the encoder matrix H_u^T followed by accumulator block provides the parity check bits P_0 of mother code. This part is same as any conventional IRA encoder. The matrices $E_1, E_2, \dots, E_{l_{max}}$ take into account extra parity bits (for decreasing code rate) provided by code extension. The puncturing block is for increasing the code rate more than the rate of mother code. For example, when rate index l is specified, the concatenated code is given by $Z = [X P_0 P_1 P_2 \dots P_l]$. Parity bits are obtained as $P_0 = XH_u^T H_p^{-T}$ and $P_i = XE_i^T$ for $i = 1, \dots, l$ [13], [15].

When sources are highly correlated or if the channel is good, only $Z = [X P_0]$ is produced and code is punctured to get the required rate. Code design for given rate index (or rate) using extension and puncturing of IRA code is discussed in section 3. The generated code Z and lossless source Y can be jointly decoded using a message passing decoder [15]. The decoding process is same as any IRA decoder, only difference being the initialization of the log-likelihood ratios (LLR) [6]. For unit code word energy and code rate r , LLR of the parity bit nodes is $f(2P - 1)4r/N_0$ and that of information bit nodes is $(2Y - 1) \ln((1 - p)/p) + f(2X - 1)4r/N_0$ where, $P = [P_0 P_1 P_2 \dots P_l]$ and function $f(x)$ indicates effect of channel on signal x .

Next we derive the relation for code rate under channel constraints and varying amount of correlation between the sources. Since Y is a vector of equiprobable binary random variable, its rate is its entropy $H(Y) = kH(y_i) = k$. From Slepian-Wolf theorem, the theoretical limit on rate after lossless compression of X is $H(Z) \geq kH(x_i|y_i) = kH(p)$. For nonideal channel we have to account for the limited channel capacity. For error free communication of a binary source of per bit entropy $H(p)$, using a channel code of rate r , the bit energy E_b required is related as $E_b/N_0 > (2^{2rH(p)} - 1)/2r$ where, $N_0/2$ is the noise power spectral density. Appropriate value

of rate r for given channel condition and conditional entropy $H(p)$ can be found by solving this equation. For finite length codes and continuous channel, the SNR (E_b/N_0) entered has to be few dB higher than actual value. This shift depends on code length and minimum BER required and can be obtained empirically. For example, in our case of IRA code of 1024 information bits and tolerable BER of 10^{-4} , shift required would be of 1.45 dB.

3 Rate Adaptive Code Design

We now elaborate the code design method which uses extended and punctured IRA code for a obtaining a specific rate r . Rate-compatible IRA codes through deterministic extending based on congruential extension sequences [13] is used in our study. Compared to other code extension methods, this method uses low-complexity, algebraic operations without any post-construction girth conditioning. As already mentioned, puncturing of code though reduces the transmitted energy; it does not reduce the memory and computational requirements of encoding and decoding operation. So, starting out with a channel code of lowest anticipated rate and puncturing saves transmission power but not computational resource. Therefore, we consider a mother code of moderate rate $r_0 = k/n_0$ and length n_0 . Corresponding parity check matrix $H = [H_u \ H_p]$ is of size $m_0 \times n_0$. Let $R = \{r_1, r_2, \dots, r_{lmax}\}$ with $r_1 > r_2 > \dots > r_{lmax}$ to be set of desired target rates lower than r_0 for corresponding parity check matrices $\{H_1, H_2, \dots, H_{lmax}\}$. For a particular rate index l , starting from H_0 , we construct

$$H_l = \begin{bmatrix} H_u & H_p & 0 \\ E & 0 & I \end{bmatrix} \tag{1}$$

where, E is obtained by concatenation of sub-matrices E_1, E_2, \dots, E_l as $E^T = [E_1^T, E_2^T, \dots, E_l^T]$. Each of the submatrices E_i is of size $(\epsilon \times k) = (\epsilon \times s\epsilon)$, 0 are all zero matrices, I is a $(\epsilon \times \epsilon)$ identity matrix. $s = \lfloor k/q \rfloor$, where $q = \epsilon$ if ϵ is odd, otherwise $q = \epsilon - 1$. ϵ is the number of rows in E_i and is chosen to have the desired code rates. Two random sequences, $\{\alpha_i: i = 0, \dots, l - 1\}$ and $\{\beta_i: i = 0, \dots, l - 1\}$, with elements from $GF(q)$ is generated. Care must be taken so that no two element in same sequence are same. A matrix $A = [a_{ij}: 0 \leq i \leq l - 1, 0 \leq j \leq s - 1]$ of dimension $l \times s$ is formed with its elements $a_{ij} = [d(\alpha_i + \beta_j)^2](mod\ q)$, where we have taken $d = 1$. Using the elements of A , a new $l \times s$ matrix L is formed with its each element $I(a_{ij})$ being $(\epsilon \times \epsilon)$ identity matrices with rows cyclically shifted to the right by a_{ij} positions. Finally, our extended matrix E is obtained from L by method of masking [9].

To obtain code rates $\{1, \dots, r_{-l}, \dots, r_{-2}, r_{-1}\}$ with $1 > \dots > r_{-l} > \dots > r_{-1} > r_0$ we use the puncturing method in [12]. In this method only parity bits are punctured and punctured node is chosen such that it is at equal distance (in terms of number of nodes) from neighboring unpunctured nodes. The punctured codes of different rates obtained from this type of deliberate puncturing are not rate compatible. To obtain rate compatibility, they have proposed this simple algorithm which performs puncturing in reverse way. In their algorithm, mother code is deliberately punctured to the highest code rate, then unpuncturing those punctured nodes to obtain lower rates.

4 Simulation Results

For simulation we have used mother code with node-wise degree-distribution of $\lambda(x) = 0.6x^3 + 0.4x^8$, rate $r_0 = 0.5$ and $k = 1024$ as in [13]. All the simulations are done for a binary phase shift keying (BPSK) modulated AWGN channel. The iterative sum product algorithm (SPA) was used for decoding, and the maximum number of decoding iterations is 100. We simulate the error performance of distributed joint source-channel coding of two correlated sources for various rates and correlation parameter. We then compare the memory requirements for obtaining various rates.

In our case mother code rate is 0.5. Code of rate 0.8 and 1/3 is obtained by puncturing and extension, respectively. Fig. 3 gives the BER plot for two different rates $r = \{1/3, 0.8\}$ and different values of correlation parameter. These plots clearly show that for given p , any suitable rate r_l can be chosen in accordance to the channel condition. As can be seen from the plots the code performance is the about 2 dB away from the Shannon's limit. This gap can be further reduced by increasing the code length, which may not be always possible for low cost sensor nodes.

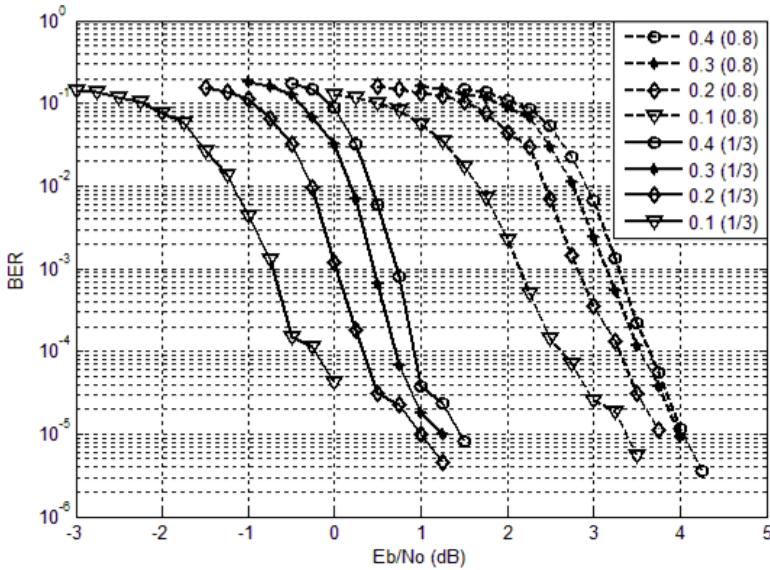


Fig. 3. Simulation results for joint source-channel decoding of source X for different value of correlation parameter p . The values in parenthesis beside correlation values are code rates.

Table 1 compares the memory requirements with an algorithm based on [10] but using systematic IRA code instead of systematic LDPC code. The numbers are indicator values only and are obtained for simulation program running on Matlab. It can be seen that if we start out with a medium rate code, our memory requirements are significantly less than the case where one starts with lower rate code and uses puncturing only.

Table 1. Memory utilization (in kilobytes) for encoding and decoding processes at different code rates. Two different rate adaptation methods are compared, our method utilizing both puncturing and extension, and other one uses code puncturing only.

Code rate	Rate adaptation by puncturing and extension (our method)		Rate adaptation by puncturing only (mother code rate is 1/3)	
	Encoder	Decoder	Encoder	Decoder
8/24	256	1351	256	1351
8/20	192	1110	256	1351
8/18	160	989	256	1351
0.5	128	869	256	1351
0.6	128	869	256	1351
0.7	128	869	256	1351
0.8	128	869	256	1351

5 Conclusion

We proposed a scheme for lossless distributed source-channel coding of correlated sources with side information using extended and punctured IRA codes. Extension and puncturing is used for achieving the required rate, which in turn depends on correlation between sources and channel condition. The simulation results confirm that proposed method can achieve any required rate for given correlation and channel condition. We have shown that the use of code extension for rate adaptation for joint source-channel coding requires lesser memory than the methods using puncturing only. Thus, besides saving transmission power by adapting to most suitable rate, our method also minimizes the memory requirement for same code rate.

References

1. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley-India, New Delhi (2006)
2. Pradhan, S.S., Ramchandran, K.: Distributed Source Coding Using Syndromes (DISCUS): Design and Construction. In: Proceedings of IEEE DCC, pp. 158–167 (1999)
3. Stankovic, V., Liveris, A.D., Xiong, X., Georghiades, C.N.: On Code Design for Slepian-Wolf Problem and Lossless Multiterminal Network. IEEE transactions on Information Theory 52(4) (2006)
4. Garcia-Frias, J., Zhao, Y.: Compression of Correlated Binary Sources Using Turbo Codes. IEEE Communication Letters 5(10) (2001)
5. Fresia, M., Vandendorpe, L., Poor, H.V.: Distributed Source Coding Using Raptor Codes for Hidden Markov Sources. IEEE Transactions on Signal Processing 57(7) (2009)
6. Liveris, A.D., Xiong, Z., Georghiades, C.N.: Joint Source-Channel Coding of Binary Sources with Side information at the Decoder Using IRA Codes. In: Proceedings of IEEE Workshop on Multimedia Signal Processing, pp. 53–56 (2002)
7. Garcia-Frias, J.: Joint Source-Channel Decoding of Correlated Sources Over Noisy Channels. In: Proceedings of IEEE DCC (2001)

8. Zhao, Y., Garcia-Frias, J.: Turbo Compression/Joint Source-Channel Coding of Correlated Sources With Hidden Markov Correlation. *Elsevier Signal Processing* 86, 3115–3122 (2006)
9. Xu, J., Chen, L., Djurdjevic, I., Lin, S., Abdel-Ghaffar, K.: Construction of Regular and Irregular LDPC Codes: Geometry Decomposition and Masking. *IEEE Transactions on Information Theory* 53(1) (2007)
10. Sartipi, M., Fekri, F.: Distributed Source Coding Using LDPC Codes: Lossy and Lossless Cases with Unknown Correlation Parameter. In: *Proceedings of Allerton Conference on Communication, Control and Computing* (2005)
11. Varodayan, D., Aaron, A., Girod, B.: Rate-adaptive Codes for Distributed Source Coding. *Elsevier Signal Processing* 86, 3123–3130 (2006)
12. Yue, G., Wang, X., Madhian, M.: Design of Rate-Compatible Irregular Repeat Accumulate Codes. *IEEE Transactions on Communications* 55(6) (2007)
13. Benmayor, D., Mathiopoulos, T., Constantinou, P.: Rate-Compatible IRA Codes Using Quadratic Congruential Extension Sequences and Puncturing. *IEEE Communications Letters* 14(5) (2010)
14. Li, J., Narayanan, K.: Rate-Compatible Low Density Parity Check Codes for Capacity-Approaching ARQ Schemes in Packet Data Communications. In: *Proceedings of IASTED CIIT*, pp. 201–206 (2002)
15. Ryan, W.E., Lin, S.: *Channel Codes: Classical and Modern*. Cambridge University Press, New York (2009)
16. Hou, J., Siegel, P.H., Milstein, L.B.: Performance Analysis and Code optimization of Low Density Parity-Check Codes on Rayleigh Fading Channels. *IEEE Journal on Selected Areas in Communication* 19(5) (2001)

Web Cam Motion Detection Surveillance System Using Temporal Difference and Optical Flow Detection with Multi Alerts

V.D. Ambeth Kumar¹ and M. Ramakrishan²

¹ Research Scholar, Sathyabama University, Chennai, India
ambeth_20in@yahoo.co.in

² Professor and Head, Velemmal Engineering College, Chennai, India
ramkrishod@gmail.com

Abstract. This paper proposes a method for detecting the motion of a particular object being observed. The Motion tracking Surveillance has gained a lot of interests over past few years. The Motion tracking surveillance system is brought into effect providing relief to the Normal video surveillance system which offers time-consuming reviewing process. Through the study and Evaluation of products and methods, we propose a Motion Tracking Surveillance system consisting of its method for motion detection and its own Graphic User Interface. Various methods are used in Motion detection of a particular interest. Each algorithm is found efficient in one way. But there exists some limitation in each of them. In our proposed system those disadvantages are omitted and combining the usage of best method we are creating a new motion detection algorithm for our proposed Motion Tracking Surveillance system. The proposed system in this paper does not have its effect usage in office alone. It also offers more convenient, effective and efficient usage in home.

Keywords: Motion detection, Surveillance, GUI, Real time environment.

1 Introduction

The task of a Motion Tracking Surveillance system is to detect a motion present in “region of awareness”, where the region of awareness, or the field of view, is defined as the “Portion of Environment (or) surrounding being monitored” [1]. The motion of moving objects is the activity of the portion of environment may be represented as Region of interest. The system captures images only when the motions exceed a certain threshold that is present in system. It thus reduces the volume of data that needs to be reviewed and is therefore a more convenient way of monitoring the environment, especially with the increasing demand for Multi-camera. Also, it helps to save Data space by not capturing static images which usually do not contain the object of interest.

There are two Main components that concern basic Motion Tracking Surveillance software, (i.e) GUI and method for Motion Detection. As part of the literature review, we evaluated 4 popular motion detection surveillance products that are currently

available in market. We considered about the terms of format and their features which are based on features that are required of a Surveillance system and the additional features that are required for purpose of motion detection. The four products, namely 'Active Webcam' by py software [2], 'Watcher' by Digi-Watcher [3], 'FG Engineering Surveillance 4 cam Basic' by FG Engineering [4], and 'Supervision Cam' by Peter Kirst [5] are chosen by considering users feedback.

Also in the literature review, the existing methods for motion detection are discussed. They include some of popular methods, such as Temporal Difference[1][6] and background modelling [1][6][7][8]; as well as methods that are not so widely used due to certain constraints eg. Optical flow [9][10] and Spatio-Temporal entropy[11].

2 Motion Detection

Methods for motion detection can be categorized into 2 Main classes, i.e Region-based and pixel-based algorithm [8]. The former, based on Spatial dependencies of neighboring pixel colours to provide results more robust to false alarms. The later based on binary difference by employing local or pixel-based model of intensity, is a simple model often used in real-time application.

2.1 Optical Flow

It is the 2-D velocity field induced in an image due to projection of moving objects onto the image plane. An optical flow shows velocity of each pixel in image. Most optical flow techniques assume that uniform illumination is present [12]. However only small movements can be accurately detected in gradient technique due to Taylor's approximation of the gradient constraint equation.

2.2 Spatio-Temporal Entropy

It works on assumption that pixel state change brought about by noises would be in small range while those brought about by motion would be large. However, it is impossible to predict all types of noises, thus accuracy of detection is most of time not to satisfied level.

2.3 Temporal Difference

The consecutive frames are compared on basis of pixel by pixel basis for calculating motion sequence and a threshold is applied for classification of objects in stationary or in motion. But when there is motion in objects, the image intensities are not varied in a short time interval [13]; so that motion is detected only in boundaries. Also it makes more false alarms as they don't show relationship of pixel with neighbourhood.

2.4 Background Modelling

Background modelling methods are classified into two types namely pixel-based and region-based models .Background subtraction is often used in case of pixel-based. An

image of the stationary background is generated by averaging the image sequence over a period of time on mixture of Gaussian distribution. With the help of Gaussian density the likelihood of each incident pixel colour is computed. On the deviation between the two pixels, they will be labelled as foreground pixels and detected as motion. Krumm et al[14] has done an excellent analysis on this method.

There are two popular methods for region-based background modelling, the three tiered algorithm which process image at the pixel, region and frame level[14], and the eigen-space decomposition method[8].

3 Proposed Method

We propose a method that uses both temporal difference and optical flow methods together with morphological filter for the purpose of motion detection. Figure 1 gives an flowchart of the proposed motion detection process.

3.1 Reducing the Number of Pixel Using Temporal Difference Method

The initial image frame is provided as the input for our proposed method. The image frame with reduced number of pixel is obtained as output.

First of all, the temporal difference method is used to obtain an initial coarse image so as to reduce the number of pixels that a downstream tracking algorithms have to process. It is used despite its shortcomings as it is a cheap and simple method. No knowledge of the background is required as in background subtraction where training periods in the absence of foreground are required for bootstrapping [14]. Also, the method's inability to detect the entire shape of objects of interest does not act as a serious obstacle considering the fact that the system is for simple office and home uses.

3.2 Image Denoising Using Optical Flow Method

The image frame with reduced number of pixels is given as the input. Noise pixels are removed in the output.

The optical flow method is then used to further analyze the detected motion area to reduce noises such as movements of trees and the capturing device. Tian et al[15] mentioned that objects of interest tend to have consistent moving directions over time. Mittal et al [16] also stated that persistent motion characteristics are shown by most objects of interest. Therefore, optical flow is a suitable method in this aspect as it is able to estimate the direction and speed of moving objects to reduce false alarms incurred by undesirable external factors like oscillating fans and trees.

3.3 Morphological Filtration

The morphological filter is then applied which is used to suppress noises while preserving the main object characteristics [17]. It consists of ways for digital image processing based on mathematical morphology which is a nonlinear approach developed based on set theory and geometry. It is able to decompose complex shapes into meaningful parts and separate them from the background. In addition, the mathematical calculation involves only addition, subtraction and maximum and minimum operations with no multiplication and division.

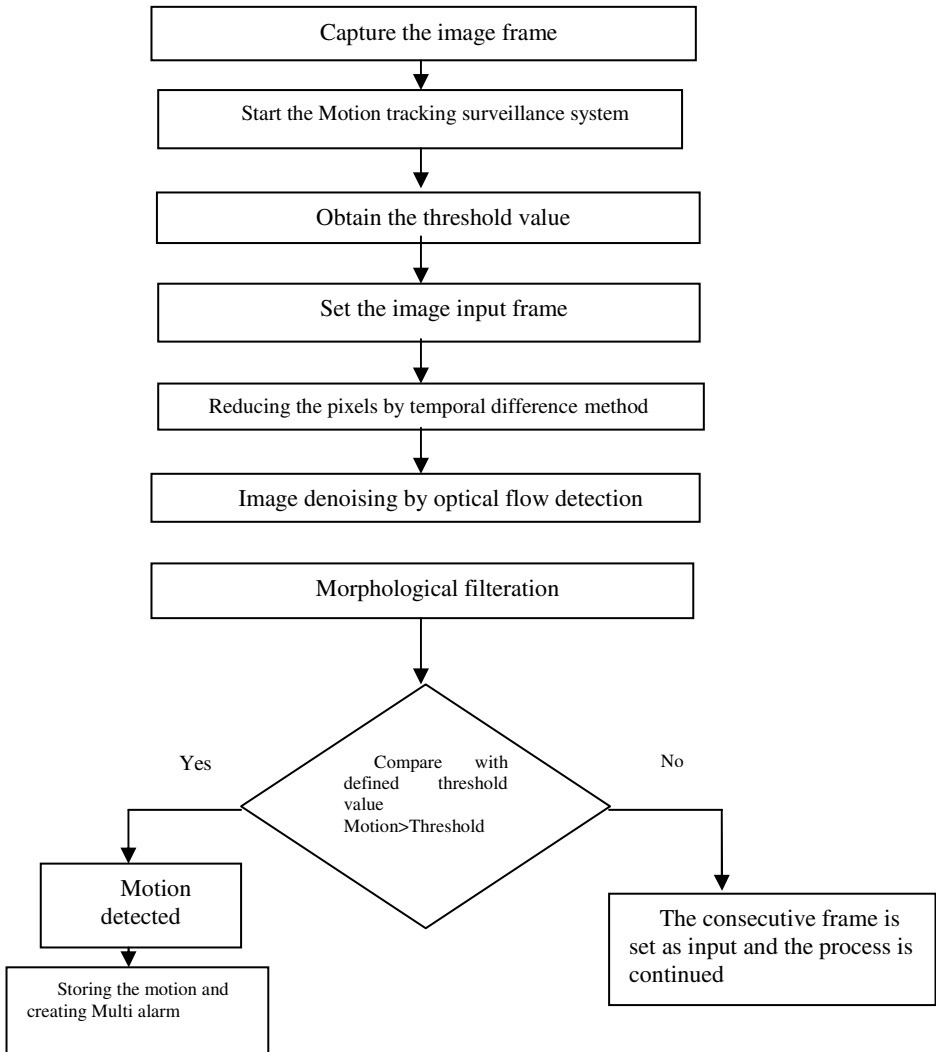


Fig. 1. Flowchart of motion detection process

The two fundamental morphological operations are dilation and erosion on which many morphological algorithms are based on. Experiments done by Lu et al [18] proved that the method is effective in preserving moving object areas and eliminating noises. Finally, a binary image is generated with the motion region coloured white and regions with no motion detected coloured black.

3.4 GUI

One guiding principle in designing the GUI of commercial software is user-friendliness. This is most of the times determined by two factors, format and the

features. Clear and neat formatting makes it easier for users to access necessary features in the system gives rise to neat formatting.

Two basic formats of GUI are used to design the format of GUI. Type 1 is the GUI in which users can easily switch from capture panel to log panel for monitoring and reviewing. Type 2 is the GUI is the more conventional GUI design whereby users have to click on the menu bar every time for another window to appear.

4 Results and Discussion

Our aim is to present a new technology with unique features. We are going to create three types of alerts: 1.Audio alert 2. Video alert 3. E-mail alert. This will be explained in detail in the forth coming paragraphs. This model is based on GUI which makes the user to familiarize the concept. The function of the settings, camera and log is shown in the picture. These functions give a easy access to the users.

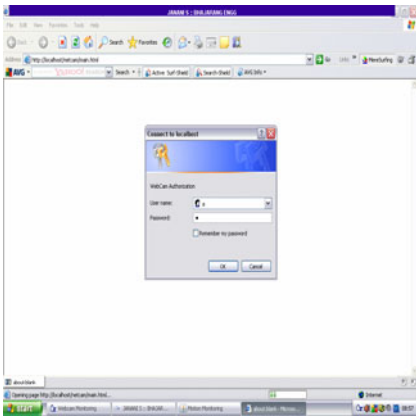


Fig. 2. Login panel

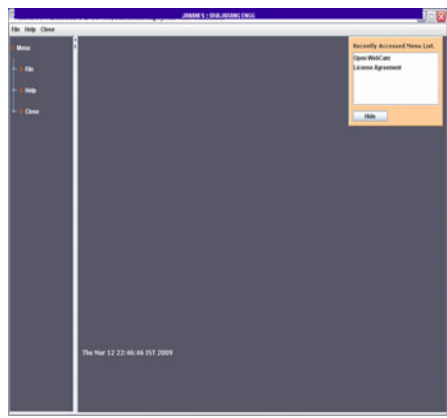


Fig. 3. Camera panel

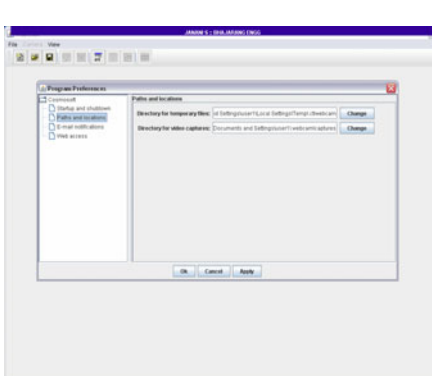


Fig. 4. Video alert

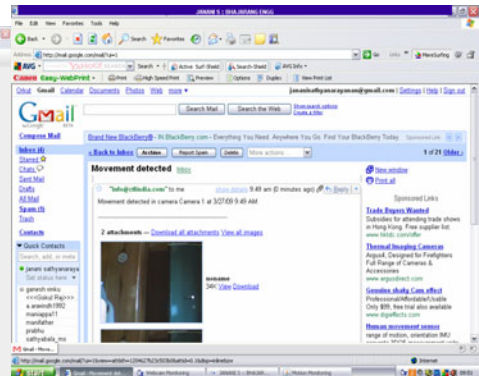


Fig. 5. E-mail alert

The “Camera” panel (shown Fig 3) displays detected motions that automatically pop up on the main window for close monitoring. The “Log” panel (shown Fig 2) displays the images saved by the system.

Now we discuss on the three types of alerts first Audio alert will produce the audible sound when any motion is detected across the prohibited region. Next the visual alert (shown Fig 4) indicates the occurrences of the motion by changing the color displayed in the monitor. When the normal video is playing a grey strip will be seen at the top of the video.

But when the motion is detected the grey color strip changes to red and the play symbol is also displayed along with that strip. Finally E-mail alert (shown in Fig 5) generates the alert as mail with frame attachments.

5 Comparison with Other Methods

The detection accuracy of motion detection using proposed algorithm is compared with other detection methods .The following table shows the comparison of number of alerts and detection accuracy of motion detection for images.

Table 1. Comparison of Alert (Performance) of the Proposed Method with Other Motion Detection Methods

Motion Detection Methods	No.of Alerts	Detection Accuracy (%)
Optical Flow [Ref.14]	01	83.34
Temporal Difference [Ref.12]	01	87.43
Proposed Method	03	94.375

The corresponding graph for the Alert for motion detection using the above mentioned methods is given as under.

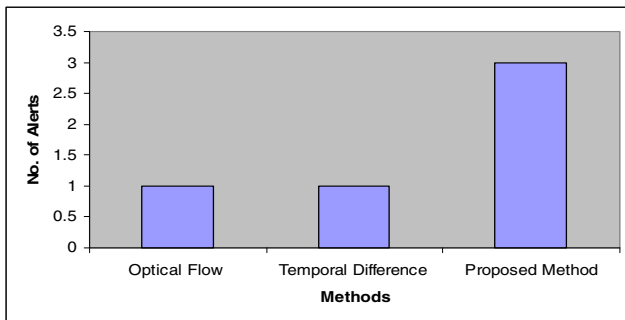


Fig. 6. Comparison of Alert of Other Motion Detection Methods with Proposed Method

The corresponding graph for the Detection Accuracy for motion detection using the above mentioned methods is given below.

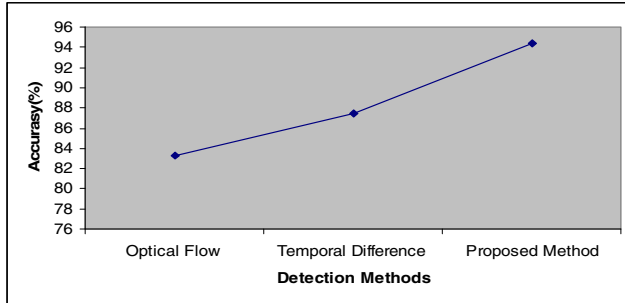


Fig. 7. Comparison of Accuracy in % of Existing Detection Methods with (OF, TD) with Proposed Method

6 Conclusion

The software system is capable of working at any circumstances. This is simple to install and access it. This system provides an enhanced feature of email alert, run in surveillance mode, live view about the happenings of that particular area etc. This system is very powerful since it uses motion detection algorithm to compare the frame. These frames are compared by pixels when there is any change even in the single pixels they indicate the alerts to the user. We are using J2SE for our implementation. This will be a more convenient way of monitoring the real time environment, especially with the increasing demand for multi-camera.

References

1. Stefan Gachter, "Motion Detection as an Application for the Omnidirectional Camera", Research Reports of CMP, Czech Technical University in Prague, Omnidirectional Visual System (7), 5-13 (2001); FPS RTD – FET, Project No: IST-1999-29017
2. About Active Web Cam, <http://www.pysoft.com/ActiveWebCamMainpage.htm>
3. Digi-Watcher Features of our webcam software for video surveillance, http://www.digi-watcher.com/surveillance_features.htm
4. <http://www.fgeng.com-vsmfeatures.htm>
5. SupervisionCam Homepage English, <http://www.supervisioncam.com/>
6. Abdelkader, M.F., Chellappa, R., Zheng, Q.: Integrated Motion Detection and Tracking for Visual Surveillance. In: Proc. Of the Fourth IEEE International Conference on Vision Systems (ICVS 2006), pp. 1–3 (2006)
7. Miezianko, R.: Motion Detection and Object Tracking in Grayscale Videos Based on Spatio Temporal Texture Changes. In: Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy. Temple University Graduate Board, pp. 8–14 (2006)

8. Sheikh, Y., Shah, M.: Bayesian Modeling of Dynamic Scenes for Object Detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 27(11), 1778–1780 (2005)
9. Rivo, J.-E.S., Cajote, E.R.: Object Motion Detection Using Optical Flow. In: *Digital Signal Processing Laboratory*, Department of Electrical and Electronics Laboratory, University of the Philippines, pp. 1–2
10. Sarker, M.H., Bechkoum, K., Islam, K.K.: Optical Flow for Large Motion Using Gradient Technique. *Serbian Journal of Electrical Engineering* 3(1), 103–113 (2006)
11. Ma, Y.-F., Zhang, H.-J.: Detecting Motion Object by Spatio-temporal Entropy. In: *Proc. IEEE International Conference on Multimedia and Expo.*, pp. 265–268 (2001)
12. Zelek, J.S.: Bayesian Real-time Optical Flow. In: *School of Engineering*, University of Guelph
13. Davis, L.: Time varying image analysis,
<http://www.umiacs.umd.edu/~lsd/426/motion.pdf>
14. Krumm, J., Toyama, K., Brumitt, B., Meyers, B.: The Ultimate Futility of Background Subtraction. *Int. Journal of Computer Vision*, 3–8 (Submitted)
15. Tian, Y.-L., Hampapur, A.: Robust Salient Motion Detection with Complex Background for Real-time Video Surveillance. *IBM T.J. Watson Research Center*, 1–6
16. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 2nd edn. Prentice-Hal, Englewood Cliffs (2002); ISBN: 0130946508
17. Mittal, A., Paragios, N.: Motion-Based Background Subtraction Using Adaptive Kernel Density Estimation. In: *Real-Time Vision and Modeling* Siemens Corporate Research Princeton and C.E.R.T.I.S. Ecole Nationale de Ponts etc Chaussees Champs sur Marne, France, pp. 1–8
18. Lu, N., Wang, J., Wu, Q.H., Yang, L.: An Improved Motion Detection Method for Real-Time Surveillance. *IAENG International Journal of Computer Science* 35, 1, *IJCS_35_1_16*, 1–10

Rts-Mirror: Real Time Synchronized Automated Rear Vision Mirror System

Kuldeep Verma¹, Ankita Agarkar¹, and Apoorv Joshi²

¹ Malwa Institute of Technology, Nipaniya, Bypass Road, Indore 452 016, India
{lucky_kd_007, ankita_agarkar}@yahoo.com

² Royal College of Technology, Opp. Pitra Parvat, Jamburdi Hapsi, Indore 452 003, India
apoorv76652@gmail.com

Abstract. In this paper, we present ‘RTS-Mirror’, an attempt to transform the overall usage and functioning of rear vision mirror to an entire stage of compliance in accordance to the automated technology of rear mirror that will leverage the present technology of movement of the rear vision mirror by a joystick or rather by a human hold. ‘RTS-Mirror’ is an intelligent system running on RTOS that tracks the face of driver in real-time and will apparently adjust the rear mirror in a contiguous synchronized manner. This project explores the use of face tracking algorithm like Haar-Classifer cascade algorithm [3], [5], Lucas-Kanade algorithm [14], [15], and joystick mechanism of rear mirror can develop a fully automated digital Rear vision mirror.

Keywords: Joystick controlled rear mirror, Facial features recognition algorithm, Synchronized movement of Mirrors, Image Tracking.

1 Introduction

The affixing of a rear vision mirror in Marmon race car at the inaugural Indianapolis 500 race [1] in 1911 ascended the importance of the newly invented automobile part. The main area of development occurred so far concerns with the optical area counting such as HCM, Electro-chromic Mirror [2], [4], [7]. Despite this, human palm interference still exists in adjusting the rear mirror.

‘RTS-Mirror’ is an endeavour to integrate a camera into the central dashboard rear mirror that will track the facial position of the driver guided by Real Time Operating System (RTOS) for computing the critical angle at real time. RTS-Mirror primarily functions with the aid of two different algorithms, namely *Haar-Classifer cascade algorithm* [3], [5] and *Lucas-Kanade algorithm* [14], [15]. Lastly, the joystick mechanism of rear vision mirror will be embedded into RTS-Mirror technology with input provided by the resultant of the algorithm.

2 Related Works

Several projects and products have tried to meld the simple rear vision mirror into a more compound form that claims to automate the rear vision mirror using joystick

interface [2], etc. But all in all, no such project has been developed that claims to have facial gestures as the primary input for adjustment of the rear vision mirror.

This project is a blend of two different projects; the facial position detection and the joystick oriented automated approach. But still it is solely a genuine project with extreme possibilities, which is still unlisted in any sort of currently developing projects.

3 Statistical Data

According to the National Highway Traffic Safety Administration (NHTSA) in Washington D.C., there are approximately 840,000 side-to-side blind spot collisions with 300 fatalities every year in the United States [8]. These accidents are due to the incorrect adjustment of the rear vision mirror. By carefully examining the accurate angle that gives the best viewing angles, three cases has arrived.

3.1 Accurate Angle Adjustment

Accurate/Critical viewing angles, showed in fig. 1, is calculated by the derivations below. These angle measurements will be patched in RTOS environment to maintain a rough idea about the actual critical angles [9], [10], [11].

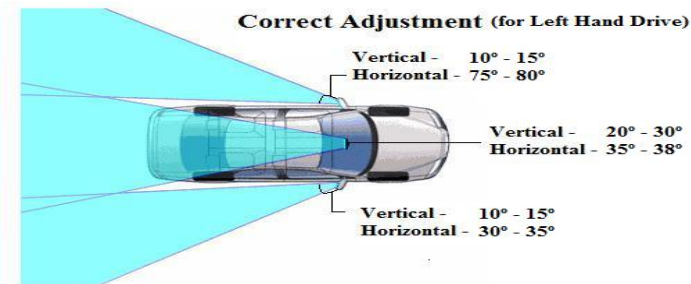


Fig. 1. Approximately calculated adjusting angles for all three rear view mirrors of car delivering the best area covered in vision

4 RTS-Mirror

4.1 What Is RTS-Mirror?

RTS-Mirror is an integrated project that delivers an intelligent environment that acts at the change in posture of rider (though slight). RTS-Mirror tracks the facial specific parts such as eyes as main concern which is represented as dots in algorithm. Afterwards, the system adjusts the mirror using Joystick controlled mechanism. The input is given by the algorithm and thus eliminating any interference of human palm.

4.2 How Does RTS-Mirror Work?

A high-level overview of how the RTS-Mirror works is shown in Figure 2 by flow chart. Firstly, the embedded camera in the rear vision mirror primarily tracks the motion in its perspectives. Consecutively, it uses Haar-Classifer Cascade Algorithm to filter the human face from the tracked motion. Successively, the next algorithm Lucas-Kanade Algorithm will tracks the features of the face. Later, it calculates the critical angle [6] which results the calculation of accurate angle, resulting in synchronization phase of all the mirrors in correspondence with the central rear view mirror. This mechanism will be handles by RTOS environment. The mirror's assembly includes an actuator, capable of transforming electrical signals into a mechanical force, most commonly a torque applied to an output shaft. The mirrors are mechanically attached to the output of the actuator through a mechanical means, such as a threaded shaft.

4.3 Calculating Critical Angle

Critical angle is an angle that provides the accurate viewing angle perfectly appropriate for the driver [6], [11]. Primarily, we consider a horizontal plane angle from the operator's eyes to the central rear view mirror, θ_{IR} , where δ_{IR} represents the horizontal plane adjustment angle for the manually adjusted internal rear view mirror. θ_{IR} can be estimated as, follows,

$$\theta_{IR}=2\delta_{IR} \quad (1)$$

Now, given θ_{IR} and applying a simple trigonometric function based upon a right triangle having perpendicular side d_{IR} and I_{IR} , the longitudinal distance from the operator's eyes to the rear view mirror.

$$I_{IR} = d_{IR} \tan (2\delta_{IR}) \quad (2)$$

Thus, the horizontal plane angle from the operator's eyes to the left side mirror, θ_{LS} , utilizing a right triangle with perpendicular sides d_{LS} and $e+I_{IR}$, can be expressed as follows, and similarly, the horizontal plane angle from the operator's eyes to the right side mirror, θ_{RS} , can be expressed as follows:

$$\theta_{LS} = \tan^{-1}(d_{LS} * e+I_{IR}) = \tan^{-1}(d_{LS} * e+d_{IR}\tan(2\delta_{IR})) \quad (3)$$

$$\theta_{RS} = \tan^{-1}(d_{RS} * e+I_{IR}) = \tan^{-1}(d_{RS} * e + d_{IR}\tan(2\delta_{IR})) \quad (4)$$

Calculated thusly, input angles from a manual adjustment of a rear view mirror can be used to generate side mirror adjustment angles [12], [13].

4.4 Implementation

The implementation phase of this project is still in its coding part. At this instance, the camera is programmed to track any motion. This part has been designed and coded with Aforge.NET framework that provides efficient classes of imaging and visual related forms, etc. The next phase will be consisting of using specified algorithms to monitor only the facial parts of human (driver). Afterwards, the synchronization of the side mirrors in accordance with the central rear mirror will be managed. Lastly,

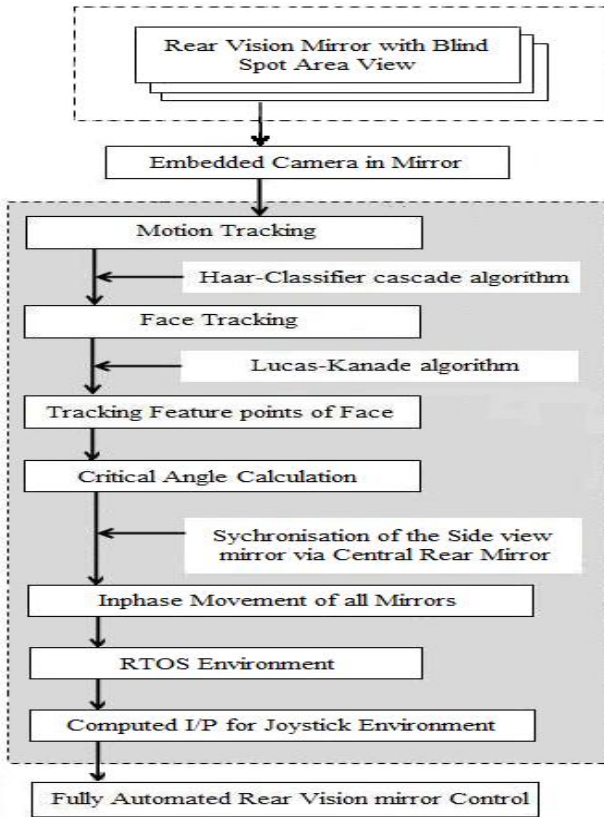


Fig. 2. Detailed System design of 'RTS-Mirror'

the project will consist of embedding the above proposed software into the present joystick technology.

The overall system design is implemented in accordance with the Flow chart (fig.2). All in all, the fully automated system will be developed with RTOS control delivering accurate angles of the mirror at real time providing efficient removal of blind spots as well.

5 Results and Conclusion

The starting result of developing RTS-Mirror has been substantial with high grade statistical data. Also, the implementation of the algorithms has provided high performance characters to the proposed system. 'RTS-Mirror' system bridges the gap between the mechanical and digital world by providing an interface, linking the existing technology of joystick mechanism with the advanced motion tracking technology. By augmenting the familiar and ubiquitous rear vision system, the system leverages existing patterns of behaviour, merging fully mechanical mechanism with a user's digital world.

Acknowledgements

We thank our Professor Divya Gautam and Lokendra Songare for their valuable feedbacks. We would also like to thank Sapan Jain, post graduate student of Texas A&M University who had contributed ideas and time to this research.

References

1. Ward's AutoWorld magazine, Ward's Auto World: Rearview Mirror webpage on Ward-sauto (2002), http://waw.wardsauto.com/ar/auto_rearview_mirror/
2. Murakami Corporation, <http://www.murakami-kaimeido.co.jp/english/product/mirror/pro04.html>
3. Adolf, F.: How-to build cascade of boosted classifiers based on Haar - like features, http://robotik.inflomatik.info/other/opencv/OpenCV_ObjectDetection_HowTo.pdf
4. Murakami Corporation, Hydrophilic Clear Mirror (HCM), <http://www.murakami-kaimeido.co.jp/english/product/mirror/pro01.html>
5. Wilson, P., Fernandez, J.: Facial Feature Detection using Haar Classifiers. Texas A&M University – Corpus Christi, JCSC 21(4) (April 2006)
6. Ryu, J., Shin, K., Litkouhi, B., Lee, J.: Automatic Rearview Mirror adjustment system for vehicle. USPC class 70149, Patent application number 20100017071
7. Murakami Corporation Electro-Chromic Mirror, <http://www.murakami-kaimeido.co.jp/english/operation/mirror/product/pro02.html>
8. AccidentMI.com, 5674 Lancaster, Commerce, MI 48382, <http://www.accidentmi.com/accidentstatistics.html>
9. Hannah, J., Palin, B., Kratsch, C., Gaynor, B.: <http://www.wikihow.com/Set-Rearview-Mirrors-to-Eliminate-Blind-Spots>
10. The Web Page of Kristopher Linquist, <http://www.linquist.net/motorsports/tech/mirrors/>
11. PhysicsForums, <http://www.physicsforums.com/showthread.php?s=63a4705cd010d3b75fca0df6d3ff6423&t=442701>
12. Adjusting Your Mirrors Correctly, <http://www.smartmotorist.com/car-accessories-fuel-and-maintenance/adjusting-your-mirrors-correctly.html>
13. Part571.111: Standard No. 111; Rearview Mirrors, <http://www.fmcsa.dot.gov/rules-regulations/administration/fmcsr/fmcsrruletext.aspx?reg=r49CFR571.111>
14. Lucas, B.D., Kanade, T.: Lucas–Kanade Optical Flow Method, http://en.wikipedia.org/wiki/Lucas%E2%80%93Kanade_Optical_Flow_Method
15. Lucas, B.D., Kanade, T.: Computer Science Department, Carnegie-Mellon University, <http://cseweb.ucsd.edu/classes/sp02/cse252/lucaskanade81.pdf>

Fuzzy Based PSO for Software Effort Estimation

P.V.G.D. Prasad Reddy¹ and CH.V.M.K. Hari²

¹ Department of Computer Science & Systems Engineering, Andhra University
prasadreddy.vizag@gmail.com

² Department of IT, Gitam Institute of Technology, GITAM University
kurmahari@gmail.com

Abstract. Software Effort Estimation is the most important activity in project planning for Project Management. This Effort estimation is required for estimation of resources, time to complete the project successfully. Many models have been proposed, but because of differences in the data collected, type of projects and project attributes, no model has been proven successful at effectively and consistently predicting software development effort due to the uncertainty factors. The Uncertainty in effort estimation controlled by using fuzzy logic and the parameters of the Effort estimation are tuned by the Particle Swarm Optimization with Inertia Weight. We proposed three models for software effort estimation using fuzzy logic and PSO with Inertia Weight. The valued effort is optimized using the incumbent archetypal and tested and tried on NASA software projects on the basis of three touchstones for assessment of software cost estimation models. A comparison of the all models is done and it is found that the incumbent archetypal cater better values.

Keywords: Person Months, Thousands of Delivered Lines of Code, Fuzzy Logic System, Particle Swarm Optimization, Software Effort Estimation.

1 Introduction

Software project management activities are mainly classified into three categories: project planning, project monitoring and control, project termination. In project planning, cost estimation is one of the most important activity. Software cost estimation is the process of predicting how much amount of effort is required to build software. The effort is measured in terms of person – months, thus later on it is converted into actual dollar – cost. The basic input parameters for software cost estimation is size, measured in KDLOC (Kilo Delivered Lines of Code). A number of models have been evolved to establish the relation between Size and Effort [9]. The parameters of the algorithms are tuned using Genetic Algorithms [3], Fuzzy models [4], Soft-Computing Techniques [5][7], Computational Intelligence Techniques[5], Heuristic Algorithms, Neural Networks, Radial Basis and Regression [8][9].

Fuzzy logic [11][12] has rapidly become one of the booming technologies in today's world for developing convoluted control systems. Fuzzy logic is so efficacious that it can address any such applications perfectly as it resembles human decision making with an ability to generate precise solutions from expurgated or unexpurgated information. It is an interlude in engineering design methods, left vacant by purely

mathematical approaches and purely logical approaches in system design. While an alternative approaches require accurate equations to statuette real -world behaviours, fuzzy design can invegle the ambiguities of real-world human language and logic. It provides both an intuitive method for describing systems in human terms and automates the conversion of those system specifications into effective models.

A paradigmatic fuzzy system comprises of three crucial components [11][12], are shown in figure 1, fuzzifier, fuzzy inference engine (fuzzy rules), and defuzzifier. The fuzzifier metamorphose the input into linguistic monikers using membership function that delineate how much a given numerical value of a particular variable fits the linguistic term mapping between the input membership functions and the output membership function using fuzzy rules that can be obtained from proficient knowledge over relationships being modelled. The greater the input membership degree, the stronger the pull towards the output membership functions can be contained in the consequents of the rules triggered.

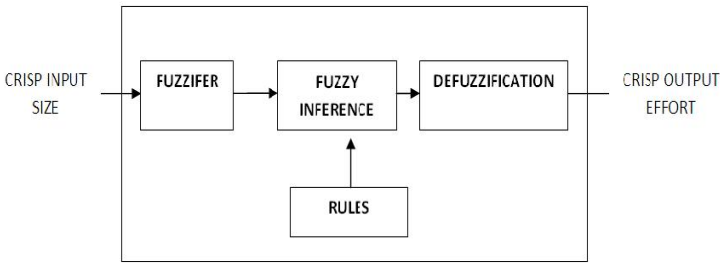


Fig. 1. Fuzzy Logic System

A defuzzifier carries out the defuzzification process to amalgamate the output into a single expression or numerical value as required.

2 Fuzzy Logic and PSO for Software Effort Estimation

The accuracy of the estimate will depend on the amount of accurate information, we have about the final product. The project is being initiated and estimated all alone, during the feasibility study. Specifications with bewilderment represent a range of possible final products and not one precisely defined product. Hence, the cost estimation cannot be accurate, the uncertainties are reduced, and more accurate cost estimates can be made. In order to reduce the skepticism, we use fuzzy logic. The cost for a project in a function of many parameters. Foremost amongst them in size of project in order to reduce the skepticism at the input level i.e. size, we use triangular membership function, this process is known as fuzzification. The parameters of the cost model equation are tuned by using PSO algorithm. By applying fuzzy inference, we get the suitable equation, finally defuzzication is done through weighted average method which is actually translates fuzzy values into output.

Particle Swarm has two primary operators: Velocity update and Position update. During each generation each particle is accelerated toward the particles previous best position (Pbest) and the global best position (Gbest). The inertia weight is multiplied

by the previous velocity in the standard velocity equation and is linearly decreased throughout the run. This process is then iterated a set number of times or until a minimum error is achieved.

The basic concept of PSO lies in accelerating each particle towards its Pbest and Gbest locations with regard to a random weighted acceleration at each time [12].

3 Model Description

In this model, we considered input sizes as fuzzy values, the parameters a, b, c, d are tuned by using PSO with inertia weight algorithm. The following in the methodology is used to estimate the effort.

3.1 Methodology (Algorithm)

Input: Size of the software project, measured effort, methodology.

Output: Optimized parameters and Estimated Efforts.

STEP 1: FUZZIFICATION: The input size in fuzzified by using triangular membership Function shown in Figure 2. The triangular membership function is defined as (α, m, β) , where α, β are left and right side of boundaries and m is the model value. It is defined as follows:

$$Y = \text{Triangle}(x; \alpha, m, \beta) = \begin{cases} 0, & x \leq \alpha \\ \frac{x-\alpha}{m-\alpha}, & \alpha \leq x \leq m \\ \frac{\beta-x}{\beta-m}, & m \leq x \leq \beta \\ 0, & \beta \leq x \end{cases} \quad (1)$$

STEP 2:- FUZZY INFERENCE: In this step we apply the following fuzzy rules to determine the Effort equation.

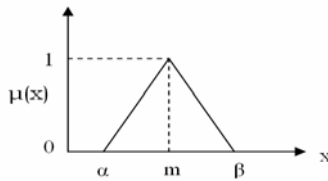


Fig. 2. Triangular Member Function

- a. If input is only size then apply the equation

$$\text{Effort} = a * (\text{size})^b \quad (2)$$

- b. Instead of having resources estimates as a function of one variable, resources estimates can depend on many different factors, giving rise to multivariable models hence create a non linearity. The cost factors are required software reliability,

Execution time constraint, Application experience, Software tools...etc. The products of all these factors are called Methodology. If input in size and Methodology (ME) then

$$\text{Effort} = a * (\text{size})^b + c * (\text{ME}) \tag{3}$$

- c. As per the above two models size and effort are directly proportional. But such a condition is not always satisfied giving rise to eccentric inputs. This can be accounted for by introducing a biasing factor (d). If input is size, methodology with bias then

$$\text{Effort} = a * (\text{size})^b + c * (\text{ME}) + d \tag{4}$$

STEP 3:- PARAMETER TUNING: The parameters “a, b, c, d” of the most model equations are termed by using Particle Swarm Optimization with inertia weight. The Update of velocity and positions of Parameter “a” is

$$V_{ai}^{k+1} = w * V_{ai}^k + c_1 * \text{rand}()_1 * (Pbest - S_{ai}^k) + c_2 * \text{rand}()_2 * (Gbest - S_{ai}^k) \tag{5}$$

$$S_{ai}^{k+1} = S_{ai}^k + V_{ai}^{k+1} \tag{6}$$

Similarly for b, c and d.

Step 4:- DEFUZZIFICATION: Finally fuzzy values are translates into actual output by using weighted average method by using Optimal Parameters obtained in Step2, Effort equation obtained in step3.

$$\text{Output} = \sum w_i * \mu_i / \sum w_i \tag{7}$$

4 Model Analysis

4.1 Implementation

We have implemented the above methodology for tuning parameters a, b, c and d in “C” language. For the parameter’ a ‘the velocities and positions of the particles are updated by applying the (5) & (6) equations with w=0.5, c1=c2=2.0. And similarly for the parameters b, c and d.

4.2 Triangular Membership Function

Fuzziness of TFN (α, m, β) is where m= model value, α = left boundary, β =right boundary. Fuzziness of TFN ($F = \beta - \alpha / 2m, 0 < f < 1$), the higher the value of fuzziness, the more fuzzy in TFN. The value of fuzziness to be taken depends upon the confidence of the estimator. A confident estimator can take smaller values of f. Let (m, 0) divides internally, the base of the triangular in ratio k=1, where k in the real positive number. If we consider F=0.1 and k=1 then $\alpha=0.9m$ and $\beta=1.1m$.

4.3 Weighted Average Defuzzification

Defuzzification is a mapping process from fuzzy logic control action area to a non fuzzy (crisp) control action area. It is done through weighted average method, by using Equation (7). The Values of a, b, c, d are obtained by using PSO with inertia

weight implemented in C language are as, $a = 2.646251$ and $b = 0.857612$ for Model 1, $a=2.771722$, $b=0.84795$ and $c = -0.007171$ for Model 2 and $a=3.131606$, $b=0.820175$, $c = 0.045208$ and $d=-2.020790$ for Model 3.

4.4 Performance Measures

We consider three performance criterions [13], Variance accounted – For (VAF), Mean Absolute Relative Error (MARE), Variance Absolute Relative Error (VARE).

5 Model Experiment

For the study of these models we have taken data of 10 NASA [10]. The following table1 shows estimated effort of our proposed model and Bailey –Basili, Alaa F. Sheta and Harish estimation models, shows using fuzzy based PSO a better software effort estimation is achieved.

Table 1. Estimated efforts of Various Models

Size	Meas ured effort	METH ODOL OGY	Bailey – Basili Estimate	Alaa F. Sheta G.E. Model Estimate	Alaa F. Sheta Model 2 Estimate	Harish model1	Harish model2	MODEL I	MODEL II	MODEL III
2.1	5	28	7.226	8.44	11.271	6.357	4.257	5.18221	5.18613	5.20013
3.1	7	26	8.212	11.22	14.457	8.664	7.664	7.23726	7.30845	7.35087
4.2	9	19	9.357	14.01	19.976	11.03	13.88	9.39037	9.55991	9.35261
12.5	23.9	27	19.16	31.098	31.686	26.252	24.702	23.92765	24.25429	24.91985
46.5	79	19	68.243	81.257	85.007	74.602	77.452	73.8252	74.34293	74.38504
54.5	90.8	20	80.929	91.257	94.977	84.638	86.938	84.59242	85.06759	84.93563
67.5	98.4	29	102.175	106.707	107.254	100.329	97.679	101.6272	101.9511	101.8466
78.6	98.7	35	120.848	119.27	118.03	113.237	107.288	115.8015	115.9854	115.7576
90.2	115.8	30	140.82	131.898	134.011	126.334	123.134	130.3124	130.4128	129.4197
100.8	138.3	34	159.434	143.0604	144.448	138.001	132.601	143.3405	143.2897	142.0119

6 Performance Analysis

It can be seen that Fuzzy based PSO models out perform the Bailey –Basili, Alaa F. Sheta and Harish Estimation models. The computed MARE, VAF and VARE for all models are given in Table 2.

Table 2. Performance Analysis of various models

Model	VAF (%)	MARE (%)	VARE (%)
Bailey –Basili Estimate	93.147	17.325	1.21
Alaa F. Sheta G.E. Model I Estimate	98.41	26.488	6.079
Alaa F. Sheta Model II Estimate	98.929	44.745	23.804
Harish model1	98.5	12.17	80.859
Harish model2	99.15	10.803	2.25
MODEL -I	98.92	6.1657	0.26
MODEL-II	98.545	6.539	0.237
MODEL-III	98.656	6.4731	0.209

7 Conclusion

Accurate software development cost estimation is very important in the budgeting, project planning and control, trade off and risk analysis of effective project management. The uncertainty in the input sizes will be reduced by using fuzzy logic. The parameters of the cost model are lined by using PSO with inertia weight will give the optimal result. The developed models are tested in NASA software project and proved to be best on the basis of VARE, MARE and VAF.

References

1. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Reading (1989)
2. Deb, K.: Multi-Objective Optimization Using Evolutionary Algorithms. John Wiley and Sons, Chichester (2002)
3. Sheta, A.F.: Estimation of the COCOMO Model Parameters Using Genetic Algorithms for NASA Software Projects. *Journal of Computer Science* 2(2), 118–123 (2006); ISSN 1549-3636
4. Sheta, A., Rine, D., Ayesh, A.: Development of Software Effort and Schedule Estimation Models Using Soft Computing Techniques. In: 2008 IEEE Congress on Evolutionary Computation, CEC 2008 (2008); 978-1-4244-1823-7/
5. Gonsalves, T., Ito, A., Kawabata, R., Itoh, K.: Swarm Intelligence in the Optimization of Software Development Project Schedule. IEEE, Los Alamitos (2008); 0730-3157/08
6. Pahariya, J.S., Ravi, V., Carr, M.: Software Cost Estimation using Computational Intelligence Techniques. 2009 World Congress on Nature & Biologically Inspired Computing (2009)
7. Attarzadeh, I., Ow, S.H.: Soft Computing Approach for Software Cost Estimation. *International Journal of Software Engineering (IJSE)* 3(1) (January 2010)
8. Huang, X., Ho, D., Ren, J., Capretz, L.F.: Improving the COCOMO model using a neuro-fuzzy approach. Elsevier, Amsterdam (2005), doi:10.1016/j.asoc.2005.06.007
9. Sheta, A., Rine, D., Ayesh, A.: Development of Software Effort and Schedule Estimation Models Using Soft Computing Techniques. IEEE, Los Alamitos (2008) 978-1-4244-1823-7/08
10. Bailey, J.w., Basili, v.R.: A meta model for software development resource expenditures. In: Fifth International conference on software Engineering, pp. 107–129. IEEE, Los Alamitos (1981), CH-1627-9/81/0000/0107500.75@ 1981
11. Ying, H.: The Takagi-Sugeno fuzzy controllers using the simplified linear control rules are nonlinear variable gain controllers. *Automatica* 34(2), 157–167 (1998)
12. Pang, W., Wang, K.-P., Zhou, C.-G., Dong, L.-J.: Fuzzy Discrete Particle Swarm Optimization for Solving Traveling Salesman Problem. In: Proceedings of the Fourth International Conference on Computer and Information Technology (CIT 2004) (2004) 0-7695-2216-5/04
13. Hari, C.V.M.K., Prasad Reddy, P.V.G.D., Jagadeesh, M.: Interval Type 2 Fuzzy Logic for Software Cost Estimation Using Takagi-Sugeno Fuzzy Controller. In: Proceedings of 2010 International Conference on Advances in Communication, Network, and Computing. IEEE, Los Alamitos (2010), doi:10.1109/CNC.2010.14, 978-0-7695-4209-6/10

SLV: Sweep Line Voronoi Ad Hoc Routing Algorithm

E. Rama Krishna¹, A. Venkat Reddy², N. Rambabu³, and G. Rajesh Kumar⁴

¹ VITS, Karimnagar, Andhra Pradesh, India
erroju.ramakrishna@gmail.com

² JITS, Karimnagar, Andhra Pradesh, India
venkat_scce@yahoo.co.in

³ SCCE, Karimnagar, Andhra Pradesh, India
neelabhi_2004@rediffmail.com

⁴ VITS, Karimnagar, Andhra Pradesh, India
gunupudirajesh@gmail.com

Abstract. Mobile Ad hoc routing is challenging issue due to dynamic and distributed nature of mobile nodes required frequent observed updations of node information to route the packets appropriately, there are several proactive and reactive routing methods and geographical hierarchical routing techniques to route the packets from source to destination effectively. Based on the characteristics of Mobile Ad Hoc Network (MANET) routing protocols the routing techniques should address the distributed, dynamic topology, periodic updates and physical security. We proposed a novel mechanism concentrating to minimize the existing DSDV routing limitations; it is originated from computational geometric concepts. Fortune's Algorithm (Sweep line method) which address the characteristic complexities of MANET routing and representing the route information in binary search method which minimizes the searching for optimal route.

Keywords: MANET, DSDV, routing, Fortune's Algorithm, Geographical hierarchical routing, Sweepline method.

1 Introduction

Mobile Ad hoc networks (MANETs) are those that instinctively form without the need of an infrastructure or centralized controller this type of peer-to-peer system infers that each node in the network can act as a data endpoint or intermediate repeater. A node in a wireless ad hoc network is connected with unrestricted nomadicity (or mobility), the topology of the network is dynamic. The random walk mobility model which is a stochastic process that models random continuous motion a mobile node moves from its current location with a randomly selected speed in a randomly selected direction. Routing evaluate packet delivery ratio, packet drop, end to end delay etc and determining performance parameters of any ad hoc communication network.

Computational geometry is the algorithmic study of geometric problems. Its emergence coincided with application areas such as computer graphics, computer-aided design/manufacturing, and scientific computing, which together provide much of the

motivation for geometric computing [1]. The basic idea of the design is to operate each Mobile Host as a specific router, which frequently advertises its view of the interconnection topology with other Mobile Hosts within the network and the idea is based on computational geometry, a sweepline algorithm or plane sweep algorithm is a type of algorithm that uses a conceptual sweep line or sweep surface to solve various problems in Euclidean space, a set of points satisfying certain relationships, expressible in terms of distance and angle. It is one of the key techniques in computational geometry.

In order to make this complete mathematically precise it must clearly define the notions of distance, angle, translation, and rotation is to define the Euclidean plane as a two-dimensional real vector space equipped with an inner product [2]

- the vectors in the vector space maps to the points of the Euclidean plane,
- the addition operation in the vector space maps to conversion, and
- the inner product implies notions of angle and distance, which can be used to define rotation

The design of network protocols for Mobile Ad hoc Networks (MANETs) is a complex issue. These networks need efficient *distributed* algorithms to determine network organization (connectivity), link scheduling, and routing. The shortest path (based on a given cost function) from a source to a destination in a static network is usually the optimal route, this idea is not easily extended to MANETs. Factors such as power expended, variable wireless link quality, propagation path loss, fading, multiuser interference, and topological changes, become relevant issues [3].

Any protocol must competently handle several inherent characteristics of MANETs:

- *Dynamic topologies, Power constrained operation*: Nodes are free to move randomly thus, the network topology which is typically multi hop may transform randomly and rapidly at erratic times, power conservation is *crucial* in mobile wireless systems since these networks typically operate off power-limited sources and may consist of both bidirectional and unidirectional links.
- *Bandwidth-constrained, variable capacity links*: Wireless links will continue to have significantly lower competence than their hardwired counterparts. In addition, the realized throughput of wireless communications after accounting for the effects of multiple access, fading, noise, and interference conditions, etc.
- *Physical security*: Mobile networks are more vulnerable to physical security threats such as eavesdropping and congestion attacks.

2 Dynamic Destination-Sequenced Distance-Vector Routing Protocol (DSDV)

The Destination-Sequenced Distance-Vector (DSDV) Routing Algorithm is based on the idea of the classical Bellman-Ford Routing Algorithm with certain improvements [4]. An n-node ad hoc network, maintains n rooted trees, one for each destination. The main role of the algorithm is to solve the routing loop problem

Advantages of DSDV

DSDV is an efficient protocol for route discovery. Whenever a route to a new destination is required, it already exists at the source. Hence, latency for route discovery is very low, DSDV also guarantees loop-free paths.

Disadvantages of DSDV

DSDV needs to send a lot of control messages. These messages are important for maintaining the network topology at each node. This may generate high volume of traffic for high-density and highly mobile networks. Special care should be taken to reduce the number of control messages.

In this paper a novel approach to minimize the limitations of the DSDV is proposed, where a systematic arrangement of mobile nodes performs periodic updates of the routing information about its clustered neighbor nodes with systematic arrangement of all possible routing paths from source node to destination node. This approach two phase process, in which Nearest neighbor search followed by Fortune's algorithm which generates Voronoi diagram from a set of points (nodes) in a plane.

3 Nearest Neighbor Search

Nearest neighbor search (NNS), also known as proximity search, similarity search or closest point search, is an optimization problem for finding closest points in metric spaces. The problem is: given a set S of points in a metric space M and a query point $q \in M$, find the closest point in S to q . In many cases, M is taken to be d -dimensional Euclidean space and distance is measured by Euclidean distance [5]. The Nearest Neighbor Search Algorithm has different solution like linear search, locality sensitive hashing, vector approximation file, Space partitioning methods.

3.1 Space Partitioning Method

Space-partitioning systems are often hierarchical, meaning that a space (or a region of space) is divided into several regions, and then the same space-partitioning system is recursively applied to each of the regions thus created. Most space-partitioning systems use planes (or, in higher dimensions, hyper planes) to divide space: points on one side of the plane form one region, and points on the other side form another. Points exactly on the plane are usually arbitrarily assigned to one or the other side.

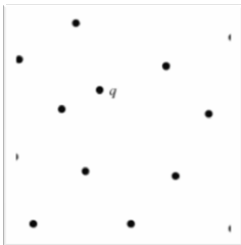


Fig. 1. A set S of n points in d dimensions query point q

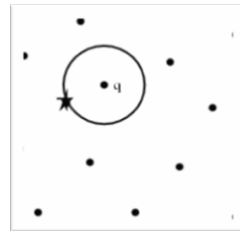


Fig. 2. Nearest neighbor graph on a set S of n points links each vertex

Recursively partitioning space using planes in this way produces a BSP tree, one of the most common forms of space partitioning [5].

4 Fortune’s Algorithm

The nearest neighbor queries as the Voronoi diagram is being computed, instead of storing it in an auxiliary data structure, using Fortune's algorithm which is a plane sweep algorithm for generating a Voronoi diagram from a set of points in a plane using $O(n \log n)$ time and $O(n)$ space. This algorithm is based on the following clever idea: rather than considering distances between the various sites, we will introduce a line that moves through the plane and use this line to facilitate a more efficient comparison of distances. We call this line the sweep line and think of it as uncovering the Voronoi diagram as it sweeps through the plane.

Let's first remember that if we are given a point p and a line l (not containing p), then the set of points whose distance to p equals its distance to l forms a parabola. We will use $P_{p,l}$ to denote this parabola. As before, it is helpful to think of the parabola as dividing the plane into two regions: one consisting of points closer to p and other consisting of points closer to l . This is an important point so, let's make it a little more clear. Consider a point q whose coordinates are $q=(q_x, q_y)$ and whose distance to p is denoted by $d(p,q)$. In what follows, the sweep line will always be a horizontal line; we will call its vertical coordinate l_y , so that the distance between q and l is $q_y - l_y$. The condition that q lies on the parabola $P_{p,l}$ is therefore $d(p,q) = q_y - l_y$.

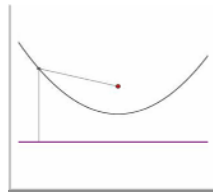


Fig. 3. Parabola dividing the plane into two regions

More generally

$$\begin{aligned}
 d(q,p) &< q_y - l_y \text{ if } q \text{ lies above } P_{p,l} \\
 d(q,p) &= q_y - l_y \text{ if } q \text{ lies on } P_{p,l} \\
 d(q,p) &> q_y - l_y \text{ if } q \text{ lies below } P_{p,l}
 \end{aligned}$$

now the horizontal sweep line l will move down through the plane. At any time, we will only consider the sites that lie above the sweep line and the parabolas $P_{p,l}$ define by those.

Fortune's algorithm: the beach line is the curve formed by the lowest parabolic arcs. That is, any vertical line will pass through several parabolas; the point at which the vertical line passes through the beach line is the lowest such point. Notice that each of the arcs that compose the beach line is associated to one of the sites above the sweep line.

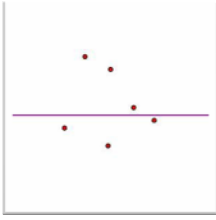


Fig. 4. Collection of sites and the position of the sweep line

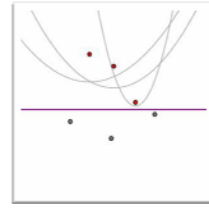


Fig. 5. The parabola

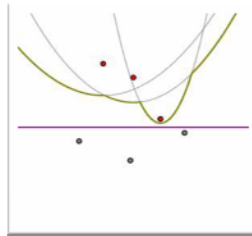


Fig. 6. Collection of sites and the position of the sweep line with parabola

The beach line is well suited for constructing the Voronoi diagram.

Let's now determine when the beach line passes through some arbitrary point q . Suppose that q is as close to site p_i as any other; that is, $d(q, p_i) \leq d(q, p_j)$ for all other sites p_j . The condition that q lies on $P_{p_i, l}$ may be expressed as

$$d(q, p_i) = q_y - l_y$$

Therefore

$$d(q, p_i) \leq d(q, p_j) = q_y - l_y$$

This means that when q appears on $P_{p_i, l}$, it cannot be above another parabola $P_{p_j, l}$, therefore, when q appears on the parabola $P_{p_i, l}$, it is on the beach line. When a point appears on the beach line, it is on a parabolic arc associated to its nearest site. The points on the beach line that lie on two parabolic arcs are called breakpoints. The breakpoints lie on the edges of the Voronoi diagram. This means that the breakpoints will sweep out the edges of the Voronoi diagram as the sweep line moves down the plane. So to construct the Voronoi diagram, we simply need to keep track of the breakpoints [6]. The Fortune's algorithm generates Voronoi diagram as an arrangement of nodes maintaining the information about the neighbor nodes these updates occur periodically which concludes the shortest and optimal path from source to destination is estimated.

5 Delaunay Triangles

The subdivision of the space determined by a set of distinct points so that each point has associated with it the region of the space nearer to that point than to any other is called **Dirichlet tessellation**. This process applied to a closed domain generates a set of convex distinct polygons called Voronoi regions which cover the entire domain. This definition can be extended to higher dimension where, for example in three dimensions, the Voronoi regions are convex polyhedrons [8]. If we connect all the pairs of points sharing a border of a Voronoi region we obtain a triangulation of the convex space containing those points. This triangulation is known as **Delaunay triangulation**.

The Delaunay triangle considered as a conceptual path to cover the nodes in particular region (Voronoi region), each node gets the information about its neighbor regions and the Delaunay property measure the existence of node.

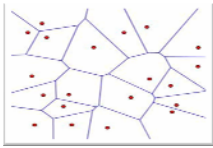


Fig. 7. Voronoi diagram (in dotted lines)

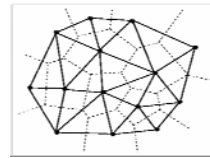


Fig. 8. Delaunay triangulation, on top of the Voronoi diagram

6 Routing Information Representation

Fortune’s algorithm or beach line method can be used to represented routing information as Beach Line Data Structure (T), derive a balanced binary search tree

- Internal nodes represent break points between two arcs
- Leaf nodes represent arcs, each arc in turn is represented by the site that has generated it

This tree structure constructed from the Sweepline (or Beach line) method of fortune’s algorithm.

The operations performed on the data structure representation are

Search: Given the current y-coordinate of the sweep line and a new site p_i , determine the arc of the beach line lies immediately above p_i . Let p_j denote the site that contributes this arc. Return a reference to this beach line entry.

Insert and split: Insert a new entry for p_i within a given arc p_j of the beach line (thus effectively replacing the single arc $\langle \dots, p_j, \dots \rangle$ with the triple $\langle \dots, p_j, p_i, p_j, \dots \rangle$. Return a reference to the newly added beach line entry (for future use).

Delete: Given a reference to an entry p_j on the beach line, delete this entry. This replaces a triple $\langle \dots, p_j, p_j, p_k, \dots \rangle$ with the pair $\langle \dots, p_i, p_k, \dots \rangle$

It is not difficult to modify a standard dictionary data structure to perform these operations in $O(\log n)$ time each [7].

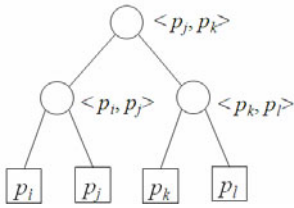


Fig. 9. Balanced binary search tree

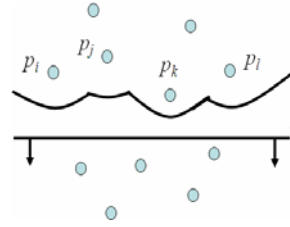


Fig. 10. Beach line method

7 Conclusion

Sweep Line Voronoi (SLV) method follows the geographic routing technique which exhibits better performance with highly dynamic and distributed natured mobile nodes which play a significant role in estimating the node location with periodic and uniform updating of the node information with a geometric relative attributes of the nodes. This enables the processing of connectivity-related parameters and routing information without flooding packets to identify the nodes to establish the route.

Sweep line or beach line is a systematic process which processes node information to establish the route in a step-by-step approach, which creates a Voronoi-based structure indicating the node range and Delaunay structure to possible route paths. These mathematical-related information can predict route alternates when the existing route is disturbed. The scan line creates a balanced binary search tree-based route information which enables the optimal processing of route discovery and node estimation.

References

1. Skiena, S.S.: The algorithm design manual
2. http://en.wikipedia.org/wiki/Euclidean_space
3. Subbarao, M.W.: Performance of Routing Protocols for Mobile Ad-Hoc Networks
4. Mahdipour, E., et al.: Performance Evaluation of DSDV Routing Protocol
5. http://en.wikipedia.org/wiki/Nearest_neighbor_search
6. Voronoi Diagrams and a Day at the Beach, David Austin Grand Valley State University
7. <http://ww.cs.umd.edu/class/spring2010/cmsc754/Lects/lect11.pdf>
8. Gorbach, M.: Image Stained Glass using Voronoi Diagrams

Hybrid Routing for Ad Hoc Wireless Networks

Ravilla Dilli¹, R.S. Murali Nath², and P. Chandra Shekar Reddy³

¹ Associate Professor, Dept. of ECE, St. Martin's Engineering College,
Secunderabad A.P., India
dilli.ravilla@gmail.com

² Professor, Dept. of CSE,
St. Martin's Engineering College
Secunderabad, A.P., India
muralinath.r.s@gmail.com

³ Professor Coordinator, Dept. of ECE,
JNT University,
Hyderabad, A.P, India
drpcsreddy@gmail.com

Abstract. The Zone Routing Protocol (ZRP) is a hybrid routing protocol that proactively maintains routes within a local region of the network (which we refer to as the routing zone). Here, we describe the motivation of ZRP and its architecture also the query control mechanisms, which are used to reduce the traffic amount in the route discovery procedure. In this paper, we address the issue of configuring the ZRP to provide the best performance for a particular network, at any time. Through NS2 simulation, we draw conclusions about the performance of the protocol.

Keywords: Zone Routing Protocol, Routing zone, Query control mechanisms.

1 The ZRP Architecture

As proactive routing uses excess bandwidth to maintain routing information, while reactive routing involves long route request delays. The Zone Routing Protocol (ZRP) aims to address the problems by combining the best properties of both approaches.

ZRP divides its network into different zones. A routing zone is defined for each node separately, and the zones of neighboring nodes overlap. ZRP refers to the locally proactive routing component as the Intra-zone Routing Protocol (IARP) [5] which maintains routing information for nodes that are within the routing zone of the node and the globally reactive routing component as the Inter-zone Routing Protocol (IERP) [2] which offer enhanced route discovery and route maintenance services based on local connectivity monitored by IARP. The protocol identifies multiple loop-free routes to the destination, increasing reliability and performance.

ZRP [3] uses a concept called *bordercasting* using the Bordercast Resolution Protocol (BRP) [6] is used in the ZRP to direct the route requests initiated by the global reactive IERP [2] to the peripheral nodes, thus removing redundant queries and maximizing efficiency. By employing *query control* mechanisms, route requests can be directed away from areas of the network that already have been covered. In order to

detect new neighbor nodes and link failures, the ZRP relies on a Neighbor Discovery Protocol (NDP) provided by the Media Access Control (MAC) layer. NDP transmits “HELLO” beacons at regular intervals. IERP uses the routing table of IARP to respond to route queries. IERP forwards queries with BRP.

1.1 Routing

A node that has a packet to send first checks whether the destination is within its local zone using information provided by IARP. In that case, the packet can be routed proactively. Reactive routing is used if the destination is outside the zone.

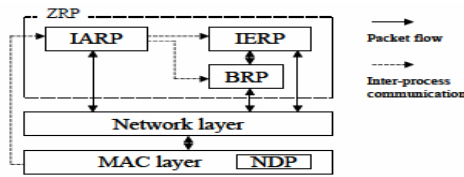


Fig. 1. ZRP architecture

The reactive routing process is divided into two phases: the *route request phase* and *route reply phase*. In the first, the source sends a route request packet to its peripheral nodes using BRP [6]. If the receiver of a route request packet knows the destination, it responds by sending a route reply back to the source. Otherwise, it continues the process by bordercasting the packet. The reply is sent by any node that can provide a route to the destination. To be able to send the reply back to the source node, routing information is accumulated when the request is sent through the network. In ZRP, the knowledge of the local topology can be used for route maintenance. Link failures and sub-optimal route segments within one zone can be bypassed. [4]

2 Query-Control Mechanisms

ZRP uses 3 types of query-control mechanisms: Query Detection, Early Termination, and Random Query- Processing Delay. To be able to prevent queries from reappearing in covered regions, the nodes must detect local query relaying activity. BRP provides two query detection methods: QD1 and QD2. QD1 is used to detect the nodes that relay the query. QD2 is used in single-channel networks to listen to the traffic by other

Table 1. Variable Simulation Parameters

Parameter	Symbol	Values	Defaults
Zone Radius (hops)	ρ	1-10	---
Node Density(neighbors/node)	δ	2-10	5
Rel. node speed(neighbors/s)	V	0.2-2.2	1.0
# of Nodes(nodes)	N	200-1500	400

nodes within the radio coverage (QD2). With Early Termination (ET), a node can prevent a route request from entering already covered regions. To reduce the probability of receiving the same request from several nodes, a Random Query-Processing Delay (RQPD) can be employed.

3 ZRP Performance Results

We expect that the amount of IARP [5] traffic should be proportional to the node density (δ) and the routing zone area (ρ^2). The IERP traffic/query decreases with the zone radius. For a given zone radius, we observe that the amount of received traffic/query increases with the zone density.

Keeping the zone radii and node density fixed, an increase in network population (N) is reflected by an increase in the network span. This has the effect of increasing the length of an IERP source route, thereby reducing the route’s reliability. Fig. 2(a) and (b) illustrates the effect of node density (δ) on the production of ZRP traffic.

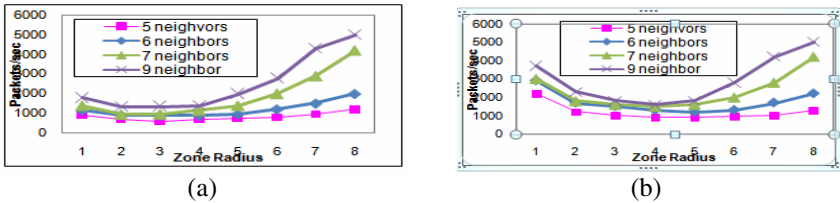


Fig. 2. ZRP traffic node ($N=1500$ nodes, $v=1.0$ neighbors/s) (a) $R_{route-usage} \ll R_{route-failure}$. (b) $R_{route-usage} \gg R_{route-failure}$.

We have seen that an increase in node density results in more IARP [5] route updates and more IERP packets per query. Fig. 2(b) demonstrates that the reduction in route query rate is more than offset by the increase in overlapping query packet transmission. In general, it appears that the ZRP traffic is an increasing function of node density. Also, the average node velocity (v) is a measure of the rate of network reconfiguration. Higher node velocities result in a linear increase in the IARP routing zone updates and IERP route failures.

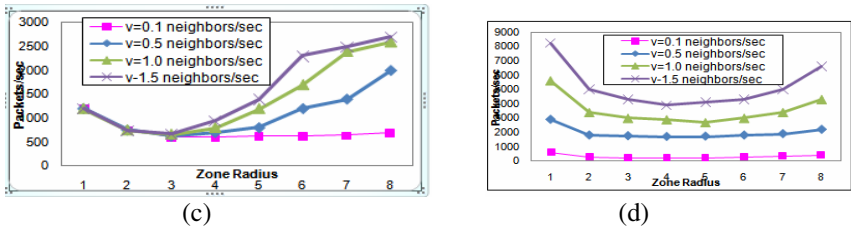


Fig. 2. (Continued): ZRP traffic node ($N=1500$ nodes, $d=5.0$ neighbors/s) (c) $R_{route-usage} \ll R_{route-failure}$. (d) $R_{route-usage} \gg R_{route-failure}$.

The overall ZRP traffic increases linearly with v , and ρ_{opt} remains constant. The node population influences the ZRP through its effect on the rate of received route queries. The addition of a new network node places additional load on the network through the extra route queries that it initiates. All other factors remaining constant, an increase in N results in an increase in network span. When the rate of route queries is driven by route failure, we have shown that larger network span reduces route reliability, further increasing the query load by all other nodes. Fig.2 (e) and (f) demonstrates this behavior.

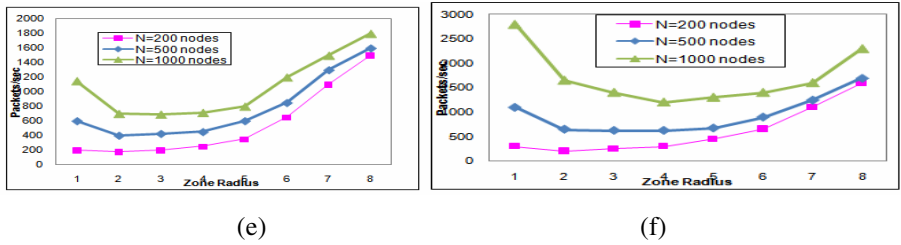


Fig. 2. (Continued) ZRP traffic node (5.0 neighbors, $v=1.0$ neighbors/s) (e) $R_{route-usage} \ll R_{route-failure}$. (f) $R_{route-usage} \gg R_{route-failure}$

As we expected, the amount of ZRP traffic increases with N and larger N favors a larger routing zone radius.

4 Conclusions

The ZRP combines two radically different methods of routing into one protocol. The tradeoff between the costs of proactive and reactive components of the ZRP determines the optimal zone radius for a given network configuration. The span of the network does not affect the amount of intrazone traffic, but the amount of reactive route query traffic increases with network span, thereby favoring larger routing zones. Relative node velocity has the potential to increase both intrazone updates and interzone route queries. We have seen that the total ZRP traffic increases with node density. We have proposed and evaluated zone radius estimation algorithm ‘min searching’, which attempts to minimize the amount of ZRP traffic based on direct measurements of the traffic. Our results demonstrate that the proposed route estimation techniques, applied in conjunction with a simple radius update protocol, allow the ZRP to perform more efficiently than traditional routing protocols without the need for centralized control or knowledge of the network operating conditions.

References

- [1] Park, V.D., Corson, M.S.: A highly adaptive distributed routing algorithm for mobile wireless networks. In: Proc. IEEE INFOCOM 1997, Kobe, Japan, pp. 1405–1413 (1997)
- [2] Haas, Z.J., Pearlman, M.R., Samar, P.: Interzone Routing Protocol (IERP). IETF Internet Draft, draft-ietf-manet-ierp-02.txt (July 2002)

- [3] Haas, Z.J., Pearlman, M.R., Samar, P.: The Zone Routing Protocol (ZRP) for Ad Hoc Networks. IETF Internet Draft, draft-ietf-manet-zone-zrp-04.txt (July 2002)
- [4] Beijar, N.: Zone Routing Protocol,
<http://www.tct.hut.fi/opetus/s38030/k02/Papers/08-Nicklas.pdf>
- [5] Haas, Z.J., Pearlman, M.R., Samar, P.: Intrazone Routing Protocol (IARP). IETF Internet Draft, draft-ietf-manet-iarp-02.txt (July 2002)
- [6] Haas, Z.J., Pearlman, M.R., Samar, P.: Bordercasting Resolution Protocol (BRP). IETF Internet Draft, draft-ietf-manet-brp-02.txt (July 2002)

Implementation of ARINC 429 16 Channel Transmitter Controller on FPGA

Debasis Mukherjee¹, Niti Kumar², Kalyan Singh³, Hemanta Mondal⁴,
and B.V.R. Reddy⁵

¹ SPSU, E&CE, Udaipur & GGSIPU, USIT, Delhi, India
debasismukherjee1@gmail.com

² GTBIT, E&CE, Delhi, India
niti221@rediffmail.com

³ AIT, RF & Microwave, New Delhi, India
kalyan_mail@yahoo.com

⁴ Freescale Semiconductor, Noida, India
hemanta.mandal@rediffmail.com

⁵ GGSIPU, USIT, Delhi, India
bvreddy64@rediffmail.com

Abstract. Multi-Channel transmitter Controller is designed on ARINC 429 Protocol. Its new architecture shows better performance and saves high percentage of total power i.e. static power and dynamic power. It is specially designed for providing a dedicated transmission link between processor and various equipments placed in aircraft by selecting the particular transmitter for transmitting the data to many equipments like Microwave Landing system, Tire Pressure Measuring Systems etc. Also, an efficient VLSI implementation of proposed system on PROASIC 3 plus ACTEL FPGA device is described.

Keywords: Total power, dynamic power and static power.

1 Introduction

Multi-Channel transmitter Controller is designed using the specifications of ARINC429. It supports up to 16 transmitter channels. It sets up unidirectional data transmission between micro-processor and ARINC429 supported equipments & has externally programmable control unit for selecting the transmission rate.

ARINC 429 supported equipments are installed in various commercial transport aircrafts. Many electronic equipments are placed in aircraft and each equipment has its unique identification number. Multi-channel receiver controller receives data from these equipments.

A VLSI implementation of the proposed architecture on FPGA device is described. The whole design was captured in VHDL language and was synthesized.

2 ARINC 429 Specifications

ARINC 429 specifications defines how the avionic equipments and systems should communicate with each other. These specifications includes electrical and data

characteristics. The transmitter is used to transmit information from the defined port to 16 receivers over a single twisted and shielded wire pairs. Data is transmitted at a bit rate of 12.5 Kbps for low speed or 100 Kbps for high speed. ARNIC429 supported electrical signal has three states, is shown in fig-1. ARINC 429 supports three states of a signal which are high state driven by +5 V, Null state driven by 0V and low state driven by -5V. Bipolar return to zero encoding scheme is used to solve the synchronization problem. Implementation of ARNIC429 specified signal is not possible through single pin because it has three states so desired signal is separated into two digital signals called Signal A and signal B as shown in the fig– 2. These signals are transmitted via two pins. The bit value of each state of signals A and B specifies the three states of desired signal. ARNIC429 supports 32 bit data word. The data word format consists of parity, label, sign/status matrix, data and source/destination identifier. MSB is the parity bit which is an odd parity. The label is used to specify the type of data. Source/destination identifier is used for multiple receivers to identify the receiver for which the data is destined.

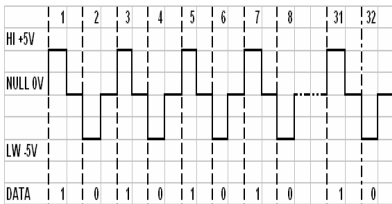


Fig. 1. ARINC Standard

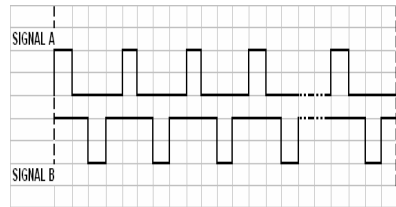


Fig. 2. Signal A and B

3 Architecture and Functional Description of Single Channel Transmitter

Single channel transmitter has two 16 bits latches .These latches are used to store the 16-16 bits data at positive and negative edge trigger of load signal. It uses signal Txsel to select the transmission rate. If Txsel is high then high transmission speed is passed and if it is low then low transmission speed is passed. This speed is given to the counter and modulator. Just as the 16 bits MSB of data is loaded, FF1 is set to enable the counter. This counter is of 6 bits but it counts only up to 36 counts where 32 counts are for transmitting the 32 bits and 4 counts are used to generate the four bit gap to distinguish between two words .The counter gives its count value to the modulator and based on this value of it check the bits of 32 bits data word and performs parallel to serial conversion. Its architecture is shown in fig.3.

4 Architecture and Functional Description of 16-Channel Transmitter Controller

16 channel transmitter controller operates on 16 MHz clock. It has four bit address bus to select the particular transmitter channel. It has a 16 bit control register and each

bit of which is assigned to each transmitter. The value of bits will decide at which transmission rate the transmitter will transmit the data serially. If bit value of control register is '0' then low frequencies such as 100 KHz and 12.5 kHz is selected and if the bit is '1' then high frequency such as 200 KHz and 100 KHz are selected. The decoder is used to select the particular channel for transmission of data: fig-4.

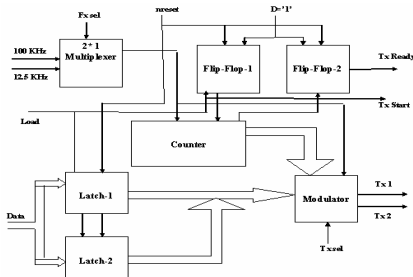


Fig. 3. Single Channel Transmitter

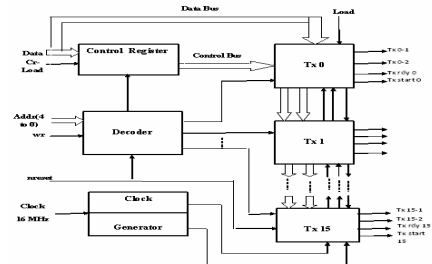


Fig. 4. 16 Channel Transmitter

5 Results and Conclusions

Family- ProASIC3, Die-A3P250, Package-208 PQFP is used.

Table 1. The comparison of Battery life and power consumption

Architecture	Transmitter Controller with 16 Control registers	Transmitter Controller with 1 Control Register
Battery Capacity	1000.000 mA*Hours	1000.000 mA*Hours
Battery Life	2.026 Hours	2.118 Hours
Total Power	1576.309 mW	1510.567 mW
Static Power	5.220 mW	5.220 mW
Dynamic Power	1571.089 mW	1505.347 mW

New architecture of ARNIC429 Transmitter and it's multi-channel transmitter controller is introduced which provides an easy interfacing with 16 bit micro-processor and controlled by micro-processor. Proposed architecture use less hardware of FPGA chip, consumes less power and increases battery life also.

References

1. Pschierer, C., Kasten, J., Schiefele, J., Lepori, H., Bruneaux, P., Bourdais, A., Andrae, R.: ARINC 424A A next generation navigation database specification. In: IEEE/AIAA 26th Digital Avionics Systems Conference, DASC 2007, pp. 2.B.6-1 – 2.B.6-8 (2007)
2. Arnic, <http://www.ARINC.com>

Segmentation of Image Using Watershed and Fast Level Set Methods

Minal M. Puranik and Shobha Krishnan

S.A.K.E.C, Mumbai

+91-9323106641

Minalpuranik@gmail.com

VESIT, Mumbai

25227638(404)

shobha.krishnan@hotmail.com

Abstract. Technology is proliferating. Many methods are used for medical imaging. The important methods used here are fast marching and fast level set in comparison with the watershed transform. Since watershed algorithm was applied to an image has over clusters in segmentation. Both methods are applied to segment the medical images. First, fast marching method is used to extract the rough contours. Then fast level set method is utilized to finely tune the initial boundary. Moreover, Traditional fast marching method was modified by the use of watershed transform. The method is feasible in medical imaging and deserves further research. In the future, we will integrate level set method with statistical shape analysis to make it applicable to more kinds of medical images and have better robustness to noise.

Categories and Subject Descriptors

I.4 Image Processing and Computer Vision.

General Terms

Design, Experimentation.

Keywords: Segmentation, Watershed transform, fast marching method and fast level set method.

1 Introduction

In computer vision, segmentation refers to the process of partitioning a digital image into multiple segments (sets of pixels) (Also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

1.1 Fast Marching Method

The *Fast Marching Method* imitates this process. Given the initial curve (shown in red), stand on the lowest spot (which would be any point on the curve), and build a little bit of the surface that corresponds to the front moving with the speed F . Repeat the process over and over, always standing on the lowest spot of the scaffold, and building that little bit of the surface. When this process ends, the entire surface has been built.

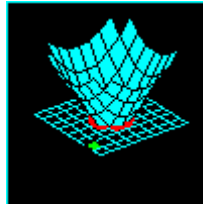


Fig. 1. Construction of stationary level set solution

Green squares show next level to be built

The speed from this method comes from figuring out in what order to build the scaffolding; fortunately, there are lots of fast sorting algorithms that can do this job quickly. Equation of the evolution of Φ , inside which our surface is embedded as the zero level set is then given by the following equation:

$$\Phi_t + F|\nabla\Phi| = 0 \quad (1)$$

The major advantages of using this method over other active contour strategies include the following [10]. First, the evolving level set function $\Phi(x, t)$ remains a function, but the propagating front $\Gamma(t)$ may change topology, break, merge and form sharp corners as Φ evolves. Second, the intrinsic geometric properties of the front may be easily determined from Φ . For example, at any point of the front, the normal vector is given by $n = \nabla\Phi$. Finally, the formulation is unchanged for propagating interfaces in three dimensions. One of the most popular level set algorithms is the so-called fast marching method. Now consider the special case of a surface moving with speed > 0 (the case where F is everywhere negative is also allowed). We then have a monotonically advancing front whose level set equation is of the following form:

$$|\nabla T|F = 1 \quad (2)$$

This simply says that the gradient of arrival time is inversely proportional to the speed of the surface. There are two ways of approximating the position of the moving surface: iteration towards the solution by numerically approximating the derivatives in Eq. (1) or explicit construction of the solution function T from Eq. (2). Fast marching method depends on the latter approach. Equation (2) is one form of the Eikonal equations. Sethian proved that it is equivalence to solve the following quadratic equation in order to get the arrival time T of the Eq. (2).

$$\left[\max(D_{ij}^{-x}T, 0)^2 + \min(D_{ij}^{+x}T, 0)^2 + \max(D_{ij}^{-y}T, 0)^2 + \min(D_{ij}^{+y}T, 0)^2 \right]^{1/2} = 1/F_{i,j} \tag{3}$$

Where $D+$ and $D-$ are backward and forward difference operators:

$$\begin{cases} D^{+x}T = \frac{T(x,y+1)-T(x,y)}{2} \& D^{-x}T = \frac{T(x,y)-T(x,y-1)}{2} \\ D^{+y}T = \frac{T(x+1,y)-T(x,y)}{2} \& D^{-y}T = \frac{T(x,y)-T(x-1,y)}{2} \end{cases} \tag{4}$$

The steps of the traditional fast marching method are as follows:

1. *Initializing step:*

- 1) *Alive points:* Let A be the set of all grid points $\{iA, jA\}$ which represents the initial curve. In our algorithm, Alive points are the seeded points users assign to. See Fig. 2;
- 2) *Narrowband points:* Let Narrowband points be the set of all grid point neighbors of A. In our Algorithm, those are the 4-nearest points of the seeded points. Set $T(x, y) = 1/F(x, y)$.
- 3) *Faraway points:* Let Faraway points be the set of all others grid points $\{x, y\}$. Set $T(x, y) = \text{TIME MAX}$, ;

1.2 Marching Forwards

- 1) Begin loop: Let $(imin, jmin)$ be the point in Narrowband with the smallest value for T;
- 2) Add the point $(imin, jmin)$ to A; remove it from Narrowband;
- 3) Tag as neighbors any points $(imin - 1, jmin)$, $(imin + 1, jmin)$, $(imin, jmin - 1)$, $(imin, jmin + 1)$ that is either in Narrowband or Faraway. If the neighbor is in Faraway, remove it from that list and add it to the set Narrowband;
- 4) Recomputed the values of T at all neighbors according to equation 3, selecting the largest possible solution to the quadratic equation;
- 5) Return to top of Loop.

1.3 Level Set Method

In implementing the traditional level set methods, it is numerically necessary to keep the evolving level set function close to a signed distance function. *Reinitialization*, a technique for periodically re-initializing the level set function to a signed distance function during the evolution, has been extensively used as a numerical remedy for maintaining stable curve evolution and ensuring usable results. However, as pointed out by Gomes, re-initializing the level set function is obviously a disagreement

between the theory of the level set method and its implementation. Moreover, many proposed re-initialization schemes have an undesirable side.

In this paper, we present a new variational formulation that forces the level set function to be close to a signed distance function, and therefore completely eliminates the need of the costly re-initialization procedure. Our variational energy functional consists of an internal energy term and an external energy term, respectively. The internal energy term penalizes the deviation of the level set function from a signed distance function, whereas the external energy term drives the motion of the zero level set to the desired Formulation has three main advantages over the traditional level set formulations. First, a significantly larger time step can be used for numerically solving the evolution PDE, and therefore speeds up the curve evolution. Second, the level set function could be initialized as functions that are computationally more efficient to generate than the signed distance function. Third, the proposed level set evolution can be implemented using simple finite difference scheme, instead of complex upwind scheme as in traditional level set formulations. The proposed algorithm has been applied to both simulated and real images with promising results. In particular it appears to perform robustly in the presence of weak boundaries.

1.4 Drawbacks Associated with Re-initialization

Re-initialization has been extensively used as a numerical remedy in traditional level set methods. The standard re-initialization method is to solve the following *reinitialization equation*

$$\frac{\partial \phi}{\partial t} = \text{sign}(\phi_0)(1 - |\nabla \phi|)$$

where ϕ_0 is the function to be re-initialized, and $\text{sign}(\phi)$ is the sign function. There has been copious literature on re-initialization methods, and most of them are the variants of the above PDE-based method. Unfortunately, if ϕ_0 is not smooth or ϕ_0 is much steeper on one side of the interface than the other, the zero level set of the resulting function ϕ can be moved incorrectly from that of the original function. Moreover, when the level set function is far away from a signed distance function, these methods may not be able to re-initialize the level set function to a signed distance function. In practice, the evolving level set function can deviate greatly from its value as signed distance in a small number of iteration steps, especially when the time step is not chosen small enough. So far, re-initialization has been extensively used as a numerical remedy for maintaining stable curve evolution and ensuring desirable results. From the practical viewpoints, the re-initialization process can be quite complicated, expensive, and have subtle side effects. Moreover, most of the level set methods are fraught with their own problems, such as when and how to re-initialize the level set function to a signed distance function. There is no simple answer that applies generally to date. The variational level set formulation proposed in this paper can be easily implemented by simple finite difference scheme, without the need of re-initialization.

2 Variational Fast Level Set Formulation of Curve Evolution without Re-initialization

2.1 General Variational Fast Level Set Formulation with Penalizing Energy

As discussed before, it is crucial to keep the evolving level set function as an approximate signed distance function during the evolution, especially in a neighborhood around the zero level set. It is well known that a signed distance function must satisfy a desirable property of $|\nabla\phi| = 1$. Conversely, any function ϕ satisfying $|\nabla\phi| = 1$ is the signed distance function plus a constant. Naturally, we propose the following integral

$$P(\phi) = \int_{\Omega} \frac{1}{2} (|\nabla\phi| - 1)^2 dx dy$$

as a metric to characterize how close a function ϕ is to a signed distance function in $\Omega \subset \mathbb{R}^2$. This metric will play a key role in our variational level set formulation. With the above defined functional $P(\phi)$, we propose the following variational formulation $E(\phi) = \mu P(\phi) + Em(\phi)$.

where $\mu > 0$ is a parameter controlling the effect of penalizing the deviation of ϕ from a signed distance function, and $Em(\phi)$ is a certain energy that would drive the motion of the zero level curve of ϕ . For a particular functional $E(\phi)$ defined explicitly in terms of ϕ , the Gateaux derivative can be computed and expressed in terms of the function ϕ and its derivatives.

2.2 Variational Level Set Formulation of Active Contours without Re-initialization

In image segmentation, active contours are dynamic curves that moves toward the object boundaries. To achieve this goal, we explicitly define an external energy that can move the zero level curve toward the object boundaries. Let I be an image, and g be the *edge indicator function* defined by

$$g = \frac{1}{1 + |\nabla G_{\sigma} * I|^2},$$

Where G_{σ} is the Gaussian kernel with standard deviation σ . We define an external energy for a function $\phi(x, y)$ as below

$$\mathcal{E}_{g,\lambda,\nu}(\phi) = \lambda \mathcal{L}_g(\phi) + \nu \mathcal{A}_g(\phi)$$

where $\lambda > 0$ and ν are constants, the Heaviside function. Now, we define the following total energy functional

$$\mathcal{E}(\phi) = \mu P(\phi) + \mathcal{E}_{g,\lambda,\nu}(\phi)$$

The external energy $\mathcal{E}_{g,\lambda,\nu}$ drives the zero level set toward the object boundaries, while the internal energy $\mu P(\phi)$ penalizes the deviation of ϕ from a signed distance function during its evolution. The gradient flow is the evolution equation of the level

set function in the proposed method. The second and the third term in the right hand side of correspond to the gradient flows of the energy functional, respectively, and are responsible of driving the zero level curves towards the object boundaries. To explain the effect of the first term, which is associated to the internal energy $\mu P(\varphi)$.

2.3 Selection of Time Step

In implementing the proposed level set method, the time step τ can be chosen significantly larger than the time step used in the traditional level set methods. We have tried a large range of the time step τ in our experiments, from 0.1 to 100.0. For example, we have used $\tau = 50.0$ and $\mu = 0.004$ for the image in and the curve evolution only takes 40 iterations, while the curve converge to the object boundary precisely.

3 Watershed Transform

Watershed transformation is a powerful tool for image segmentation. In this paper, the different morphological tool used in segmentation is reviewed together with an abundant illustration of the methodology through examples of image segmentation coming from various areas of image analysis. The gradient image is often used in the watershed transformation, because the main criterion of the segmentation is the homogeneity of the grey values of the object present in the image. But, when other criteria are relevant, other functions can be used. In particular, when the segmentation is based on the shape of the objects, the distance functions is very helpful. The watershed transform finds the catchment basins and ridge lines in such a grayscale image.

4 Experimental Details (Results)

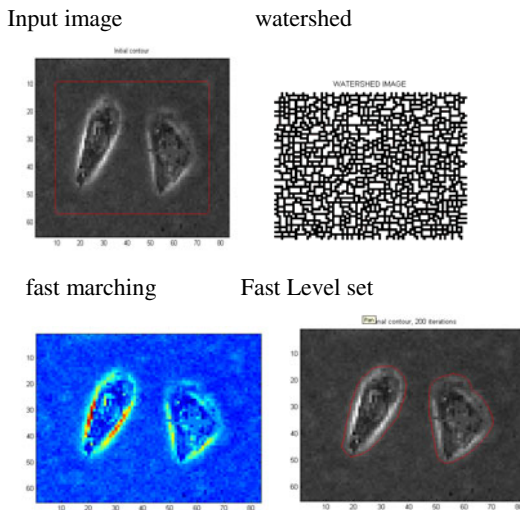


Fig. 2. Result for microscopic cell image

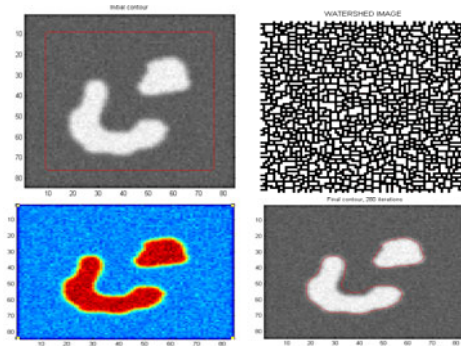


Fig. 3. Result for microscopic cell image

5 Conclusion

In this paper, we present a new variational level set formulation that completely eliminates the need of the reinitialization, results demonstrate desirable performance of our method in extracting weak object boundaries, which is very difficult for the traditional level set method, watershed and fast marching method.

References

1. Qiu, P.: Some recent developments on edge detection and image reconstruction based on local smoothing and nonparametric regression. In Book Recent Research Developments in Pattern Recognition 1, 41–49 (2000)
2. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: a survey. *Medical Image Analysis* 1(2), 91–108 (1996)

Tree Structured, Multi-hop Time Synchronization Approach in Wireless Sensor Networks

Surendra Rahamatkar¹, Ajay Agarwal², Praveen Sen¹, and Arun Yadav³

¹ Dept. of Computer Science, Nagpur Institute of Technology, Nagpur, India
rahamatkar_s@rediffmail.com

² Dept. of MCA, Krishna Institute of Engg. & Technology, Ghaziabad, India

³ Dept. of Computer Sc., BMAS Engineering College, Agra,(U.P.), India

Abstract. Time synchronization for wireless sensor networks (WSNs) has been studied in recent years as a fundamental and significant research issue. Time synchronization in a WSN is a critical for accurate time stamping of events and fine-tuned coordination among the sensor nodes to reduce power consumption. This paper proposes a reference based, tree structured multi-hop time synchronization service for WSNs in which sensor nodes synchronize by collecting reference points with reference to Root node of the logically constructed tree structure of the network. Proposed approach is lightweight as the number of required broadcasting messages is restricted to one broadcasting domain.

Keywords: Ad Hoc Tree Structure, Clock synchronization, Wireless sensor networks, Hierarchical sensor network.

1 Introduction

Time synchronization for wireless sensor networks (WSNs) has been studied in recent years as a fundamental and significant research issue. Many applications based on these WSNs assume local clocks at each sensor node that need to be synchronized to a common notion of time. Time synchronization, which aims at creating a common time scale in the network, is a necessity for many applications.

Two of the most prominent examples of existing time synchronization protocols developed for the WSN domain are the Reference Broadcast Synchronization (RBS) algorithm [4] and the Timing-sync Protocol for Sensor Networks (TPSN) [5]. Tiny-SeRSync [7] protocol works with the ad hoc deployments of sensor networks. This protocol proposed two asynchronous phases: Phase I –secure single-hop pair wise synchronization, and Phase II–secure andresilient global synchronization to achieve global time synchronization in a sensor network.

2 Tree Structured Multi-hop Time Synchronization Approach

In this Section we proposed Tree Structured Referencing multi-hop Time Synchronization scheme, which is based on the protocol, proposed by [1], [2] that aims to

minimize the complexity of the synchronization. Thus the needed synchronization accuracy is assumed to be given as a constraint, and the target is to devise a synchronization algorithm with minimal complexity to achieve given precision. The proposed scheme works on two phases. First phase used to construct an ad hoc tree structure and second phase used to synchronize the local clocks of sensor nodes followed by network evaluation phase.

The goal of the TSRT is to achieve a network wide synchronization of the local clocks of the participating nodes. We assume that each node has a local clock exhibiting the typical timing errors of crystals and can communicate over an unreliable but error corrected wireless link to its neighbors as mentioned in [6]. This scheme synchronizes the time of a sender to possibly multiple receivers utilizing a single radio message time-stamped at both the sender and the receiver sides. Linear regression is used in this scheme to compensate for clock drift as suggested in [4].

2.1 Tree Structured Multi-hop Time Synchronization Phase

The proposed approach synchronizes the time of a sender with respect to multiple receivers by utilizing a single synchronization message time-stamped at both sides (sender and receiver). We assume that the root of the network is dynamically decided through the election algorithm, which maintains the global time of the network and all other nodes synchronize their clocks with reference to the root node. A logical tree structure from a designated source point (root) will be formed using the algorithm suggested at [2]. The nodes form an adhoc tree structure to transfer the global time from the root to all the nodes.

2.2.1 Time Synchronization

Practically the network radius of WSN network is generally greater than one hop. In the case of network-wide synchronization, the proposed multi-hop approach can be applied; the detailed description is described in this section. We assume that every node contains a unique identification.

Nodes in multi-hop synchronization, utilize *reference points* to achieve synchronization. A reference point contains a pair of global and local time stamps where both of them refer to the same time instant. Periodically by sending and receiving the synchronization message, reference points are generated. The transmissions of synchronization message are either initiated by the *root node* of the logically constructed adhoc tree of the network or any *synchronized node* in the network. The root node of the network is dynamically decided through the election algorithm and with reference to the root node; the whole network is being synchronized.

A node within the broadcast range of the root can collect reference points directly from the root. Nodes outside the broadcast range of the root can gather reference points indirectly through other synchronized nodes, are synchronized with reference to the root. When a node collects enough consistent reference points, by estimating the clock offset and skews of its own local clock and becomes synchronized. The newly synchronized node can then broadcast synchronization messages to other nodes in the network.

The proposed scheme is explained in more detail as follows: Periodically a Root Node namely the reference node broadcasts the Synchronization message to all other

nodes of its transmission range. After receiving the consistent number of Synchronization messages, the node (say node-1 and node-2 in fig. 1(A)) generates the reference points and by estimating the clock offset and skews of own local clock can become synchronized. The nodes from beyond the range of Root Node may synchronize by collecting the reference points from the nearest synchronized node from broadcast range. As per the fig. 3(B), Nodes 5,6,7,8 and 9 may collect the reference points from the nodes 3 and 4. Nodes 3 and 4 are synchronized with reference to Root Node and periodically broadcast the Synchronization message to all other nodes of their transmission range. The nodes 5 & 6 may be synchronized with reference to the node 3; similarly nodes 7 and 9 may be synchronized with reference to synchronized node 4.

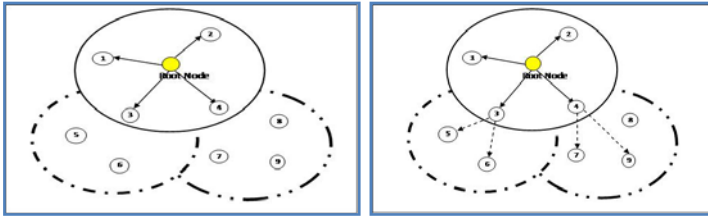


Fig. 1. (A) Reference node broadcasts & neighbours are synchronized

(B) Other layer nodes are synchronized

3 Comparison and Conclusion

The performance comparison of proposed scheme and TPSN in terms of the average number of message exchanges M with respect to the number of beacons N when network wide sync error probability P_s is assumed as 0.01% and 1% respectively. This comparison is based on the linear network model where the depth of the network $B = 5$, $\epsilon_{\max} = 10ms$, $\sigma_{\epsilon_0} = 16.67\mu$, $d = 10ms$, $t = 400ms$, and $\sigma_{\epsilon_s} = 1.58\mu$ have assumed.

It can be observed that proposed scheme requires a less number of timing messages than TPSN when there multiple numbers of beacon transmissions are required. Moreover, the gap of the average number of required timing messages between proposed scheme and TPSN significantly increases as N increases, and thus proposed scheme is by far more efficient than TPSN for large value of N . It can be also seen that a few number of beacons is enough to minimize M . Besides, as expected, a larger number of beacons required to meet a more strict constraint of the network-wide error probability P_s . In practice, a lower number of N is highly preferable since it is proportional to the synchronization time, i.e., a lower N induces better latency performance, although, it may not be optimal in terms of energy consumption.

While the proposed approach is especially useful in WSN which are typically, extremely constrained on the available computational power, bandwidth and have some of the most exotic needs for high precision synchronization [3]. The proposed synchronization approach was designed to switch between TPSN and RBS. These two algorithms allow all the sensors in a network to synchronize themselves within a few microseconds of each other, while at the same time using the least amount of

resources possible. In this work two varieties of the algorithm are presented and their performance is verified theoretically and compared with existing protocols. The comparison with RBS and TPSN shows that the proposed synchronization approach is lightweight since the number of required broadcasting messages is constant in one broadcasting domain.

References

1. Rahamatkar, S., Agarwal, A.: An Approach towards Lightweight, Reference Based, Tree Structured Time Synchronization Scheme in WSN. In: COSIT, Part I. CCIS, vol. 131, pp. 189–198. Springer, Berlin (2011)
2. Rahamatkar, S., Agarwal, A., Sharma, V.: Tree Structured Time Synchronization Protocol in Wireless Sensor Network. *J. Ubi. Comp. & Comm.* 4, 712–717 (2009)
3. Rahamatkar, S., Agarwal, A., Kumar, N.: Analysis and Comparative Study of Clock Synchronization Schemes in Wireless Sensor Networks. *Int. J. Comp. Sc. & Engg.* 2(3), 523–528 (2010)
4. Elson, J.E., Girod, L., Estrin, D.: Fine-Grained Network Time Synchronization using Reference Broadcasts. In: 5th Symp. on Operating Systems Design and Implementation, pp. 147–163 (2002)
5. Ganeriwal, S., Kumar, R., Srivastava, M.B.: Timing-Sync Protocol for Sensor Networks. In: First ACM Conference on Embedded Networked Sensor System (SenSys), pp. 138–149 (2003)
6. Dai, H., Han, R.: TSync: a lightweight bidirectional time synchronization service for wireless sensor networks. *SIGMOBILE Mob. Comp. Comm. Rev.* 8(1), 125–139 (2004)
7. Sun, K., Ning, P., Wang, C.: TinySeRSync: Secure and Resilient time synchronization in wireless sensor networks. In: 13 ACM Conf. on Comp. Comm. Security, pp. 264–277 (2006)

A New Markov Chain Based Cost Evaluation Metric for Routing in MANETs

Abhinav Tiwari, Nisha Wadhawan, and Neeraj Kumar

School of Computer Science and Engineering
Shri Mata Vaishno Devi University, Katra (J&K), India
abhinavtiwari.1003@gmail.com, nishadhamija74@yahoo.com,
nehra04@yahoo.co.in

Abstract. In recent time, routing remains a major concern in Mobile adhoc networks (MANETs) due to the mobility of nodes. In particular, the link stability remains a major concern in existing protocols. In this paper, we propose a Markov Chain based Cost Evaluation Metric (MCCE) which is an extension of existing WCETT metric with an enhancement of link stability. The proposed routing metric is capable of providing better link stability and higher packet delivery fraction than the other proposed metrics in this category.

Keywords: Cost Evaluation function (CE), MCCE (Markov Chain Cost Evaluation), Mobile Adhoc Networks, ETX, ETT, WCETT.

1 Introduction

A mobile Adhoc Network (MANET) is that network which does not have any centralized control. Each node in MANET can send data to any node within its transmission range through a series of intermediate nodes, i.e., that every node can act as a sender or a receiver or a router. These networks are used in areas where a fast and efficient network deployment is required such as in military operations and in floods & earthquakes affected areas where the deployment of any wired or wireless network quickly is not feasible [1,2]. Since in MANETs all the nodes are mobile, network topology keeps on changing with respect to mobility of nodes in the network [3]. It is very important to route data to the destination through the most optimal route in order to improve the performance of the network [4] [5].

Routing in MANETs is a very active area of research. Lot of efforts have been made to improve the process of transferring data on the basis of various routing metrics which are used to select the best route for the data transfer between any two nodes. Various metrics have been proposed in order to find high quality paths taking into consideration topology changes which can be chosen depending upon the type of environment which is required to be built. [1]

The remainder of this paper is organized as follows. Section 2 discusses the related work, Section 3 describes the system architecture and proposed approach, Section 4 describes the simulation results and comparison, and finally section 5 concludes the article.

2 Related Works

A number of MANETs routing protocols have been proposed [6, 7, 8] which are based on the minimum hop-count as their routing metric. But hop-count does not provide a path with maximum throughput in many scenarios [9]. To remove the problems associated with hop-count metric, a new metric called Expected Transmission Count (ETX) was proposed [9]. WCETT [10] was an enhancement over the ETT metric which contained all the features of ETT metric and it also included channel diversity for the first time. WCETT along a path with n hops and k different channels is given by the following formula

$$WCETT = (1 - \beta) \times \sum_{i=1}^n ETT_i + \beta \times \max_{1 \leq j \leq k} X_j, \text{ Where } X_j = \sum_{\text{hop } i \text{ is on channel } j} ETT_i$$

For $1 \leq j \leq k$ and $0 \leq \beta \leq 1$. Finally, ETT is estimated as

$$ETT = ETX \times \frac{S}{B}$$

Where S is the packet size and B is the bandwidth of the link. This metric did not account for link stability in its calculation. Therefore, we propose an improvement over WCETT metric called Markov Chain Based Cost Evaluation (MCCE) metric.

3 Proposed Approach

3.1 System Architecture

The system architecture in the proposed approach is shown in figure 1. It consists of mobile devices and wireless connection. There is no centralized control as shown in figure1.

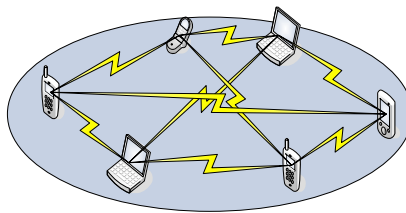


Fig. 1. Scenario in mobile adhoc network

We propose a new metric called MCCE which introduces a new feature of link stability in the selection of the path. The link stability in a network can be defined as-

$$\text{Link Stability} = 1 - e^{-\lambda t} \tag{1}$$

Where λ is the job arrival rate at any instant of time t. The Cost Evaluation (CE) function for a link between nodes i and j can be calculated as follows-

$$CE(i, j) = J_i \times \text{Link Stability} \tag{2}$$

Where J_i is the jitter value which is equal to some threshold value initially and then later it is calculated as

$$J_i = | (CE(i, j)_{\text{time} = t-1}) - (CE(i, j)_{\text{time} = t-2}) | \tag{3}$$

Therefore the metric value for a link between nodes i and j in the path can be calculated as-

$$MCCE(i, j) = WCETT \times CE(i, j) \tag{4}$$

The probability of successful transmission of a link between any two nodes i and j in a network of n number of nodes can be represented using a Markov Chain Probability Matrix (MCPM) as:

$$MCPM = [a_{ij}] , \text{ where } a_{ij} = \frac{MCCE(i, j)}{\sum_{j=1}^n MCCE(i, j)} \tag{5}$$

Pseudo Code

Input: - λ , WCETT value of the link (i, j) , N . **Output:** - $MCCE(i, j)$

Repeat

- Calculate the link stability in the network using equation (1).
- Calculate jitter in the link (i, j) using last two values of CE from equation (3).
- Calculate the value of CE (i, j) at time t using the equation (2).
- The metric value $MCCE(i, j)$ for the link (i, j) is calculated from equation (4).

While $(N \neq \phi)$

The Markov Probability Matrix is calculated using the equations (5) and (6).

4 Simulation Results and Comparison

The performance of proposed algorithms is evaluated using ns-2 [11] with respect to packet delivery fraction and average throughput. Figure 2 shows the impact of the proposed scheme on existing WCETT based scheme. The results obtained show

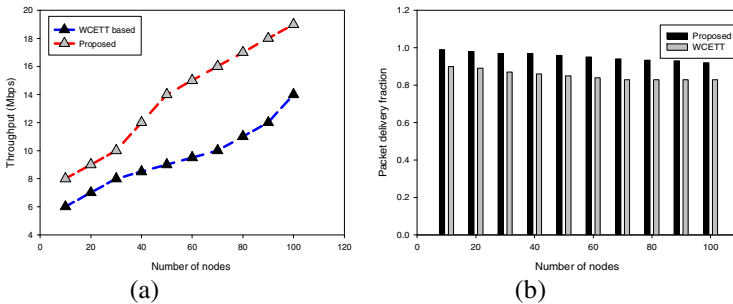


Fig. 2(a)-(b). Throughput and packet delivery fraction in proposed and existing schemes

Table 1. Relative comparison of routing metrics

Routing Metrics	State Independent Link stability	Load Balancing	Routing Decision
Hop Count [1999]	No	No	Per hop basis
ETX [2003]	No	No	Number of retransmission
ETT [2004]	No	No	Transmission time
WCETT [2004]	No	No	Transmission time
MCCE [2010]	Yes	Yes	Link stability

that the proposed scheme has higher throughput and packet delivery fraction than WCETT scheme with an increase in number of nodes.

5 Conclusions

MCCE metric can enhance the functionality of WCETT metric by introducing link stability in the calculation of optimal paths. The metric is very useful for applications in real time applications where the nodes are highly mobile and it is difficult to estimate the exact position of a node. Moreover, the proposed metric determines more reliable paths with less packet loss and delivery time. The results obtained also show that the proposed scheme is quite effective with respect to the existing scheme in terms of packet delivery fraction and throughput.

References

1. Safwat, A.M., Hassanein, H.S.: Infrastructure-based routing in wireless mobile ad hoc networks. *Computer Communications* 25(3), 210–224 (2002)
2. IETF MANET Working Group Charter, <http://www.ietf.org/html.charters/manetcharter.html>
3. Raimundo, J., Macêdo, A., Assis Silva, F.M.: The mobile groups approach for the coordination of mobile agents. *J. Parallel Distributed Computing* 65, 275–288 (2005)
4. Das, S.R., Perkins, C.E., Royer, E.M.: Performance Comparison of Two On demand Routing Protocols for Ad Hoc Networks. In: *IEEE INFOCOM* (March 2000)
5. Spohn, M., Garcia-Luna-Aceves, J.J.: Neighborhood aware source routing. In: *Proceedings of ACM MobiHoc, 6th IEEE/ ACIS International Conference on Computer and Information Science, LongBeach, CA* (2001)
6. Johnson, D.B., Maltz, D.A.: Dynamic source routing in ad hoc Wireless networks. *Mobile Computing*, 153–181 (1996)
7. Perkins, C., Belding-Royer, E., Das, S.: Ad hoc On-Demand Distance Vector (AODV) Routing. In: *IETF.RFC*, vol. 3561 (2003), <http://tools.ietf.org/html/rfc3561>
8. Tønnessen, A.: Implementing and extending the Optimized Link State Routing Protocol. *UniK University Graduate Center, University of Oslo* (2004)
9. Couto, D.S.J.D., Aguayo, D., Bicket, J., Morris, R.: A high throughput path metric for multi-hop wireless routing. In: *Proc. MOBICOM* (2003)
10. Draves, R., Padhye, J., Zill, B.: Routing in multi-radio, multi-hop wireless mesh network. In: *Proc. MOBICOM* (2004)
11. Fall, K., Varadhan, K. (eds.): NS notes and documentation. The VINT project, LBL (February 2000), <http://www.isi.edu/nsnam/ns/S>

Texture Image Classification Using Gray Level Weight Matrix (GLWM)

R.S. Sabeenian¹ and P.M. Dinesh²

¹ Professor and Centre Head Sona SIPRO, ² Research Associate, Sona SIPRO,
ECE – Department, Sona College of Technology, Salem, India
sabeenian@sonatech.ac.in, pmdineshece@live.com

Abstract. The texture analysis plays an important role in image processing and image classification field. Texture is an important spatial feature useful for identifying objects in an image. The local binary pattern and entropy are the most popular statistical methods used in practice to measure the textural information of images. Here, we proposed new statistical approach for the classification of texture images. In this method, the local texture information for a given pixel and its neighborhood is characterized by the corresponding texture unit and the global textural aspect of an image is revealed by its texture spectrum. The proposed method extracts the textural information of an image with a more complete respect of texture characteristics.

Keywords: Textures Spectrum, Local binary pattern, Images, Statistical methods, Spatial Features, Gray Level Weight Matrix.

1 Introduction

In remote sensing data with a high spatial resolution, some of the image elements are represented by a group of pixels, not by only one pixel. This means that image classification and interpretation based on the analysis of individual pixels will result in a relatively high rate of classification and will no longer be sufficient to satisfy the needs. A good understanding or a more satisfactory interpretation of remotely sensed imagery should include descriptions of both the spectral and textural aspects.

The purpose of this study is to present a new statistical method of texture analysis [2], which is focused on texture characterization and discrimination. The concept of texture unit is proposed first. It may be considered, as the smallest complete unit, which best characterizes the local texture aspect of a given pixel and its neighborhood in all eight directions of a square raster. Then a texture image is characterized by its features like local binary pattern, texture spectrum and entropy, which describe the distribution of all the texture units within the image. Some natural images have been used to evaluate the discriminating performance of the texture spectrum.

2 Previous Method

2.1 Local Binary Pattern (LBP)

LBP is invariant against any monotonic gray scale transformation. The method is rotation variant like most existing texture measures. LBP does not address the contrast of texture, which is important in the discrimination of some textures. For this purpose, we can combine LBP with a simple contrast measure, and consider joint occurrences of LBP. LBP and LBP/Contrast perform well also for small image regions (e.g., 16x16 pixels), which is very important. A LBP is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice-versa.

2.2 Entropy

If an image is interpreted as a sample of a “gray level source” that emitted it, we can model that source’s symbol probabilities using the gray-level histogram of the observed image and generate an estimate, called the first-order estimate, H , of the source entropy

$$H = - \sum_{k=1}^L P_r r_k \log P_r(r_k) \quad \text{where } k = 1, 2, \dots, L$$

2.3 Texture Spectrum

The previously defined set of 6561 texture units describes the local-texture aspect of a given pixel. Thus, the statistics of the frequency of occurrence of all the texture units over a large region of an image should reveal texture information. We termed the texture spectrum the frequency distribution of all the texture units, with the abscissa indicating the texture unit number N_{TU} and the ordinate representing its occurrence frequency. Where N_{TU} is the texture unit number, E_i is the i^{th} element of texture unit set

$$N_{TU} = \sum_{i=1}^8 E_i 3^{i-1}$$

3 Proposed Method

In LBP the signs of the eight differences are recorded into an 8-bit number. The original 3x3 neighborhood is thresholded by the value of the center pixel [4].

In proposed method instead of thresholding the image transforming is made with neighborhood to a texture unit with the texture unit number under the ordering way as shown in figure1.

The transforming conditions

$$E_i = \begin{cases} 0 & \text{if } V_i < V_o \\ 1 & \text{if } V_i = V_o \\ 2 & \text{if } V_i > V_o \end{cases}$$

Where:

V_i = The center pixel value
 V_o = The neighboring pixel value

Neighborhood (V_i)

62	85	92
29	40	36
67	36	66



Texture unit (E_i)

2	2	2
0	1	0
2	0	2

Fig. 1. Texture unit transformation

The values of the pixels in the transformed texture unit neighborhood are multiplied by the weights given to the corresponding pixels. Finally, the values of the eight pixels are summed to obtain a number for this neighborhood. This method considers the center pixel value till end of the process for each 3 x 3 matrix. The equation involved in the calculation of GLWM is shown below.

$$GLWM = \sum_{i=0}^8 E_i \times 2^i$$

4 Results and Discussion

In order to evaluate the performance of the texture spectrum in texture characterization and classification, several experimental studies have been carried out on 9 of Brodatz's images [1]. These images were selected because they are broadly similar to one another and also that they resemble parts of remotely sensed images.

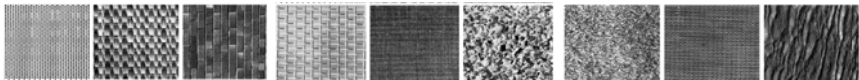


Fig. 2. Nine Brodatz's text images

Table 1. Analysis of Texture Image Classification

Method	LBP	TSO	Entropy	GLWM	Method	LBP	TSO	Entropy	GLWM
Image1	94.41	91.25	78.09	95.63	Image6	84.24	82.5	87.5	89.25
Image2	96.6	86.15	81.57	94.32	Image7	82.67	93.5	93.6	94.32
Image3	85.82	87.82	82.68	91.35	Image8	93.51	96.8	76.55	94.64
Image4	93.25	93.5	88.55	93.64	Image9	80.46	87.8	86.5	88.65
Image5	95.37	92.5	96.52	95.37					

We evaluated local binary pattern, texture spectrum, entropy and GLWM. The results obtained are tabulated. According to tabulation results the GLWM technique has the best one. The GLWM technique is calculated with overlapping of each 3 x 3 of the image hence the result of this technique is good when compared to other.

We extended our database to 60 image of the Brodatz's texture database [1]. The classification rate of image at different categories are tabulated below

Table 2. Classification percentage for 60 images of database

Image Size	No of Sample	Classification Rate
512 X 512	60 X 1	100%
256 X 256	60 X 4	98%
128 X 128	60 X 16	95%
64 X 64	60 X 64	94%
Total Classification Rate		96.75 %

5 Conclusion

Based on the concept of GLWM, Local Binary Pattern, texture spectrum operator and entropy measure, texture analysis has been presented. We used 60 images from the Brodatz's database. From them we took 5100 samples 512 X 512 size of 60 samples, 256 X 256 size of 240 samples, 128 X 128 size of 960 samples, 64 X 64 size of 3840 samples. Totally 5100 samples among them GLWM able to classify 4935 samples successfully. Evaluations show that the GLWM is able to reveal texture information in digital images.

References

- [1] Lin, H., Wang, J.-y., Liu, S.: The Classification Study of Texture Image Based on the Rough Set Theory. In: 2010 IEEE International Conference on Proceeding of the IEEE xplore Granular Computing (GrC), August 14-16 (2010)
- [2] Khelifi, R., Adel, M., Bourennane, S.: Texture classification for multi-spectral images using spatial and spectral Gray Level Differences. In: 2010 2nd International Conference on Proceeding of the IEEE xplore Image Processing Theory Tools and Applications (IPTA), July 7-10 (2010)
- [3] Sabeenian, R.S., Palanisamy, V.: Texture Based Medical Image Classification of Computed Tomography images using MRCSF. Published in International Journal of Medical Engineering and Informatics (IJMEI) Published by Inderscience Publications 1(4), 459-472 (2009)
- [4] Sabeenian, R.S., Palanisamy, V.: Crop and Weed Discrimination in Agricultural Field using MRCSF Published. In the International Journal of Signal and Imaging Systems Engineering (IJSISE) Published by Inderscience Publishers 3(1), 61-69

Formal Verification of IEEE802.11i WPA-GPG Authentication Protocol

K.V. Krishnam Raju¹ and V. Valli Kumari²

¹ Dept. of Computer Science Engineering, SRKR Engineering College, Bhimavaram, Andhra Pradesh, India

kvkraju.srkr@gmail.com

² Dept. of Computer Science & Systems Engineering, Andhra University, Visakhapatnam, Andhra Pradesh, India, 5300 03

vallikumari@gmail.com

Abstract. IEEE802.11i is the standard designed to provide secured communication of wireless LAN. The IEEE802.11i specification contains WPA-GPG authentication protocol. It allows a wireless station to gain access to a protected wireless network managed by access point. This paper models the WPA-GPG authentication protocol by formal verification using CasperFDR and analyzes the output. A few attacks are found in this protocol. The specifications through which these attacks are found are presented.

Keywords: Formal Verification, WPA-GPG, CasperFDR.

1 Introduction

WPA protocol is a subset of IEEE 802.11i standard. The IEEE 802.11i standard is designed to provide secured communication of wireless LAN as defined by all the IEEE 802.11 specifications. IEEE 802.11i enhances the WEP (Wire line Equivalent Privacy), a technology used for many years for the WLAN security, in the areas of encryption, authentication and key management. IEEE 802.11i is based on the Wi-Fi Protected Access (WPA), which is a quick fix of the WEP weaknesses [1].

The IEEE 802.11i standard includes authentication, encryption and data integrity check of three components, is a complete security program. The core is 802.1x (Port Access Control) and TKIP (Temporal Key Integrity Protocol). The IEEE 802.1X offers an effective framework for authenticating and controlling user traffic to a protected network, as well as dynamically varying encryption keys. 802.1X ties a protocol called EAP (Extensible Authentication Protocol) to both the wired and wireless LAN media and supports multiple authentication methods, such as token cards, Kerberos, one-time passwords, certificates, and public key authentication.

WPA-GPG is a modification on WPA-PSK authentication protocol that allows a wireless station (also known as supplicant or STA) to gain access to a protected wireless network, managed by an Access Point (also known as AP or authenticator), by means of its personal GPG key [2]. WPA-GPG verifies that the authenticating STA is the owner of the key that it is providing to the AP, this represents the core of the

authentication process. WPA-GPG does not require specific hardware nor firmware since, its modifications to the original protocol are important but not radical. WPA-GPG ensures the same level of security of WPA-PSK but it also guarantees in-network user privacy and non-repudiation of network traffic. With the term in-network user privacy we mean that network traffic is encrypted in a way that even STAs which have the right to authenticate are not able to decrypt it. As we will show this is not possible in WPA-PSK. As a direct consequence WPA-GPG also allows the AP to know exactly which STA has generated what traffic, ensuring non-repudiation.

The goal of WPA-PSK four-way handshake protocol is to create a PTK known both to supplicant and authenticator while not revealing the PMK. WPA-GPG has exactly the same goal, with the difference that no PSK is shared between STA and AP but STA uses its GPG key to authenticate. In WPA-GPG the PMK is randomly generated by the AP, encrypted and sent to the authenticating STA. Both AP and STA will therefore be able to derive a PTK from the PMK, exactly as it happens with WPA-PSK, with the main difference that each PTK cannot be derived by other STAs.

Modeling and analysis of security protocols with Communicating Sequential Process (CSP) and Failure Divergence Refinement (FDR) have been proven to be effective and have helped the research community find attacks in several protocols. However, modeling directly in CSP is time-consuming and error-prone. Lowe thus designed Casper [3], which takes more abstract descriptions of protocols as input and translates them into CSP. CSP was first described by Hoare in [4] [5], and has been applied in many fields.

First, we formally model and analyze the WPA-PSK protocol with CasperFDR. Next, we use CasperFDR to show that there are no other known attacks on WPA-PSK protocol.

The rest of the paper is organized as follows. Section 2 deals with related works. In Section 3, WPA-GPG protocol is modeled with CasperFDR and is analyzed and finally we conclude in Section 4.

2 Related Works

Johnson and Walker are among the first researchers who discussed the security issues in IEEE 802.16 [6]. Sen Xu, Chin-Tser Huang, Manton M. Matthews [7], have analyzed security issues on the PKMv1, PKMv2 protocols using Casper FDR and proposed solutions. K.V.Krishnam Raju, V.Valli Kumari, N.Sandeep Varma, KVSVN Raju [8], also have analyzed security issues on the PKMv3 protocol using Casper FDR.

According to [9] Wi-Fi Protected access (WPA) is really less secure. Generally wireless protection is enabled using MAC filtering. According to survey there is no clear insight into this protocol by the research community. But much work has been done on other wireless protocols like PKMV1, PKMV2, PKMV3 that belong to 802.16 standards [10]. On the other hand in this paper we have done verification process for WPA-GPG authentication protocol using some standard verification tool like CasperFDR.

3 Modeling and Analysis of WPA-GPG Authentication Protocol

3.1 WPA-GPG Authentication Protocol Structure

Message 1: The first message is sent by the AP and contains the AP random nonce (ANonce). This message is not encrypted and no MIC is attached (the PTK cannot be derived yet).

Message 2: After receiving the first message, the STA will send its nonce (SNonce) and attaches the STA public key (GPG Key) for authentication, and attaches the signature of ANonce and received RSN IE(Information elements) .The signed RSN IE is added to allow AP to verify that the STA has received a correct element and to prevent version rollback attack.

Message 3: After receiving the second message, the AP will verify the validity of STA signature on ANonce, using the key attached to message 2. The AP will generates a random 32 byte PMK which is kept secret and derives the PTK. Attaches IE and GTK to the message and computes message MIC. Encrypt the generated PMK using STA public key.

Message 4: After receiving the third message, the STA will decrypt the PMK with its public key and derives the PTK and verifies the MIC value of message 3. If MIC code verification is successful STA can send the acknowledgment which contains nothing but a MIC and AP can authorize it to access the network.

Message1	AP → STA : Na
Message2	STA → AP : Ns GPG KEY Sign(Na + IE)
Message3	AP → STA : GTK MIC Cry(PMK)
Message4	STA → AP : ACK MIC

Fig. 1. WPA-GPG Authentication Protocol [2]

3.2 Modeling WPA-GPG Authentication Protocol in CasperFDR

The modeled WPA-GPG authentication protocol in CasperFDR is shown in Fig.2. In the specification the initiator A and responder S represents Authenticator and Supplicant.

3.3 Analysis of WAP-GPG Authentication Protocol with CasperFDR

After compiling and checking the above model in CasperFDR tool, attacks were found for two of the four properties declared in specification part in Fig.2. Out of four property1 and property2 are related to secret specifications next property3 and property4 are related to authentication specifications. CasperFDR tool found attacks on every specification in the specification part.


```

#Free variables
A, S : Agent
na, ns: Nonce
kpm, kpt, gtk: SessionKey
gpg : Agent →PublicKey
sk : Agent→ SecretKey
f: HashFunction
InverseKeys = (gpg,sk), (kpm,kpm), (kpt,kpt), (gtk,gtk)
#Processes
INITIATOR(A,na,kpt,kpm,gtk) knows gpg, sk(A)
RESPONDER(S,ns,kpt,kpm,gtk) knows gpg, sk(S)
#Protocol description
0. → A : S
[A!=S]
1. A→S : na
2. S→A : ns, gpg(S), {na}{sk(S)}
3. A →S: gtk, {kpm}{gpg(S)}
4. S→A : f(kpm)
#Specification
Secret(A, na, [S])
Secret(S, ns, [A])
Agreement(A, S, [na,ns])
Agreement(S, A, [na,ns])
#Actual variables
Authenticator, Supplicant, Mallory : Agent
Na, Ns, Nm : Nonce
Kpm, Kpt, Gtk : SessionKey
InverseKeys = (Kpm, Kpm), (Kpt, Kpt), (Gtk, Gtk)
#Inline functions
symbolic gpg,sk
#System
INITIATOR(Authenticator, Na, Kpt, Kpm, Gtk)
RESPONDER(Supplicant, Ns, Kpt, Kpm, Gtk)
#Intruder Information
Intruder = Mallory
IntruderKnowledge={ Authenticator,Supplicant,Mallory,Nm,gpg,sk(Mallory)}

```

Fig. 2. WPA-GPG key agreement Protocol Specification

```

0.          → Authenticator : Mallory
1. Authenticator → I_Mallory : Na
1. I_Supplicant → Supplicant : Nm
2. Supplicant   → I_Supplicant : Ns, gpg(Supplicant),{Nm}{sk(Supplicant)}
2. I_Mallory    → Authenticator : Nm, gpg(Mallory), {Na}{sk(Mallory)}
3. Authenticator → I_Mallory : Gtk, {Kpm}{gpg(Mallory)}
3. I_Supplicant → Supplicant : Gtk, {Kpm}{gpg(Supplicant)}
4. Supplicant   → I_Supplicant : f(Kpm)

```

Fig. 3. Attack generated by CasperFDR for property1

An attack on the property1 Secret(A, na, [S]) as shown in Fig.3. It represents the Supplicant believes Ns is secret but the intruder also knows Ns value.

An attack on the property2 Secret(S, na, [A]) as shown in Fig.4.It also represents the Supplicant believes Ns is secret but the intruder also knows Ns value.

- | | |
|------------------|--|
| 0. | → Authenticator : Mallory |
| 1. Authenticator | → I_Mallory : Na |
| 1. I_Supplicant | → Supplicant : Nm |
| 2. Supplicant | → I_Supplicant : Ns, gpg(Supplicant), {Nm}{sk(Supplicant)} |
| 2. I_Mallory | → Authenticator : Nm, gpg(Mallory), {Na}{sk(Mallory)} |
| 3. Authenticator | → I_Mallory : Gtk, {Kpm}{gpg(Mallory)} |
| 3. I_Supplicant | → Supplicant : Gtk, {Kpm}{gpg(Supplicant)} |
| 4. Supplicant | → I_Supplicant : f(Kpm) |

Fig. 4. Attack generated by CasperFDR for property2

An attack on the property3 Agreement(A, S, [na,ns]) as shown in Fig.5. Top level trace generated by CasperFDR is Supplicant believes that it has completed a run of the protocol, taking role RESPONDER, with Supplicant, using data items Nm, Ns. Authenticator also believes that it is running the protocol.

- | | |
|------------------|---|
| 0. | → Authenticator : Mallory |
| 1. I_Supplicant | → Supplicant : Nm |
| 1. Authenticator | → I_Mallory : Na |
| 2. Supplicant | → I_Supplicant : Ns, gpg(Supplicant),{Nm}{sk(Supplicant)} |
| 2. I_Mallory | → Authenticator : Nm, gpg(Mallory), {Na}{sk(Mallory)} |
| 3. Authenticator | → I_Mallory : Gtk, {Kpm}{gpg(Mallory)} |
| 3. I_Supplicant | → Supplicant : Gtk, {Kpm}{gpg(Supplicant)} |
| 4. Supplicant | → I_Supplicant : f(Kpm) |

Fig. 5. Attack generated by CasperFDR for property3

- | | |
|------------------|---|
| 0. | → Authenticator : Supplicant |
| 1. Authenticator | → I_Supplicant : Na |
| 1. I_Mallory | → Supplicant : Na |
| 2. Supplicant | → I_Mallory : Ns, gpg(Supplicant), {Na}{sk(Supplicant)} |
| 2. I_Supplicant | → Authenticator : Nm, gpg(Supplicant), {Na}{sk(Supplicant)} |
| 3. Authenticator | → I_Supplicant : Gtk, {Kpm}{gpg(Supplicant)} |
| 3. I_Mallory | → Supplicant : Gtk, {Kpm}{gpg(Supplicant)} |
| 4. Supplicant | → I_Mallory : f(Kpm) |
| 4. I_Supplicant | → Authenticator : f(Kpm) |

Fig. 6. Attack generated by CasperFDR for property4

An attack on the property4 Agreement(S, A, [na,ns]) as shown in Fig.6. Top level trace generated by CasperFDR is Authenticator believes (s)he has completed a run of

the protocol, taking role INITIATOR, with Supplicant, using data items Na, Nm. Supplicant also believes that it is running the protocol.

From the above attacks we can find that not only the Authenticator and the Supplicant but the Intruder can also derive GTK key from PMK that is the key used between Authenticator and Supplicant for message encryption.

4 Conclusions and Future Work

In this paper, the WPA-GPG protocol is modeled using CasperFDR. The compilation was done with CasperFDR. Attacks were found in this protocol. The attacks that are interpreted by CasperFDR and the message sequence results are reported. In future we will fix the attacks found in the WPA-GPG protocol.

References

- [1] IEEE 802.11i. WLAN Security Standards, <http://www.javvin.com/protocol80211i.html>
- [2] WPA-GPG. Wireless authentication using GPG Key, Gabriele Monti, December 9 (2009)
- [3] Lowe, G.: Casper: A compiler for the analysis of security protocols. *Journal of Computer Security* 6, 53–84 (1998)
- [4] Hoare, C.A.R.: Communicating sequential processes. *Commun. ACM* 21(8), 666–677 (1978)
- [5] Hoare, C.A.R. (ed.): *Communicating Sequential Processes*. Prentice Hall International (1985)
- [6] Johnston, D., Walker, J.: Overview of IEEE 802.16 security. *IEEE Security & Privacy* (2004)
- [7] Xu, S., Matthews, M.M., Huang, C.-T.: Modeling and Analysis of IEEE 802.16 PKM Protocols using CasperFDR. In: *IEEE ISWCS 2008* (2008)
- [8] Krishnam Raju, K.V., Valli kumari, V., Sandeep varma, N., Raju, K.V.S.V.N: Formal Verification of IEEE802.16m PKMv3 Protocol Using CasperFDR. In: *International Conference on advances in information and communication technologies, ICT 2010, Kerala, India* (September 2010)
- [9] Wpa-Wpa2, <http://www.lylebackemort.com/blog/2008/5/10/wpa-wpa2-as-insecure-as-i-expected> (accessed on March 14, 2010)
- [10] Xu, S., Huang, C.-T., Matthews, M.M.: Modeling and Analysis of IEEE 802.16 PKM Protocol using CasperFDR. University of South California, Columbia SC, 29208

Adaptive Steganography Based on Covariance and Dct

N. Sathisha¹, Swetha Sreedharan², R. Ujwal², Kiran D'sa², Aneeshwar R. Danda²,
K. Suresh Babu³, K.B. Raja³, K.R.Venugopal³, and L.M. Patnaik⁴

¹ Department of Electronics and Communication Engineering,
R L Jalappa Institute of Technology, Doddaballapura, Bangalore Rural Dist. 561 203, India
nsathisha@gmail.com

² Department of Electronics and Communication Engineering, BMSIT, Bangalore, India
swethasree2@yahoo.com

³ Department of Computer Science and Engineering,
University Visvesvaraya College of Engineering, Bangalore University, Bangalore 560 001,

⁴ Defence Institute of Advanced Technologies, Pune, India

Abstract. The steganography is a covert communication to transfer confidential information over an internet. In this paper we propose Adaptive Steganography based on Covariance and Discrete Cosine Transform (ASCDCT) algorithm. The Average Covariance of the Cover Image (ACCI) is computed. The ACCI of 0.15 is considered as the threshold value. The cover image is segmented into 8*8 cells and DCT is applied to derive coefficients. The payload Most Significant Bits (MSBs) are embedded into the cover image based on ACCI and DCT coefficients. It is observed that the capacity, Peak Signal to Noise Ratio (PSNR) and security is better compared to the existing algorithm.

Keywords: Covariance, Cover Image, DCT, Payload, Steganography.

1 Introduction

Steganography is an art and science of hiding confidential information into a cover media to ensure the security of information over the communication channel. The cover media can be text, audio, video and image. Weiqi Luo et al., [1] embeds the secret message into sharper edge regions of cover image adaptively according to size of the message and the gradients of the content edges of cover image. Cheng-Hsing Yang et al., [2] developed a technique to embed the secret information by Pixel Value Differing (PVD) method. The number of secret bits embedded depends on the difference between two consecutive pixels. Bo-Luen Lai and Long-Wen Chang [3] proposed a transform domain based adaptive data hiding method using haar discrete wavelet transform. The image was divided into sub-bands (LL1, HL1, LH1 and HH1) and most of the data is hidden in the edge region as it is insensitive to the Human eye. If these sub-bands were complex, then further division of the bands were performed so that more number of data bits could be embedded. Raja et al., [4] proposed a high capacity, secure steganographic algorithm in which the payload bits are encrypted and embedded in the wavelet coefficients of the cover image. This method utilizes the

approximation band of the wavelet domain to improve robustness. Carlos velsco et al., [5] presented an adaptive data hiding method using convolutional codes and synchronization bits in DCT domain. The cover image is divided into suitable and ineligible blocks based on the DCT energy features from the horizontal, vertical and diagonal frequency information. The suitable blocks are used for embedding data using Quantization Index Modulation (QIM). The two synchronization bits are used for desynchronization problem and convolution codes are used for decoding errors.

2 Algorithm

Problem definition: The algorithm is given in the Table 1 and objectives are (i) The payload is to be embedded into the cover image to derive stegoimage (ii) The high capacity and security with reasonable PSNR.

Table 1. Algorithm of ASCDCT

Input: Cover Image and Payload; Output: Stego Image.	
Step 1)	A cover image of any size and format is considered and if it is color image convert it into grayscale image.
Step 2)	Applying pixel management to the cover image to avoid overflow and underflow of the pixel values 0 and 255.
Step 3)	Covariance of cover image is determined and average is computed to get average covariance.
Step 4)	The average covariance of cover image value is fixed as 0.15, if ACCI > 0.15 go to step 5 else step 6
Step 5)	<ul style="list-style-type: none"> (i) The cover image is segmented into 8*8 matrix and DCT is applied on each matrix. (ii) Payload bit length L to be embedded based on DCT coefficients: $L=4$, if $Co \geq 2^5$; $L=3$, if $2^4 \leq Co \leq 2^5$; $L=2$, if $2^3 \leq Co \leq 2^4$; else $L=1$; (iii) The stego image obtained in the DCT domain is converted back to the spatial domain using IDCT.
Step 6)	<ul style="list-style-type: none"> (i) The cover image is segmented into 8*8 matrix and DCT is applied on each matrix. (ii) Payload bit length L to be embedded based on DCT coefficients: $L=5$, if $Co \geq 2^5$; $L=4$, if $2^4 \leq Co \leq 2^5$; $L=3$, if $2^3 \leq Co \leq 2^4$; else $L=2$; (iii) The stego image obtained in the DCT domain is converted back into the spatial domain using IDCT.

3 Performance Analysis

The payload is embedded into the DCT coefficients of cover image based on ACCI and length L . The performance parameter such as PSNR between cover image and stegoimage is computed and given in the Table 2. It is observed that the PSNR depends on the ACCI of the cover image and also the PSNR decreases as the Hiding

Table 2. ACCI and PSNR for different cover images with payload boat

Cover Image	ACCI	H C	H C	H C	H C
		12.5%	20.0%	25.0%	34.0%
		PSNR	PSNR	PSNR	PSNR
Lena	0.064	46.31	41.51	40.51	39.41
Old image	0.122	45.41	41.95	40.85	39.09
Baboon	0.154	47.25	42.64	41.88	40.71
Barbara	0.203	47.08	43.48	42.54	41.17
Ranch House	0.280	46.06	41.50	40.46	39.20
Bridge	0.358	46.13	41.78	40.96	39.69
Casa	0.504	47.56	43.33	42.57	41.35

Capacity (HC) increases. The PSNR value is maintained around 42 dB for the capacity of 34%.

The Maximum Hiding Capacity (MHC) and the PSNR between the Cover Image (CI) and stego image for different Payload (PL) is tabulated for existing algorithm *An Adaptive Steganographic Technique Based on Integer Wavelet Transform (ASIWT)* [6] and the proposed algorithm ASCDCT is given in the Table 3. It is observed that the PSNR is improved in the proposed algorithm compared to the existing algorithm is due to ACCI.

Table 3. PSNR of existing and proposed techniques for a MHC of 47%

Image	Existing Method (ASIWT)	Proposed Method (ASCDCT)
	PSNR	PSNR
CI: Lena; PL: Barbara	31.80	39.35
CI: Baboon; PL: Cameraman	30.89	37.96

4 Conclusions

The steganography is used to transfer secret message over open channel. In this paper ASCDCT is proposed. The cover image covariance is computed to consider number of MSBs of payload to be embedded based on DCT coefficients. The cover image is divided into 8*8 cells and converted into DCT coefficients to determine the length of the payload bits to be embedded into the cover image. It is observed that the capacity, security and the PSNR values are improved compared to the existing algorithm. In future the robustness of algorithm can be verified.

References

1. Luo, W., Huang, F., Huang, J.: Edge Adaptive Image Steganography Based on LSB Matching Revisited. *IEEE Transactions on Information Forensics and Security* 5, 173–178 (2010)
2. Yang, C.-H., Weng, C.-Y., Wang, S.-J., Sun, H.-M.: Adaptive Data Hiding in Edge Areas of Images with Spatial LSB Domain Systems. *IEEE Transactions on Information Forensics and Security* 3, 488–497 (2008)

3. Lai, B.-L., Chang, L.-W.: Adaptive Data Hiding for Images Based on Harr Discrete Wavelet Transform. In: Chang, L.-W., Lie, W.-N. (eds.) PSIVT 2006. LNCS, vol. 4319, pp. 1085–1093. Springer, Heidelberg (2006)
4. Raja, K.B., Vikas, Venugopal, K.R., Patnaik, L.M.: High Capacity Lossless Secure Image Steganography using Wavelets. In: International Conference on Advances Computing and Communications, pp. 230–235 (2006)
5. Velasco, C., Nano, M., Perez, H., Martinez, R., Yamaguchik: Adaptive JPEG Steganography using Convolutional Codes and Synchronization Bits in DCT Domain. In: 52nd IEEE International Midwest Symposium on Circuits and Systems, pp. 842–847 (2010)
6. El Safy, R.O., Zayed, H.H., El Dessouki, A.: An Adaptive Steganographic Technique Based on Integer Wavelet Transform. In: IEEE Proceedings on International Conference on Networks and Media, pp. 111–117 (2009)

Image Segmentation Using Grey Scale Weighted Average Method and Type-2 Fuzzy Logic Systems

Saikat Maity¹ and Jaya Sil²

¹ Dr. B.C. Roy Engineering College
Tel.: +919932252306
saikat.maity@bcrec.org

² Bengal Engineering & Science University (B.E.S.U)
Tel.: +919433283641
js@cs.becs.ac.in

Abstract. In the paper, the difficulty in image segmentation based on the popular level set framework to handle an arbitrary number of regions has been addressed. There is very few work reported on optimized segmentation with respect to the number of regions. In the proposed model, first the image is classified using type-2 fuzzy logic to handle uncertainty in determining pixels in different color regions. Grey scale average (GSA) method has been applied for finding accurate edge map to segment the image that produces variable number of regions.

Keywords: Segmentation, Grey Scale Average (GSA), Type-2 fuzzy logic, Fuzzy weighted average (FWA).

1 Introduction

Edge of an image represents the sudden change in the pixel intensity value creating two separate regions of different intensities. However, other factors like poor focus or refraction may be the cause of formation of an edge in the image [1]. Relatively early, the problem of image segmentation has been formalized by Health [5] as the minimization of an energy function that penalizes deviations from smoothness within regions and the length of their boundaries [4]. In the level set approach [2], N regions are represented using exactly $\log_2 N$ level set functions. Grey scale average (GSA) method has been proposed here for finding accurate edge map to segment the image that produces variable number of regions [3], described in figure-1. The paper has been divided into five sections. Section 2 describes type-2 fuzzy logic to classify the color regions while in section 3, fuzzification and defuzzification method in image segmentation and how to determine the grey scale average has been discussed. The proposed algorithm with masking and GSA method is presented in section 4 followed by experimental results and discussions in section 5.

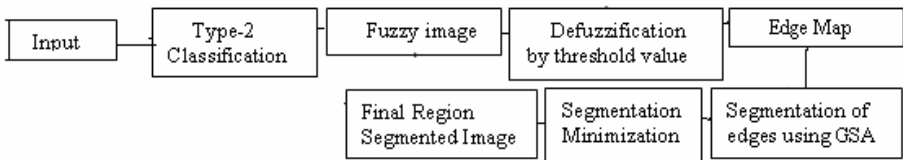


Fig. 1. Complete Block Diagram of the System

2 Classification of Image

Fuzzy Gaussian filter [6] is used to remove noise and sharpen the image, a preprocessing step of edge detection with the objective of enhancing the edges [7]. After removal of noise, the image is classified using type-2 fuzzy rule based classifier.

Let the pixel intensity of an input image $f(x,y)$ is represented by a vector $X = (x_1, x_2, \dots, x_m)$ where m is the number of pixels in the image. Assume r is the total number of fuzzy rules and one of such rule is represented as: If x_1 is R_{j1} and x_2 is R_{j2} and \dots and x_m is R_{jm} then $X = (x_1, x_2, x_3, \dots, x_m)$ belongs to class k with $CF = CF_j$ where $j=1,2,\dots,r$ and $k=1,2,\dots,M$ is the number of color classes classified using equation (1).

$$c = \arg \max_j G_i(x).CF_j \tag{1}$$

$CF_j \in [0,1]$ is the certainty factor of the j -th rule such that

$G_i(x) = \prod_{i=1}^m \mu_R(x_i)$ where $\mu_R(x_i)$ is the degree of membership value of pixel i in type-2 fuzzy set R .

3 Grayscale Weighted Average Method

The classified regions are now segmented by detecting edges using GSA method. All the edge points of the image constitute a set, called an edge map. Edge map is the specific region bounded by neighborhood pixels within the same object, shown in figure-2 where the color lines are the edges separated by different b/w color intensity.

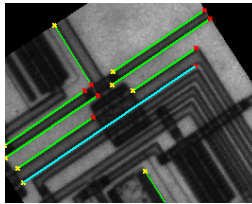


Fig. 2. Edge Map in B/W Image

Pixels with membership value ‘1’/‘0’ definitely belong/don’t belong to the edge map set. However, pixels with intermediate membership values may or may not belong to the edge map, which are determined with certainty depending upon a prescribed threshold value. After thresholding, a binary image is obtained, which is the edge map representation of the original image.

The generalized version of the segmentation problem with an arbitrary number of regions N is considered by providing the energy function of the model in (2).

$$E(\Omega_i, p_i, N) = \sum_{i=1}^N \int \log p_i dx + \frac{\nu}{2} \int_{\Gamma_i} ds + \lambda \tag{2}$$

Where $\Omega_i = i$ -th region, $p_i =$ a priori probability of i^{th} pixel, $\Gamma_i = i^{\text{th}}$ region boundary, $\nu =$ weighted parameter of boundary Γ , ds is the deviation of distance between two region and the additional term of this energy functional penalizes the number of regions with the parameter λ . Starting with the entire image domain Ω as a single region, the two-region segmentation is applied in order to determine the best splitting of the domain. If energy decreases by splitting, two regions are formed, which are again divided and so on, until the energy does not decrease by further splits and thus the optimum number of regions is determined.

GSA deals with the grey scale value having the same range between 0-1 like FWA method. For the grey scale, the weighted average GS_{avg} is computed as given in (3)

$$GS_{avg} = \sum_{i=1}^{255} w_i p_i / \sum_{i=1}^{255} w_i = f(w_1, \dots, w_{255}, p_1, \dots, p_{255}) \quad (3)$$

Where w_i is the weighted intensity and p_i is the pixel intensity.

4 Proposed Algorithm with GSA

Begin

Read input image $f(x, y)$ of size $M \times N$

Create the mask $W (m \times n)$ with mask coefficients, using

Sparse matrix so that sum of all coefficients of each

Let mask=0 ;

Mask weighted average, $a=(m-1)/2$ // small size

Mask weighted average, $b=(n-1)/2$ // large size.

$g(x, y) = 0$ // output image

For $y = b$ to $(N - b - 1)$ do

For $x = a$ to $(M - a - 1)$ do

Calculate the (*largest value*) among all the maximum column values;

Calculate the (*smallest value*) among all the minimum column values;

$GSA_range = (largest\ value) - (smallest\ value) / n$

For $x = 1$ to N do

for $y = 1$ to M do

$g(x, y) = (f(x, y) - smallest\ pixel\ value) \times 255 / GSA_range;$

End for

Select a color image and convert it into grey scale input image;

Store pixel values of the image along x and y coordinates in matrix form;

Generate the *Convolution mask* for different gradient operators and store it ;

$SUM_{coeffmatrix} = 0;$ //Set all the points as black //

Each mask along the horizontal and vertical direction is *convolved* with the input image; Calculate magnitude of the gradient vector;

If $E_{threshold} < 0.55$ // Threshold is required to determine whether the point belongs to a specific region or not //

Truncate unwanted edges from edge map information;

Else include edge in the edge map set;

End_for

End.

5 Conclusion and Result Discussion

The proposed algorithm is reasonably fast where 169×250 size image took 22.5 seconds on an Intel Atom 1.6 Ghz Processor having 1 GB of RAM size. By keeping the advantages of the level set framework, its main problem has been solved in the paper. The customized Gaussian filter [8] has a good contrast and sharpness characteristics, which is required to sharpen an image. This paper shows that the proposed segmentation method exhibits better performance in edge detection compare to the conventional method of segmentation [9].

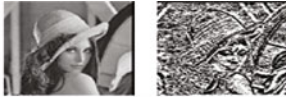


Fig - 2a Edge detection using Sobel Filter



Fig - 3 Edge detection by Custom Filter



Fig-4 Segmentation by Conventional Method



Fig-5 Segmentation using Proposed Method



Fig-6 CMS

Fig-7 PMS

References

- [1] Argyle, E.: Techniques for edge detection. Proc. IEEE 59, 285–286 (1971)
- [2] Chidiac, H., Ziou, D.: Classification of Image Edges. In: Vision Interface 1999, pp. 17–24. Troise-Rivieres, Canada (1999)
- [3] Hueckel, M.: A local visual operator which recognizes edges and line. J. ACM 20(4), 634–647 (1973)
- [4] Malik, J., Belongie, S., Shi, J., Leung, T.K.: Textons, contours and regions: cue integration in image segmentation. In: Proc. IEEE International Conference on Computer Vision, Corfu, Greece, pp. 918–925 (September 1999)
- [5] Heath, M., Sarkar, S., Sanocki, T., Bowyer, K.W.: Comparison of Edge Detectors: A Methodology and Initial Study. Computer Vision and Image Understanding 69(1), 38–54 (1998)
- [6] Shin, M.C., Goldgof, D., Bowyer, K.W.: Comparison of Edge Detector Performance through Use in an Object Recognition Task. Computer Vision and Image Understanding 84(1), 160–178 (2001)
- [7] Peli, T., Malah, D.: A Study of Edge Detection Algorithms. Computer Graphics and Image Processing 20, 1–21 (1982)
- [8] Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 2nd edn. Prentice-Hall, Inc., Upper Saddle River (2002)
- [9] Pratt, W.K.: Digital Image Processing, 4th edn. John Wiley & Sons, Inc., Hoboken (2007)

Cluster Analysis and Pso for Software Cost Estimation

Tegiyot Singh Sethi¹, CH.V.M.K. Hari², B.S.S. Kaushal¹, and Abhishek Sharma¹

¹ Dept. of CSE, Gitam University, Visakhapatnam, India
{tjss401, kaushalb09, abhisheksharma9129}@gmail.com

² Dept. of IT, Gitam University, Visakhapatnam, India
kurmahari@gmail.com

Abstract. The modern day software industry has seen an increase in the number of software projects. With the increase in the size and the scale of such projects it has become necessary to perform an accurate requirement analysis early in the project development phase in order to perform a cost benefit analysis. Software cost estimation is the process of gauging the amount of effort required to build a software project. In this paper we have proposed a Particle Swarm Optimization (PSO) technique which operates on data sets which are clustered using the K-means clustering algorithm. The PSO generates the parameter values of the COCOMO model for each of the clusters of data values. As clustering encompasses similar objects under each group PSO tuning is more efficient and hence it generates better results and can be used for large data sets to give accurate results. Here we have tested the model on the COCOMO81 dataset and also compared the obtained values with standard COCOMO model. It is found that the developed model provides better estimation of the effort.

Keywords: Particle Swarm Optimization (PSO), K-Means, Software Cost Estimation, Constructive Cost Model (COCOMO).

1 Introduction

The software industry today is all about efficiency. The provident allocation of the available resources and the judicious estimation of the requisites form the basis of any planning and scheduling activity. With the increase in the expanse and impact of modern day software projects, the need for accurate requirement analysis early in the software development phase has become pivotal. For a given set of requirements, it is desirable to cognize the amount of time and money required to deliver the project prolifically. The chief aim of software cost estimation is to enable the client and the developer to perform a cost – benefit analysis. The cost / effort estimates are determined in terms of person-months(pm) which can be easily commuted to actual currency cost. The cost of the software varies depending on both complexity and lines of code and to estimate the cost we make use of Particle Swarm Optimization on clustered data. A common approach to the estimation of the software effort is by expressing it as a function of the project size[1] and the Effort Adjustment Factor (EAF). The equation of effort in terms of size and methodology is considered as follows:

$$\text{Effort} = a * (\text{size})^b * (\text{EAF}) + c \quad (1)$$

Here a, b, c are constants. The constants are usually determined by regression analysis applied to historical data[11]. There are a number of models proposed for tuning parameters using Neural Networks, Machine learning techniques, fuzzy techniques and Genetic algorithms[2][3][5][6][8].

PSO is a robust stochastic optimization technique [4][9][10] based on the movement of intelligent swarms . PSO applies the concept of social interaction to problem solving. It uses a number of agents (particles) that constitutes a swarm moving around in the search space looking for the best solution. Each particle is treated as a point in an N- dimensional space which adjusts its movement. According to its own flying experience (P_{best} - personal best) as well as the flying experience of other particles (G_{best} –global best). The basic concept of PSO lies in accelerating each particle towards its Pbest and Gbest locations with regard to a random weighted acceleration at each time. The modifications of the particle’s positions can be mathematically modeled by making use of the following equations:

$$V^{k+1} = wV_i^k + c_1 \text{rand}_1 (P_{\text{best}} - S_i^k) + c_2 \text{rand}_2 (G_{\text{best}} - S_i^k) \quad (2)$$

$$S_i^{k+1} = S_i^k + V_i^{k+1} \quad (3)$$

Where, S_i^k is current search point; S_i^{k+1} is modified search point; V_i^k is the current velocity; V^{k+1} is the modified velocity; V_{pbest} is the velocity based on P_{best} ; V_{gbest} = velocity based on G_{best} ; w is the weighting function; c_j is the weighting factors; rand_j are uniformly distributed random numbers between 0 and 1. In the particle swarm optimization technique, the particles searches the solutions in the solution space within the range [-s,s]. The PSO works better when operated on datasets having similar valued objects. Hence to enhance the PSO computations we have clustered the given dataset into groups of similar projects using the K Means clustering algorithm. K-Means clustering[12] is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. Here an iterative refinement method is employed to find the means(centroids) of the cluster. Any new value entered is first ascertained to be of a particular cluster and then the PSO estimation is performed on it. The parameter thus estimated gives lesser error and closer results than those without using clustering.

2 Model Descriptions

In this model we have considered “The standard PSO with inertia weights” which works on data clustered by using the K Means clustering algorithm. The effort is given by Equation 1, in which a, b and c are parameters to be tuned by PSO.

In PSO, swarm behavior is used for tuning the parameters of the Cost/Effort estimation. As the PSO is a random weighted probabilistic model the previous benchmark data is required to tune the parameters. Based on that data, swarms develop their intelligence and empower themselves to move towards the solution .We initially develop clusters having similar values and then apply the PSO to each cluster individually to

obtain the parameter value. The following is the methodology employed to tune the parameters in the proposed models following it.

3 Methodology

Here we describe the methodology for our model which uses K means clustering algorithm and implements PSO on these clusters.

Input: k the number of clusters, d-data set containing size of software projects, measured efforts, EAF (complexity factor).

Output: Set of k clusters, Optimized parameters for the clusters.

Step 1: Choose k objects from d as initial centroids.

Step 2: Assign each object to a cluster based on the minimum distance between the centroid and the value.

Step 3: Update the cluster mean by assigning centroid as the mean of all data values as the new centroid.

Step 4: Repeat the above steps until we get the stable clusters. These clusters obtained are tuned using PSO.

Step 5: *Initialization:* Initialize particles with random positions and velocity vectors of tuning parameters. We also need the range of velocity between $[-V_{max}, V_{max}]$.

Step 6: *Evaluation of Fitness Function:* For each particle position with values of tuning parameters, evaluate the fitness function. The fitness function here is Mean Absolute Relative Error (MARE). The objective in this method is to minimize the MARE by selecting appropriate values from the ranges specified in step 1.

Step 7: *Finding the P_{best} – Personal best:* If fitness (p) better than fitness (P_{best}) then: $P_{best} = p$. Here the P_{best} is determined for each particle by evaluating and comparing measured and estimated effort values of the current and previous parameters values.

Step 8: *Finding the G_{best} (global best):* Set the best of ' P_{best} ' as global best – G_{best} . The particle value for which the variation between the estimated and measured effort is the least is chosen as the G_{best} particle.

Step 9: *Update values:* Update the velocity and positions of the tuning parameters with equations (2) & (3).

Step 10: Repeat steps 2 to step 5 until "particles exhaust".

Step 11: Give the Gbest values as the optimal solution.

Step 12: Stop.

4 Model Analysis

We have implemented the above methodology, for tuning the parameters a, b and c in "C" language. For the parameter 'a, b and c' the velocity and position of the particles are updated by applying the equations (2) & (3), with parameters $w=0.5$, $c1=c2=2.0$. The data is clustered by using K Means. The final allocation matrix defines sets of values in each cluster. We apply PSO individually on each cluster and output a set of parameter values for each cluster.

5 Performance Criterion

We consider three performance criterions which are Variance Accounted –For (VAF), Mean Absolute Relative Error (MARE) and Variance Absolute Relative Error (VARE) [7].

6 Experimental Study

For the study of this model we have taken 45 data values from COCOMO81 dataset, for both training and testing. We have considered 3 clusters (0, 1, 2) the clusters are indicated in Table 1. By running the ‘C’ implementation of the above methodology we have obtained the following parameters for the proposed model.

Cluster 0: a=0.314800; b=1.862067; c= -5.062736

Cluster 1: a=0.079505; b=2.137003; c= -3.657698

Cluster 2: a=4.182602, b=0.963337; c= -1.449817

The Measured and Estimated Efforts Corresponding to Cluster PSO & COCOMO are given in Table 1 and the corresponding graph is depicted in Fig 1. The Performance Criterion is given in Table 2.

Table 1. Measured (ME) and Estimated (EE) Efforts Corresponding to Cluster PSO & COCOMO

Cluster No.	Size	EAF	ME	COCOMO	EE	Cluster No.	Size	EAF	ME	COCOMO	EE
0	16	0.66	33	39	31.22	1	28	0.96	83	102	90.80
0	18	2.38	321	214	157.87	1	30	1.14	87	130	126.33
0	20	2.38	218	243	193.19	1	32	0.82	106	100	103.67
0	24	0.85	79	108	94.36	1	28	0.45	50	47	40.62
0	13	2.81	98	133	99.89	2	4	2.22	43	30	33.85
0	22	1.76	230	201	170.01	2	6.9	0.4	8	9.8	9.30
0	13	2.63	82	161	93.16	2	3	5.86	73	60	69.18
0	12	0.68	55	33	16.82	2	3.9	3.63	61	52	54.88
0	15	0.35	12	20	12.00	2	3.7	2.81	40	38	40.00
0	19.5	0.63	45	46	45.00	2	1.9	1.78	9	10.7	12.37
0	23	0.38	36	33	36.00	2	9.4	2.04	88	89	72.43
0	24	1.52	176	193	172.73	2	2.14	1	7.3	7	7.25
0	15	3.32	237	239	156.79	2	1.98	0.91	5.9	5.9	5.90
0	25	1.09	130	145	132.51	2	6.2	0.39	8	8.4	8.01
0	21	0.87	70	68	74.30	2	2.5	0.96	8	8.1	8.26
1	46	1.17	240	212	328.92	2	5.3	0.25	6	4.7	3.76
1	30	2.39	423	327	268.86	2	10	3.18	122	114	120.79
1	37	1.12	201	238	196.26	2	8.2	1.9	41	55	58.88
1	48	1.16	387	239	357.48	2	5.3	1.15	14	22	22.53
1	50	3.14	1063	962	1063.00	2	4.4	0.93	20	14	14.76
1	40	2.26	605	529	472.89	2	6.3	0.34	18	7.5	6.92
1	34	0.34	47	44	47.00	2	6.7	2.53	57	60	64.67
						2	9.1	1.15	38	42	38.92

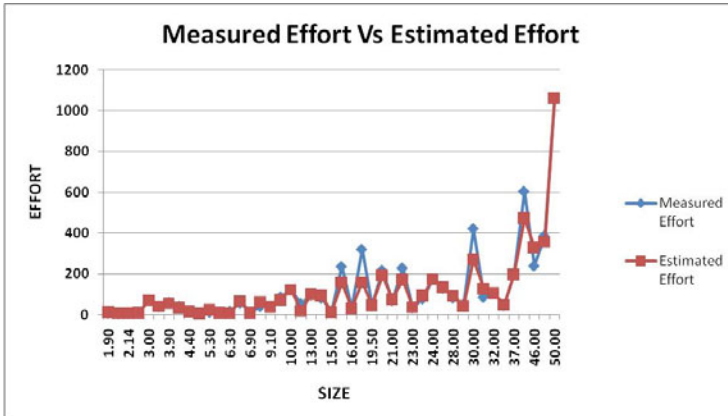


Fig. 1. Measured Effort Vs Estimated Effort

Table 2. Performance Criterion

Model	VAF (%)	MARE (%)	VARE (%)
Cluster-PSO Model	94.6618	17.3564	3.7287
COCOMO Model	95.4722	20.7606	4.1224

7 Conclusion

Software cost estimation is based on a probabilistic model and hence it does not generate exact values. However if good historical data is provided and a systematic technique is employed we can generate better results. In this study we have proposed a new model to estimate the software effort. In order to cluster the values into similar groups we have used K-means clustering algorithm and in order to tune the parameters we use particle swarm optimization methodology algorithm. It is observed that the clustered-PSO gives more accurate results when juxtaposed with the standard COCOMO. On testing the performance of the model in terms of the error rate the results were found to be useful. This method can be applied to large datasets to generate efficient values. These techniques can be applied to other software effort models.

References

1. Bailey, J.w., Basili, V.R.: A meta model for software development resource expenditures. In: Fifth International conference on software Engineering, pp. 107–129. IEEE, Los Alamitos (1981), CH-1627-9/81/0000/0107500.75@
2. Briand, L.C., El Emam, K., Bomarius, F.: COBRA: A Hybrid Method for Software Cost Estimation, Benchmarking, and Risk Assessment, International Software Engineering Research Network Technical Report ISERN-97-24 (Revision 2), PP 1-24 (1997)
3. Gruschke, T.: Empirical Studies of Software Cost Estimation: Training of Effort Estimation Uncertainty Assessment Skills. In: 11th IEEE International Software Metrics Symposium (METRICS 2005), 1530-1435/05. IEEE, Los Alamitos (2005)

4. Jørgensen, M.: Evidence-Based Guidelines for Assessment of Software Development Cost Uncertainty. *IEEE Transactions on Software Engineering* 31(11), 942–954 (2005)
5. Sheta, A.F.: Estimation of the COCOMO Model Parameters Using Genetic Algorithms for NASA Software Projects. *Journal of Computer Science* 2(2), 118–123 (2006)
6. Auer, M., Trendowicz, A., Graser, B., Haunschmid, E., Biffli, S.: Optimal Project Feature Weights in Analogy-Based Cost Estimation: Improvement and Limitations. *IEEE Transactions on Software Engineering* 32(2), 83–92 (2006)
7. Hari, C.V.M.K., Prasad Reddy, P.V.G.D., Jagadeesh, M.: Interval Type 2 Fuzzy Logic for Software Cost Estimation Using Takagi-Sugeno Fuzzy Controller. In: *Proceedings of 2010 International Conference on Advances in Communication, Network, and Computing*. IEEE, Los Alamitos (2010), doi:10.1109/CNC.2010.14, 978-0-7695-4209-6/10
8. Jørgensen, M., Shepperd, M.: A Systematic Review of Software Development Cost Estimation Studies. *IEEE Transactions on Software Engineering* 33(1), 33–53 (2007)
9. Poli, R., Kennedy, J., Blackwell, T.: Particle swarm optimization An overview. In: *Swarm Intell.*, pp. 33–57. Springer, Heidelberg (2007), doi:10.1007/s11721-007-0002-0
10. Felix, T.S.C., Tiwari, M.K.: *Swarm Intelligence: Focus on Ant and Particle Swarm Optimization*, pp. 1–548. I-TECH Education and Publishing (2007), ISBN 978-3-902613-09-7
11. Keung, J.W., Kitchenham, B.A., Jeffery, D.R.: Analogy-X: Providing Statistical Inference to Analogy-Based Software Cost Estimation. *IEEE Transactions on Software Engineering* 34(4), 471–484 (2008)
12. Bin, W., Yi, Z., Shaohui, L., Zhonghi, S.: CSIM: A Document Clustering Algorithm Based on Swarm Intelligence, pp. 477–482. IEEE, Los Alamitos (2002), 0-7803-7282-4/02@2002

Controlling Crossover Probability in Case of a Genetic Algorithm

Parama Bagchi¹ and Shantanu Pal²

¹ Department of Computer Science and Engineering,
Dream Institute of Technology,
Kolkata, India

² AKC School of Information Technology,
University of Calcutta,
Kolkata, India

{paramabagchi, shantanu.smit}@gmail.com

Abstract. Genetic Algorithms (GAs) are commonly used today worldwide. Various observations have been theorized about genetic algorithms regarding the mutation probability and the population size. Basically these are the search heuristics that mimic the process of natural evolution. This heuristic is routinely used to generate useful solutions for optimization and search problems. GAs belong to the larger class of evolutionary algorithms (EAs), which generate solutions to maximize problem solving by using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In this paper we study of a simple heuristic in order to control the crossover probability of a GA. We will also explain how stress factors in on the crossover probability and why it is an important phenomenon in case of a GA and how it can be controlled effectively. Experimental results show that, for reaching lower probability from higher probability, we can get faster optimal solutions for any problem. These experimental values are derived by taking the values at the high probability and then slowly yet steadily decreasing them.

Keywords: genetic algorithm, crossover probability, heuristic, evolutionary algorithms.

1 Introduction

GAs is generally portrayed as a search procedure which can optimize functions based on a limited sample of function values. A function based on a minimal spanning tree of data points can be used for clustering and GAs in an attempt to optimize the specified objective function in order to detect natural grouping in a given data set [1]. The method has also found to provide good results for various real life data sets [2]. Unlike conventional search methods GAs deals with multiple solutions simultaneously and computes the fitness function values for these solutions [2]. GAs was inspired by the way living things evolved into more successful organisms in nature [3]. In this paper we work on the general ideas of the GAs and the experimental values gained from the input values regarding the proposed field. We do not try to define any

re modification of traditional GAs, rather we shows that taking values from higher probability to lower probabilities will make a successive change to reach the optimal solution of any problems. This is intern the basic motivation if this paper. The rest of the paper is organized as follows: Section 2 provides an understanding of the simulation and natural selection. Section 3 recalls the enlisted crossover techniques. In section 4 we extend our experimental results which is the motivate aim of our work. Finally section 5 draws our conclusions.

2 Simulation and Natural Selection

To simulate the process of natural selection in a computer, representation of an individual and fitness function is essential to determine. They are as follows:

2.1 Representation of an Individual

At each point during the search process in [4], *Filho et all* maintains a "generation" of "individuals." Each individual is a data structure representing the "genetic structure" of a possible solution. In GAs, the alphabet $\{0, 1\}$ is usually used for 2 clusters. This string is interpreted as a solution to the problem in [4] trying to solve, if we want to find the optimal quantity of the three major components, we can use the alphabet $\{1, 2, 3 \dots, 9\}$ denoting the number of each ingredient.

2.2 Fitness Function

Given an individual, we must assess how good a solution it is so that we can rank individuals. A possible function is $\text{Fitness}(i) = (\text{sum of edge weight of MSTs}) / (\text{Total no of points in each cluster})$. This function has a value between 0 and 1 and is monotonically increasing [5].

3 The Crossover Technique Enlisted

The crossover technique [6] is one of the most fundamental operations in the evaluation of GA. Generally if there is a medium - large population a great responsibility is own to this crossover operator. This is primarily because of the fact that in comparison to the mutation operator, the crossover operator owes and brings about a great change in the structure of the chromosome which is finally responsible for the convergence of the GA. Normally in case of a very small population; crossover might not have a very significant importance as compared to mutation. The test so as to ensure that crossover works best when it is kept at a high range of probability values was done.

4 Experimental Results

As the main aim was to show the result generated by the experiment which reach the optimal solution quicker. It is also cleared that we have started the experiment with a

high probability and gradually decreased it to a lower probability region, which indeed reflect the quicker optimal solution for a problem. The following observations were seen by conducting the underneath experiment. Here we have taken clusters consisting of a number of points. For the points correct clustering is found when slowly the crossover probability is decreased from high to low. Excellent results are found if the change between a higher probability and a lower probability is decreased by an amount of 0.1. Decreasing the crossover probability, slowly and step by step would help us to actually see the increasing change in the nature of chromosomes, which finally help us to achieve convergence. By decreasing the crossover probability in this fashion, the sample string changed slowly, till we got the final convergence. In contrast to this when the crossover probability was kept at a low range say [0.5], and slowly increased, we did not time taken to converge was longer. The following table gives a snapshot of the discussions made above.

Table 1. Crossover Probability by the Experiment

Size of Data	Number of Iterations	Number of Points Correctly Clustered	Percentage (%) of Correct Classification
20 Points	700	12	40% (Crossover \rightarrow 0.6)
	800	11	30%
	900	11	30%
	5000	11	45% (Crossover \rightarrow 0.8)
	6000	11	45%
	7000	20	100%
	30 Points	900	21
1000	18	40%	
2000	30	100%	
9000	21	30% (Crossover \rightarrow 0.8)	
10000	21	30%	
15000	21	30%	

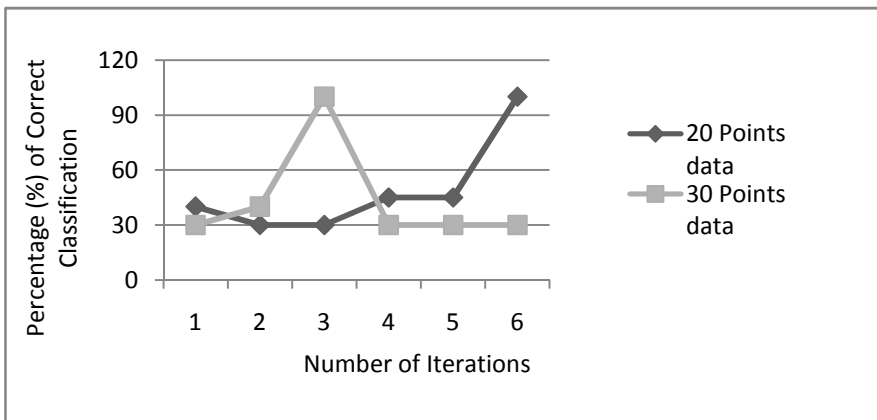


Fig. 1. Percentage of Correct Classification for 20 and 30 Points Data

The above experimental values were actually derived by taking the values at a high probability and then slowly made to decrease steadily. The final experimental results show that the probabilities have been taken starting from a high value to steadily low values. The graph below shows the results.

5 Conclusion

If the conception of a computer algorithms being based on the evolution of organism is surprising, the extensiveness with which this algorithms is applied in so many areas is no less than astonishing. These applications, be they commercial, educational and scientific, are increasingly dependent on this algorithms, the GAs. Its usefulness and gracefulness of solving problems has made it a more favorite choice among the traditional methods, namely gradient search, random search and others. In conclusion, it can as well be said that by increasing the no of iterations it is possible to get a better set of results. The size of the population should be increased to see the optimal solution and to be able to deduce the fittest chromosome for a large population. The stopping criterion for genetic algorithms is still not clear. A significant work can also be done in that region.

References

1. Togelius, J., Preuss, M., Yannakakis, G.N.: Towards Multiobjective Procedural Map Generation. In: PCGames Monterey, CA, USA, June 18 (2010)
2. Chowdhury, N., Jana, P.: Finding the Natural Groupings in a Data Set Using Genetic Algorithms. In: Manandhar, S., Austin, J., Desai, U., Oyanagi, Y., Talukder, A.K. (eds.) AACC 2004. LNCS, vol. 3285, pp. 26–33. Springer, Heidelberg (2004)
3. Mathew, T V.: Genetic Algorithm,
http://www.civil.iitb.ac.in/tvm/2701_dga/2701-ga-notes/gadoc/gadoc.html (last access November 2009)
4. Filho, J.R., Alippi, C., Treleaven, P.: Genetic Algorithm Programming Environment's. ACM Journal Computer 27(6) (June 1994), doi:10.1109/2.294850
5. Rokach, L., Maimon, O.: Clustering Methods. In: Data Mining and Knowledge Discovery Handbook, vol. 3, pp. 321–325 (2005), doi:10.1007/0-387-25465-X_15
6. Genetic Algorithms: Principles of Natural Selection Applied to Computation. Science, 261, 872-878

A Qualitative Survey on Unicast Routing Algorithms in Delay Tolerant Networks

Sushovan Patra¹, Anerudh Balaji¹, Sujoy Saha²,
Amartya Mukherjee¹, and Subrata Nandi¹

¹ Department of Computer Science and Engg., ² Department of Computer Application National
Institute of Technology, Durgapur, 713209, India
{bubususpatra, anerudhbalaji, sujoy.ju, mamartyacsel,
subrata.nandi}@gmail.com

Abstract. DTN is an emerging research area that takes a different approach to (inter)networking and allows working in stressed as well as in highly heterogeneous environments. DTN features a number of unique properties which make this concept applicable to challenging environments in which traditional communication paradigms would fail or perform poorly. In DTN intermediate nodes takes the responsibility to store-and-forward delivery as well as physical data carriage using deterministic and/or probabilistic routing. In this paper we present a comprehensive up to date survey of different routing protocols as well as perform qualitative comparison of different routing strategies with respect to important issues in DTN. Further we highlight some of the upcoming issues related to design of DTN routing strategy.

1 Introduction

In wired and wireless network a large amount of wire-less bandwidth capacity remains unused because the current communications paradigm (i.e. the Internet) has not been designed to make use of local and intermittent connectivity. DTN was an attempt to extend the reach of the network which enables communication between ‘Challenged’ networks which is a technical by-product of absence of end to end connection and unpredictable node mobility. DTN has greater application in *disaster management* [11] but as the cost of the handle devices like laptops, mobiles phones, pager has decreased, it can be introduced in wide applications of modern technologies.

Routing Protocols in Delay Tolerant Networks classifies the routing family in two categories - Forwarding based and Flooding based and compares the protocols proposed in the flooding families and the forwarding families on the basis of resource usage, effectiveness amongst other factors [8][15][1]. Previous survey papers do not consider the latest works on Delay tolerant network routing and are updated till the year 2006 only this has motivated us to come up with this survey that encompasses the most recent work. Our paper also provides a different perspective on sub classification of routing techniques in Delay Tolerant Networks. The paper is divided into three sections. In the following section we present the Taxonomy and produce the summary in tabular form (Table 1) and section 3 presents the Conclusion and future directions. The figure below represents our proposed taxonomy.

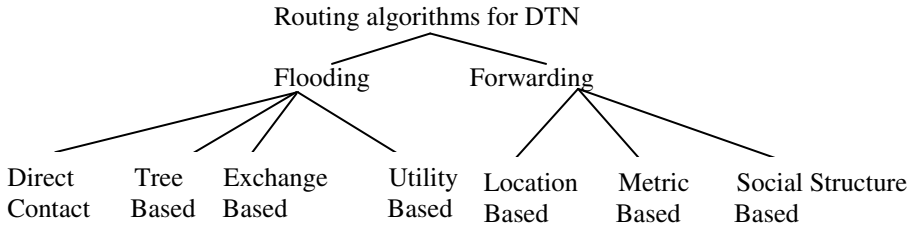


Fig. 1. Classification of DTN routing algorithms based on a new taxonomy

2 Proposed Taxonomy of Dtn Routing Strategies

The routing strategy for Delay Tolerant networks can be categorized based on two properties namely Replication (flooding) and Knowledge (forwarding) [1].

2.1 Flooding Based

The *relays* [1] in DTN store the messages until they connect with the destination, at which point the message is delivered.

2.1.1 Direct Contact

Waits until source comes into contact with the destination before forwarding the data but does not necessarily consider shortest path as it does not use knowledge of the network [1][8]. It is assumed to be *degenerated* case of forwarding as it does only use single copy of the message.

2.1.2 Tree Based Flooding

Tree Based Flooding distributes the task of making copies to the relayed node; with an indication for the number of copies required [1]. The concept is quite similar to diffusion computation algorithm. *Two hop relay* is a special case of Tree based flooding since this strategy is a case of tree based flooding with depth 1.

2.1.3 Exchange Based

This category includes those routing strategies in which exchange of information plays a pivotal part in the procedure. Three algorithms – PROPHET [2], Epidemic [1] and Practical Routing [10] fall under this category. *Epidemic routing* [8] maintains a *Summary Vector* that will keep track of the entire message IDs. Two nodes send each other the list of all the messages IDs they have in their buffers on contact and exchange the messages they do not have using IDs. Extreme flooding leads to huge amount of resources consumption [1], [8]. *PROPHET*[2] is extension of epidemic routing which along with *Summary Vector* also contain the delivery predictability information stored at the nodes to update the internal delivery predictability vector [2] [8]. Extreme flooding of previous one is here controlled by *threshold* value. *Practical Routing* [10] uses only observed information about the network. A metric is designed that estimates how long a message will have to wait before it can be transferred to the

next hop. Messages are exchanged if the topology suggests that a connected node is “closer” (has better metric) than the current node.

2.1.4 Utility Based Routing

Flooding leads to network congestion and drastic fall in network throughput, which forms the basis for utility based routing schemes. In strategies that fall under utility based routing, the utility of the message is considered while making routing decisions. *ORWAR* [13] proposes a multi copy routing scheme, using controlled replication and a fixed number of copies distributed over network. At each contact the node tries to forward half of the message copies, keeping the rest for itself, messages with best utility per bit ratio selected first and forwarded. *RAPID* [14] optimizes a specific routing metric such as worst case delivery delay or the fraction of packets that are delivered within a deadline. It models DTN routing as a utility driven resource allocation problem.

In Direct contact [1] [8] delivery ratio is dependent on the proximity of the destination with respect to the source. Tree based flooding [1] supports multi path routing of the messages. Epidemic routing promises a high delivery ratio and exhibits average effectiveness as it consumes much more buffer space with compare to any other routing. Utility based routing have the highest delivery ratio while creating a far lesser overhead. A qualitative comparison as per the above mentioned category is illustrated in the table 1.

2.2 Forwarding Based

It uses the network topology information to select the best path, and the message is then forwarded from node to node along this path.

2.2.1 Location Based Routing

This forwarding approach requires the least information about the network and assigns coordinates (can be GPS) to each of the nodes. A distance function is used to estimate the *cost* of delivering messages from one node to another. In general a message is forwarded to a potential next hop if that node is closer in than the current custodian with respect to destination node [1].DTN routing in *mobility space* [15] defines a generic routing scheme for DTNs using a high-dimensional Euclidean space constructed upon nodes' mobility patterns, and assigns coordinates to the nodes in this space based upon their probability of being found in each possible location.

2.2.2 Metric Based Routing

In this approach, a *weight* is assigned to each link based on its suitability in delivering messages to a given destination as in *Gradient Routing*[9].Each node must store a metric for all potential destinations. When the custodian of a message contacts another node that has a better metric for the message's destination, it passes the message to it [1], [9]. *Friendship based routing* [7] finds a metric that reflects the node relations more accurately by considering the following three behavioral features of close friendship: high frequency, longevity, regularity. These properties addressed in new metric called Social Pressure Metric, SPM which gives the information regarding delay for message passing. Then information in SPM is use to choose best forwarder.

2.2.3 Social Structure Based Routing

This class of routing algorithms discovers the heterogeneity of human interactions such as *community* formation from real world human *mobility traces*.

Predictability of the node mobility motivates the idea of *BUBBLE Rap* [3]. Each node has a global ranking across the system and a local ranking for its local community and may belong to multiple communities with multiple local rankings. The algorithm detects the community using K-clique by *Palla et al.* and *weighted network analysis* (WNA) by *Newman*. The forwarding is done based on popularity of the node.

Table 1. Summary of flooding and forwarding based routing schemes. Sub classification has been done according to the propose taxonomy (fig 1).

Routing Family	Routing Categories	Motivation	Remarks
Flooding	Direct Contact [1]	The need for a simple and somewhat effective model of flooding.	1. Simple and does not consume much resource. 2. Works only if source contacts destination.
	Tree based [1]	The need to reduce the number of copies in the network and reduce network congestion	1. Can deliver messages to destinations that are multiple hops away. 2. Tuning of parameters is a major hindrance.
	Utility based [13], [14]	The need to optimize specific routing metric such as average delivery delay.	1. It does minimize retransmissions. 2. Effective use of resources at system level.
	Exchange based [1],[2],[10]	It was originally proposed so that synchronization of the databases is possible	1. Huge amount of resources-buffer space, bandwidth are used. 2. No knowledge of network required. 3. High delivery ratio.
Forwarding	Location based [1],[15]	Run time determination of destination location to avoid router maintenance cost.	1. Less knowledge requirement 2. Location does not necessarily correspond to network topology. 3. Change in network topology complicates routing.
	Metric based [1], [7], [9]	Using information's like last contact with destination, battery energy, mobility in from of metrics.	1. Requires more knowledge compare to Location based. 2. Achieve higher Delivery Probability than Direct contact 3. Fail to balance load.
	Social structure based [3], [4], [5], [6]	Mobility traces of the real world have been used and heterogenous social relationship in term of group and individual has been identified for first time.	1. Delivery ratio achieved is better than PROPHET [3]. 2. Requires a setup period. 3. Latency is reduced by 20% compared to PROPHET

Incompetent assumption of the network graph is fully connected in a social network motivates the idea of SimBet which is based on the concept of Ego network. A node forwards messages to a node with higher Social Similarity [4] or Higher Between's [4]. The major disadvantage of SimBet is that it models social relationship in a binary form, which is not realistic and leads to inconsistencies which lead to SimBetAge [5]. It improves on similarity estimation and calculation of between's, by using weighted time dependant graph and exponential decay function [5]. Fair routing [6] attempts to overcome the existing problems of security and load balancing. Interaction Strength is an indicator of the likelihood of a contact to be sustained over time Nodes will only accept forward request from those nodes of equal or higher status (*popularity*) i.e. higher Queue length which may lead to starvation for new node in the network.

Location based routing offers very little flexibility and have scalability issues but consume minimal resources. Metric based routing offers comparatively lesser delay and a better delivery ratio. Social Structure based routing, achieves a better delivery ratio than the other routing categories and also addresses load balancing and security issues as in Fair Routing [6]. A qualitative illustration is presented in the Table 1.

3 Conclusions and Future Directions

In spite of the continuous research in DTN in last few years, none of the works presented here pursue dynamic allocation of buffering spaced based on applications. Other than Fair routing [6], none of the papers surveyed here considers the issue of security, which is of paramount importance while using Delay tolerant networks for applications such as Social Networks. The techniques proposed for load balancing such as Message splitting [12] and Fair Routing [6] are not convincing. Even though they achieve significant improvement in terms of balancing, the overall performance of the routing protocol is comparable to that of PRoPHET [2]. The recent improvement in research arena has shown the glimpse that DTN is quite sustainable in large network structure. So we have identified the issues *security*, *flexibility*, *scalability*, and *mobility support*, *load balancing* require to be addressed so the DTN can be used as a platform to host upcoming application like Social Network, bulk data transfer, multimedia data transfer and streaming in Sparse Network.

References

1. Jones Paul, E.P.C., Ward, A.S.: Routing Strategies for Delay Tolerant Networks. Submitted to Computer Communication Review (2008)
2. Lindgren, A., Doria, A., Schelen, O.: Probabilistic Routing in Intermittently Connected Networks, vol. 3126, pp. 239–254 (2004)
3. Hui, P., Crowcroft, J., Yoneki, E.: BUBBLE Rap: Social-based Forwarding in Delay Tolerant Networks. In: MobiHoc, Hong Kong SAR, China, May 26–30 (2008)
4. Daly, E., Haahr, M.: Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs. In: MobiHoc, Montreal, Canada, September 9-14 (2007)
5. Link, J.Á.B., Viol, N., Goliath, A., Wehrle, K.: SimBetAge: Utilizing Temporal Changes in Social Networks for Pocket Switched Networks, U-NET, Rome, Italy, December 1 (2009)

6. Pujol, J.M., Toledo, A.L., Rodriguez, P.: Fair Routing in Delay Tolerant Networks. In: Proc. of IEEE Infocom Rio de Janeiro, Brazil, April 19-25 (2009)
7. Bulut, E., Szymanski, B.K.: Friendship Based Routing in Delay Tolerant Mobile Social Networks. In: IEEE GLOBECOM Miami Florida USA (December 2010)
8. Shen, J., Moh, S., Chung, I.: Routing Protocol in Delay Tolerant Network: A Comparative Survey. In: The 23rd International Technical Conference on Circuits/System Computer and Communication (2008)
9. Poor, R. D.: Gradient routing in ad hoc networks. MIT Media Laboratory (2000), <http://www.media.mit.edu/pia/Research/ESP/texts/poorieepaper.pdf> (unpublished manuscript)
10. Jones, E.P.C., Li, L., Ward, P.A.S.: Practical routing in delay-tolerant networks. In: Proc of the ACM SIGCOMM workshop on Delay-tolerant networking (2005)
11. Uddin, Y.S., Nicol, D.M., Abdelzaher, T.F., Kravets, R.H.: A Post-Disaster Mobility Model for Delay Tolerant Networking. In: Proc. Winter Simulation Conference, Austin, Texas (December 2009) (invited)
12. Jain, S., Fall, K., Patra, R.: Routing in a Delay Tolerant Network. In: SIGCOMM, August 30-September 3 (2004)
13. Sandulescu, G., Tehrani, S.: Opportunistic DTN Routing with window aware adaptive replication. In: Proceedings of the ACM 4th Asian Conference on Internet Engineering (AIN-TEC), Bangkok, Thailand (2008)
14. Balasubramanian, A., Levine, B., Venkatramani, A.: Replication Routing in DTN: A Resource allocation approach. In: SIGCOMM, Kyoto, Japan, August 27-31 (2007)
15. Leguay, J., Friedman, T., Conan, V.: DTN Routing in a mobility pattern space. In: ACM WDTN, pp. 276–283 (2005)

Designing and Modeling of CMOS Low Noise Amplifier Using a Composite MOSFET Model Working at Millimeter-Wave Band

Adhira Raj¹, Balamurugan Karthigha², and M. Jayakumar³

¹M. Tech, VLSI Design, Amrita Vishwa Vidyapeetham, Coimbatore, India
adhira.raj@gmail.com

² Assistant Professor, Amrita Vishwa Vidyapeetham, Coimbatore, India
b_karthigha@cb.amrita.edu

³ Associate Professor, Amrita Vishwa Vidyapeetham, Coimbatore, India
m_jayakumar@cb.amrita.edu

Abstract. In this paper, MOSFET modeling for millimeter wave integrated circuits is discussed. High frequency MOSFET is built using BSIM3v3 as intrinsic core and the parasitics due to HF are designed as extrinsic subcircuit. The proposed methodology is then used in designing a low power, mm-wave CMOS low noise amplifier. The operation of the circuit is simulated using a circuit simulator. The wideband characteristics are verified by implementing the LNA circuit with and without composite model.

Keywords: CMOS millimeter-wave integrated circuits, LNA-low noise amplifier, high frequency (HF) behaviour, composite model, Bsim3v3.

1 Introduction

Millimeter wave WLAN is getting popular due to the enhancement band of 7GHz in the frequency of 53-67 GHz. The major advantage of this band is its high oxygen attenuation making it interface free for short range indoor WLAN application. This requires realization of semiconductor technology band devices operating at millimeter waves. Also CMOS implementation promises higher level of integration, low cost and low power consumption. The advanced performance of MOSFET could be used for high frequency circuit design in view of a system-on-a-chip realization. To acquire an efficient design environment, accurate MOSFET model working at millimeter wave frequencies is essential.

An approach to model a MOSFET at HF is to build a subcircuit based on the intrinsic MOSFET that has been modelled for low-frequency analog applications. With the added parasitic components at the gate, at the drain, at the source and at the substrate, these models could be made to work at millimeter wave frequencies. Once the MOSFET model is designed, then the working could be examined by using the model to construct LNA circuit.

In this paper the section 2 describes about the concepts of MOSFET modeling at HF and designing of composite MOSFET model. Section 3 addresses the design

parameters of LNA and issues in selecting a multistage low noise amplifier. To demonstrate the effectiveness of the proposed MOSFET model, Section 4 describes the implementation of the composite model in LNA and finally section 5 illustrates the simulation results obtained from MICRCAP 10 circuit simulator.

2 MOSFET Modelling

2.1 AC Small-Signal Modeling

In this, a four-terminal MOSFET can be divided into two portions: intrinsic part and extrinsic part. The concepts of equivalent circuits representing both intrinsic and extrinsic components in a MOSFET are analyzed to obtain a physics-based RF model.

2.2 Modeling of the Intrinsic MOSFET

A BSIM3v3 model as the intrinsic component is selected. BSIM3v3 has been widely accepted as a standard CMOS model. This model is mainly selected as it is found to be more advantageous, as the Drain current could easily be modified by current equation and this model mainly comprises of an accurate capacitance model.

2.3 Subcircuit Model

As we discussed above, a MOSFET contains many extrinsic components such as gate resistance, source/drain series resistance, substrate resistance and capacitance, and gate overlap capacitance.

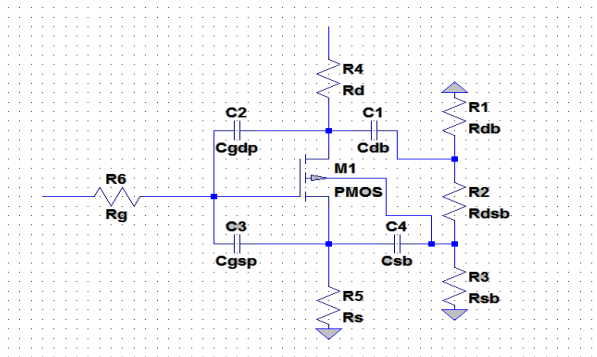


Fig. 1. Composite MOS Model

3 LNA Architecture

3.1 Multistage LNA Architecture

For many applications such as the automatic gain control in WLAN receivers with single signal path, variable gain of amplifiers is needed. In this paper, we propose a

millimeter wave CMOS multistage LNA. A multistage LNA is selected over a single stage LNA. LNA is a critical block in radar systems. The resistive loss is to be minimized thus using interconnection lines and the inductors. The LNA consists of four cascaded common-source stages. In the resistive-feedback second stage, the peaking inductor and the gate-source capacitance of transistor M2 were series resonance at the geometric mean frequency. LNA based on inductive-source cascade. This proposed architecture provides gate-to-source and source-to-body parasitic capacitances in two stage gain cascaded architecture.

4 Experimental Results

The composite MOSFET model suitable for high frequency IC is done. Using the composite MOSFET model, a multistage LNA is constructed and simulated using circuit simulator MICROCAP-10. The simulated results depicts that the gain from 4.5 dB increases to 12 dB with the inclusion of the composite MOS model. Similarly the composite MOS model is included in a common source amplifier whose gain from 20 dB increases to 40 dB. The amplifier without subcircuit is shown in figure 2 and amplifier with subcircuit is shown in figure 3.

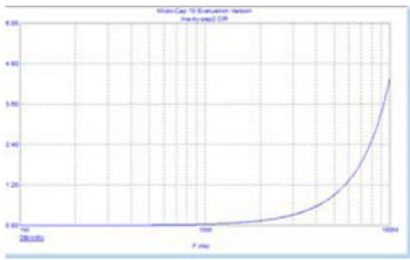


Fig. 2. Amplifier gain without subcircuit

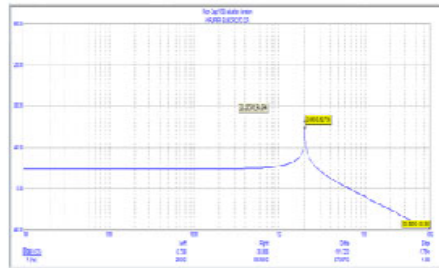


Fig. 3. Amplifier gain with subcircuit

The comparison of multistage and single stage low noise amplifier is done. It is found that a increase in gain is found when using a multistage circuitry. The Table 1 shows the comparative study results of different low noise amplifier design with subcircuit and without subcircuit. The comparison considering single stage and multistage is also shown and thus showing that the gain with multistage is more efficient than single stage low noise amplifier.

Table 1. Comparative Study on Lna

Parameters	Multi -stage LNA	Multi -stage LNA	Single -stage LNA
CMOS technology	.18µm	.13µm	.13µm
GAIN(dB)without composite MOS model	11.9 (at 53GHz)	6.03 (at100 MHz)	6.04 (at15 GHz)
GAIN(dB)with composite MOS model	19.9 (at60 GHz)	12.9 (at60 GHz)	6.5 (at60 GHz)

5 Conclusion

The modelling of both intrinsic and parasitic components in MOSFETs is crucial to describe the HF behaviour of MOS devices operated at GHz frequencies. The brief discussion of the MOSFET modeling at millimetre wave frequency is done. The LNA design consisting of multistage is found to have better gain and with the inclusion of the designed composite MOS model it is found that the circuit gives better performance in the aspect of gain. Thus a high frequency MOSFET model for low noise amplifier is designed and discussed. This work can be extended to study the distortion behaviour of MOSFET working at millimetre wave frequencies.

References

1. Rezaul Hasan, S.M.: Analysis and Design of a Multistage CMOS Band-Pass Low-Noise Preamplifier for Ultrawideband RF Receiver. *IEEE Transactions on Very Large Scale Integration (vlsi) Systems*, TVLSI.2009.2014166 (2009)
2. Choi, W., Jung, G., Kim, J., Kwon, Y.: Scalable Small-Signal Modeling of RF CMOS FET based on 3-D EM-Based Extraction of Parasitic Effects and its application to Millimeter-Wave Amplifier Design. *IEEE Transaction on Microwave Theory and Techniques* 57(12) (2009)
3. Cheng, Y., Jamal Deen, M., Chen, C.-H.: MOSFET Modeling for RF IC Design. *IEEE Transaction Electronic Devices*, TED.2005.850656 (2005)
4. Shigematsu, H., Hirose, T., Brewer, F., Rodwell, M.: Millimeter-wave CMOS circuit Design. *IEEE Tran. Microwave Theory Tech.* 53(2) (2005)
5. Yang, M.T., Ho, P.P.C., Wang, Y.J., Yeh, T.J., Chi, Y.T.: Broadband Small-Signal Model and Parameter Extraction For deep sub-micron MOSFETs valid up to 110GHz. In: *IEEE Radio Frequency Integrated Circuits Symposium* (2003)
6. Ou, J.J., et al.: CMOS RF modelling for GHz communication ICs. In: *Symposium on VLSI Technology Digest of Technical Papers*, pp. 94–95 (1998)
7. Zirath, H., Ferndahl, M., Motlagh, B.M., Masud, A., Angelov, I., Vikes, H.O.: CMOS devices and circuits for microwave and millimeter wave applications. In: *Proceedings of Asia-Pacific Asia Microwave conference* (2006)
8. Shahriar Rashid, S.M., Ali, S.N., Roy, A., Rashid, A.B.M.H.: A 36.1GHz Single stage Low Noise Amplifier using 0.13 μ m CMOS Process. *World Congress on Computer and Information* (2009)
9. Yu, Y.-H., Yang, Y.-S., Chen, Y.-J.E.: A Compact Wideband CMOS Low Noise Amplifier With Gain Flatness Enhancement. *IEEE journal of solid-state circuits* 45(3) (2010)

Designing Dependable Business Intelligence Solutions Using Agile Web Services Mining Architectures

A.V. Krishna Prasad¹, S. Ramakrishna², B. Padmaja Rani³,
M. Upendra Kumar⁴, and D. Shravani⁵

¹ Research Scholar Computer Science S.V. University Tirupathi A.P. India
kpvambati@gmail.com

² Professor Computer Science S.V. University Campus Tirupathi A.P. India
drsramakrishna@yahoo.com

³ Associate Professor CSE JNTU CEH Hyderabad A.P. India
padmaja_jntuh@yahoo.co.in

⁴ Research Scholar JNTU Hyderabad A.P. India
uppi_shravani@rediffmail.com

⁵ Research Scholar Rayalaseema University Kurnool A.P. India
sravani.mummadi@yahoo.co.in

Abstract. Next generation Business Intelligence web application development uses integrated and intensified technologies like Web 2.0 architectures, Agile Modeling, and Service-orientation (or Web Services). Applying Web Services Mining strategies to Agile Modeled Web architectures will eventually provide valuable insights to Business Intelligence users, Operational Business Decision makers, and more importantly Web application architects. These insights are important in maintenance of these developed applications and also in their scalability purposes. Our research focuses on applying Mining for Software (or Web) Engineering for designing dependable solutions for these integrated technologies, which will eventually improve the Web Engineering process in terms of architecture, its security, requirements etc. In this paper, we discuss about our Mining approach for Business Intelligence to improve insights of Web Engineering applications. We validate our approach with a suitable exemplar.

Keywords: Web Mining, Business Intelligence, Agile Modeling, Web 2.0 architectures, Web Services Mining.

1 Introduction to Next Generation Business Intelligence Web Applications

Now a days, most of the Business Intelligence applications are developed as Web based applications with little Web Engineering principles used in developing them. Next generation Business Intelligence applications development are using an integration of specific technologies like Web 2.0 architectures, Agile modeling, and Service-orientation (or Web Services).[1] Our research is based on the premise that, Applying Web Services Mining approaches to Agile Modeled Web 2.0 architectures will eventually provide valuable insights to Business Intelligence users, Decision makers, and

importantly Web application architects. Our research focuses on designing technical solutions for these integrated technologies, which will eventually improve the Web Engineering process in terms of architectures security requirements. When all these different technologies are integrated together for the development of Business Intelligence web based applications, it creates many research challenges pertaining to insights of decision making regarding architectures of the developed application, its inherent security of that architecture or its requirements etc. These insights are required for maintenance of this application or in its future scalability issues. Moreover, Business Intelligence has to be shifted from enhancing the data warehousing and data mining techniques such as OLAP (Online Analytical Processing), OLAM (Online Analytical Mining), multi dimensional modeling, design methodologies, optimization, indexing and clustering techniques, to how to securely protect these knowledge capitals from being tampered with by unauthorized use. [9]

Agile Modeling. Agile modeling embraces change as a part of the software development process.[2] In most approaches, change is usually considered a bad word. Agile developers work in pairs, create many prototypes of their solutions, and incorporate user's feedback throughout the entire process. Agile software development has encouraged developers to tailor their methods to meet their specific needs. Agile modeling using Unified Modeling Language is geared towards small development projects with tight deadlines, like building Web front ends.

Web 2.0 Architectures. The relationship between Web 2.0 design patterns, models, and architecture artifacts are based on Web 2.0 technologies like search engine optimization, web services, wikis etc. to name a few for our consideration. Models guides Reference architectures and finally specialized architectures refines reference architectures, accounts for domain specific requirements, and also enables solution patterns. [3]

Web Services Architectures. Service Oriented Architectures (and their implementations Web Services) uses a series of independent services that can communicate business logic with one another. These services can be used independently or together to form business platforms that come together to form business platforms that come together and provides value.

2 Mining Approaches for These Business Intelligence Applications

Mining Agile Architectures. Agile software development methods are used to build secure systems. There are different methods defined in agile development as extreme programming (XP), scrum, feature driven development (FDD), test driven development (TDD), etc. Agile processing includes the phases as agile analysis, agile design and agile testing. These phases are defined in layers of Model Driven Architecture (MDA) to provide security at the modeling level which ensures that "security at the system architecture stage will improve the requirements for that system".

Mining Web 2.0 architectures. Traditionally, to mine web 2.0 architectures in general, we use the methodology for mining patterns from examples, hence capture the knowledge, and then construct models and architecture based on the commonalities in

the patterns. Design patterns are micro architectures that have proved to be reliable, easy to implement and robust. Three types of Design Patterns i.e. Creational Design Patterns, Structural Design Patterns and Behavioral Design Patterns. Design Patterns are described by listing the intents, motivations, applicability, structure (UML diagrams), participants, collaborations, consequences, implementation details are known as related patterns. The structure of the patterns is represented in Graphs or Matrices (Abstract Class Matrix, Generalization Matrix, Association Matrix etc). The pattern descriptions are easy to modify to suit the needs of users by using Design Pattern Markup Language. Reality Mining is a mashup pattern in terms of MIT which states that, it is the collection of machine-sensed environmental data pertaining to human social behavior.

Web Services Mining Architectures. Web Service mining is a search process aiming at the discovery of interesting and useful compositions of existing web services. Recall is the fraction of relevant services in the collection that were returned by the system and precision is the fraction of the returned results that are relevant. The Business Process Execution Language (BPEL) deals with Web service composition and attempts to solve the problem of composing a number of web services into a business process. In Business Process Query, there is a need to interact with web services. Web services interaction mining provides three levels of abstraction that represent three complementary web services. The levels are Web service Operations level, Web service Interactions level and web service work flow level.

Designing Dependable Solutions for Business Intelligence. Software Engineering problems must be treated by both theoretical and empirical methodologies. The former is characterized by abstract, inductive, mathematics-based, and formal-inference centered studies; while the latter is characterized by concrete, deductive, data-based, and experimental-validation-centered studies. We propose to build a qualitative or descriptive model along with appropriate notation or tool for providing specific solutions with validations of a case study. Dependability attributes can be seen from different perspectives, depending on the application. Eventually dependability is intended to prevent errors from becoming failures.

Insights of Web Engineered Business Intelligence Applications. General Business objectives and their functionality for Rich Security Model that users can administer are: Provide more effective mechanisms to move work between business entities, such as self-service for customers or partners or enabling outsourcing by providing business partners a collaborative environment or business data on an extranet. The valuable insights from this approach include ease of use, scalability, disconnected from processes for these developed applications, improved customer satisfaction, increased business agility, reduced time to market, increased revenue and operational efficiency and improvements.

Implementations and Validations. Software Engineering for Web (Web Engineering) covers the definition of processes, techniques and models suitable for its environment to guarantee quality of results. An important design artifact in any software development project is the Software Architecture. Software Architecture's important part is the set of architectural design rules. A primary goal of the architecture is to capture the architecture design decisions. An important part of these design decisions

consists of architectural design rules. In an MDA (Model-Driven Architecture) context, the design of the system architecture is captured in the models of the system. MDA is known to be layered approach for modeling the architectural design rules and uses design patterns to improve the quality of software system. And to include the security to the software system, security patterns are introduced that offer security at the architectural level. Moreover, agile software development methods are used to build secure systems. We had implemented various case studies like Web Services Mining Dashboard application CRM application with spatial capabilities as primary work. Later on we worked on a case study of Design of Agile Modeled Web Services design for secure stock exchange, with focus on mined security architectures insights. For details of these implementations, please refer to the website <http://sites.google.com/site/upendramgitcse>

3 Conclusions

In this paper we discussed mining approach for Business Intelligence to improve insights of Web Engineering applications. Future work includes developing a formal security plan for every connected computer. Performing security inspections of requirements and specifications. Utilizing high-security programming language such as “E”.

References

1. Czernicki, B.: Silverlight 4 Business Intelligence Software. Apress, USA (2010)
2. Mordinyi, R., Kuhn, E., Schatten, A.: Towards an Architectural Framework for Agile Software Development. In: IEEE 17th International Conference and Workshop on Engineering of Computer Based Systems (ECBS), pp. 276–280 (2010)
3. Governor, J., Hinchcliffe, D., Nickull, D.: Web 2.0 Architectures. Oreilly Publishers, USA (2009)
4. Arsanjani, A., Zhang, L.-J., Ellis, M., Allam, A., Chennabasavaiah, K.: S3 A Service-Oriented Reference Architecture. In: IEEE IT Pro., pp. 10–17 (May/June 2007)
5. Chivers, H., Paige, R.F., Ge, X.: Agile Security Using an Incremental Security Architecture. In: Baumeister, H., Marchesi, M., Holcombe, M. (eds.) XP 2005. LNCS, vol. 3556, pp. 57–65. Springer, Heidelberg (2005)
6. Yee, C.G., Radha Krishna Rao, G.S.V.: An Introductory study on Business Intelligence Security, pp. 204–217. Idea Group Inc., USA (2007)
7. Fernandez, E.B., Yoshika, N., Washizaki, H., Jurjens, J., VanHilst, M., Pernul, G.: Using Security Patterns to Develop Secure Systems. IGI Global, pp. 16–31 (2011), doi:10.4018/978-1-61520-837-1.ch002
8. Chung, W.: Designing Web-based Business Intelligence Systems: A Framework and Case Studies. In: DESRIST, California, CA, USA, February 24–25, pp. 147–171 (2006)
9. Chivers, H., Paige, R.F., Ge, X.: Agile Security Using an Incremental Security Architecture. In: Baumeister, H., Marchesi, M., Holcombe, M. (eds.) XP 2005. LNCS, vol. 3556, pp. 57–65. Springer, Heidelberg (2005)

A Modified Continuous Particle Swarm Optimization Algorithm for Uncapacitated Facility Location Problem

Sujay Saha¹, Arnab Kole², and Kashinath Dey³

¹ Assistant Professor, CSE Department, Heritage Institute Of Technology, Kolkata, India
sujay.saha@heritageit.edu

² M. Tech Student, CSE Department, Heritage Institute Of Technology, Kolkata, India
arnab.kole@heritageit.edu

³ Associate Professor, CSE Department, University Of Calcutta, Kolkata, India
kndey55@rediffmail.com

Abstract. A continuous version of particle swarm optimization (CPSO) is employed to solve uncapacitated facility location (UFL) problem which is one of the most widely studied in combinatorial optimization. The basic algorithm had already been published in the Research Article “A Discrete Particle Swarm Optimization Algorithm for Uncapacitated Facility Location Problem” [1]. But in addition to that, the algorithm is slightly modified here to get better result in a lesser time. To make a reasonable comparison, the same benchmark suites that are collected from OR-library [6] are applied here. In conclusion, the results showed that this modified CPSO algorithm is slightly better than the published CPSO algorithm.

Keywords: Swarm Intelligence, Continuous Particle Swarm Optimization, UFL, Inertia Factor, Acceleration Coefficient.

1 Introduction

Efficient supply chain management has led to increased profit, increased market share, reduced operating cost and improved customer satisfaction for many businesses. One strategic decision is related to physical distribution structure in supply chain management including locating facilities and allocating customers to them is facility location. Facility Location, also known as Location Analysis, is a branch of Operation Research [4] concerning itself with mathematical modeling and solution of problems concerning optimal placement of facilities in order to minimize the transportation costs, avoid placing hazardous materials near housing, outperform competitors’ facilities, etc. We are trying to solve a ‘Uncapacitated Facility Location Problem’ which finds the number of facilities to be established and specify those facilities such that the total cost will be minimized considering the total cost means a fixed cost of setting up a facility in a given site and a transportation cost of satisfying the customer requirements from a facility. We maintain the constraints that there is no limit of capacity for any candidate facility and the whole demand of each customer has to be assigned to one of the facility.

The organization of the paper is as follows: in Section 2, a basic idea is given for Particle Swarm Optimization Technique. Section 3 reports about the definition of the Uncapacitated Facility Location Problem along with the constraints. Section 4 describes the published CPSO algorithm. Section 5 describes the modified CPSO algorithm for solving the Uncapacitated Facility Location problem. Section 6 provides the comparison results of these two algorithms. Finally, Section 7 presents the conclusion driven.

2 Particle Swarm Optimization (PSO)

The class of complex systems sometimes referred to as swarm systems is a rich source of novel computational methods that can solve difficult problems efficiently and reliably. When swarms solve problems in nature, their abilities are usually attributed to swarm intelligence [3]; perhaps the best-known examples are colonies of social insects such as termites, bees, and ants. One of the best-developed techniques of this type is particle swarm optimization (PSO). In PSOs, which are inspired by flocks of birds and shoals of fish, a number of simple entities, the particles, are placed in the parameter space of some problem or function, and each evaluates the fitness at its current location. Each particle then determines its movement through the parameter space by combining some aspect of the history of its own fitness values with those of one or more members of the swarm, and then moving through the parameter space with a velocity determined by the locations and processed fitness values of those other members, along with some random perturbations. The members of the swarm that a particle can interact with are called its social neighborhoods. Together the social neighborhoods of all particles form a PSOs social network. More precisely, in the canonical version of PSO, each particle is moved by two elastic forces, one attracting it with random magnitude to the fittest location so far encountered by the particle, and one attracting it with random magnitude to the best location encountered by any of the particle's social neighbors in the swarm. If the problem is N dimensional, each particle's position and velocity can be represented as a vector with N components (one for each dimension). Starting with the velocity vector, $v = (v_1, \dots, v_N)$, each component, v_i , is given by:

$$v_i(t+1) = \omega v_i(t) + \psi_1 R_1 (x_{s_i} - x_i(t)) + \psi_2 R_2 (x_{p_i} - x_i(t)) \quad (1)$$

where x_{s_i} is the i th component of the best point visited by the neighbors of the particle, $x_i(t)$ is the i th component of the particle's current location, x_{p_i} is the i th component of its personal best, R_1 and R_2 are two independent random variable uniformly distributed over $[0, 1]$ is a constant known as the inertia weight, and ψ_1 and ψ_2 are two constants, known as the acceleration coefficients, which control the relative proportion of cognition and social interaction in the swarm. The same formula is used independently for each dimension of the problem, and synchronously for all particles. The position of a particle is updated every time step using the equation:

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (2)$$

The next iteration takes place after all particles have been moved. Eventually, the swarm as a whole, like a flock of birds collectively foraging for food, is likely to move close to the best location.

3 UFL Definition

In a UFL problem, there are a number of customers, m , to be satisfied by a number of facilities, n . Each facility has a fixed cost, fc_j . A transport cost, c_{ij} , is accrued for serving customer, i , from facility, j . There is no limit of capacity for any candidate facility and the whole demand of each customer has to be assigned to one of the facilities. We are asked to find the number of facilities to be established and specify those facilities such that the total cost will be minimized. The mathematical formulation of the problem can be stated as follows:

$$Z = \min(\sum_{i=1}^m \sum_{j=1}^n c_{ij} \cdot x_{ij} + \sum_{j=1}^n fc_j \cdot y_j). \tag{3}$$

subject to:

$$\sum_{j=1}^n x_{ij} = 1 \quad \forall i \text{ in } m. \tag{4}$$

$$0 \leq x_{ij} \leq y_j \quad y_j \in \{0; 1\}. \tag{5}$$

where $i = 1 \dots m; j = 1, \dots, n; x_{ij}$ represents the quantity supplied from facility i to customer j ; y_j indicates whether facility j is established ($y_j = 1$) or not ($y_j = 0$). Constraint (4) makes sure that all customers demands have been met by an open facility and (5) is to keep integrity. Since it is assumed that there is no capacity limit for any facility, the demand size of each customer is ignored and therefore (3) established without considering demand variable.

4 Earlier Related Work for UFL Problem

According to Sevkli and Guner [1], CPSO considers each particle has three key vectors: position (X_i), velocity (V_i), and open facility (Y_i). $X_i = [X_{i1}, X_{i2}, \dots, X_{in}]$ denotes the i th position vector in the swarm, where X_{ik} is the position of the i th particle with respect to the k th dimension. Similarly $V_i = [V_{i1}, V_{i2}, \dots, V_{in}]$ is the i th velocity in the swarm, where V_{ik} is the velocity of the i th particle with respect to the k th dimension. Y_i represents the opening or closing facilities based on the position vector (X_i), $Y_i = [Y_{i1}, Y_{i2}, \dots, Y_{in}]$ where Y_{ik} represents opening or closing k th facility of the i th particle. For an n -facility problem, each particle contains n number of dimensions. Initially the positions and velocities are generated as continuous uniform random variables using the following rules:

$$\begin{aligned} x_{ij} &= x_{\min} + (x_{\max} - x_{\min}) \times r_1 \\ v_{ij} &= v_{\min} + (v_{\max} - v_{\min}) \times r_2 \end{aligned} \tag{6}$$

where $x_{min} = -10.0$, $x_{max} = 10.0$, $v_{min} = -4.0$, $v_{max} = 4.0$, $r_1, r_2 \in [0, 1]$. The position vectors do not represent a candidate solution to calculate the total cost (fitness value). In order to create a candidate solution the position vector is converted to a binary variable $Y_i \leftarrow X_i$, which is also a key element of a particle. This conversion is done using the following formula:

$$y_i = \lfloor \lfloor x_i \rfloor \pmod{2} \rfloor \tag{7}$$

For example, if the position value of one particle with respect to one dimension is -7.47, then the corresponding open facility vector will be

$$\lfloor \lfloor -7.47 \rfloor \pmod{2} \rfloor = \lfloor \lfloor 7.47 \pmod{2} \rfloor \rfloor = \lfloor \lfloor 1.47 \rfloor \rfloor = 1 \tag{8}$$

After generating the position vectors and open facility vectors, the total cost of each particle is calculated as per equation (3) and initially set these cost as local best or personal best (*Pti*) for each particle and the minimum of these personal bests is set as global best (*Gi*) of the swarm. Next, the velocity of each particle is updated as per equation (1) and the position is updated as per equation (2) where ω is the inertia factor used to control the impact of the previous velocities on the current one. After updating the position value for all particles, the corresponding open facility vector can be determined to start a new iteration if the predetermined stopping criterion is not yet met. The stopping criteria can be either getting an optimal solution or reaching the maximum number of iterations chosen for obtaining the result in a reasonable CPU time.

5 Modifications

To reduce the execution time & the cost, the above algorithm is modified in the following manners: first, to generate the position and the velocity in equation (6), if $r_1 = 0.5$ then $x_{ij} = 0$ and similarly if $r_2 = 0.5$ then $v_{ij} = 0$. So instead of doing the addition and multiplication which take some time, the values of both x_{ij} and v_{ij} can be set to zero for the above said value of r_1 and r_2 . Secondly, instead of keeping the fixed inertia factor, the value of ω can be varied from 0.9 to 0.4 in various iterations to get the lesser optimal cost. So the modified CPSO algorithm is as follows:

Modified CPSO Algorithm for UFL

Begin

Initialize positions (population) randomly For each particle as:

Generate r_1 randomly in $[0, 1]$

If $r_1 = 0.5$ then $x_{ij} = 0$, else

$$x_{ij} = x_{min} + (x_{max} - x_{min}) \times r_1$$

End

Initialize velocities randomly For each particle as:

Generate r_2 randomly in $[0, 1]$

If $r_2 = 0.5$ then $v_{ij} = 0$, else

$$v_{ij} = v_{min} + (v_{max} - v_{min}) \times r_2$$

End

```

Calculate open facility vector as per equation (7)
Calculate fitness value using open facility vector as per equation (3)
Set to position vector and fitness value as personal best (Pti )
Select the best particle and its position vector as global best (Gt)
End
Do {
  For each particle
    Update velocity as per equation (1)
    Update position as per equation (2)
    Find open facility vectors & calculate the fitness value using open facility
    vector
    Update personal best (Pti )
    Update the global best (Gt) value with position vector
  End
  Decrement the value of  $\omega$  by:  $(0.9 - 0.4) / \text{total no. of iterations}$ 
} While (Maximum Iteration is not reached)

```

Table 1. An illustration of deriving open facility vector from position vector for a 4-customer 3-facility problem

ith particle vectors	Particle dimension (k)		
	1	2	3
Position Vector (Xi)	3.8	-0.93	5.42
Open Facility Vector (Yi)	1	0	1

Table 2. An example of 3-facility to 4-customer

Facility Locations		1	2	3
Fixed Cost		15	8	3
Customers	1	12	3	7
	2	2	8	5
	3	4	6	14
	4	9	1	10

Considering the 3-facility to 4-customer problem shown in Table 1, the total cost of open facility vectors can be calculated as follows:

$$\begin{aligned}
 \text{Total Cost} &= \{\text{open facilities fixed cost } (f_{c_j}) + \min(\text{cost of supply from open facilities to customers } i[c_i])\} \\
 &= \{(15+3) + \min(12, 7) + \min(2, 5) + \min(4, 14) + \min(9, 10)\} = \{18+2+7+4+9\} = \{40\}.
 \end{aligned}$$

6 Comparison of Results

The modified CPSO algorithm is coded with C language in Linux Platform and run on an Intel Core 2 Duo 2.4 GHz Laptop with 1GB memory. The results are shown in the following table:

Table 3. Experimental Results gained with modified CPSO algorithm

Benchmark					
Problems	Size (m x n)	As Per Earlier CPSO Algorithm		As Per Modified CPSO Algorithm	
		Optimal Cost	ACPU Time	Optimal Cost	ACPU Time
Cap 71	16 x 50	932615.75	0.1218	932597.00	0.9700
Cap 72	16 x 50	977799.40	0.1318	977779.00	0.7500
Cap 73	16 x 50	1010641.45	0.1865	1010619.00	0.5900
Cap 74	16 x 50	1034976.98	0.1781	1034956.00	0.5300
Cap 101	25 x 50	796648.44	0.8818	796626.00	1.3800
Cap 102	25 x 50	854704.20	0.7667	854682.00	1.2400
Cap 103	25 x 50	893782.11	0.9938	893758.00	0.9900
Cap 104	25 x 50	928941.75	0.6026	928918.00	0.7700
Cap 131	50 x 50	793439.56	3.6156	794137.00	2.1300
Cap 132	50 x 50	851495.33	3.5599	853149.00	1.8900
Cap 133	50 x 50	893076.71	3.7792	894072.00	1.8900
Cap 134	50 x 50	928941.75	3.3333	928918.00	1.2400

7 Conclusion

In this paper, CPSO algorithm is applied to solve UFL problems. We have slightly modified this algorithm for initializing the particle's positions and velocities for a particular value of the random numbers r_1 and r_2 to reduce the average CPU time. Secondly, instead of keeping a fixed ω value, we have varied the value in the range of 0.9 to 0.4 over the total number of iterations. This modified algorithm has been tested on 12 files of OR-Library and the better results either in terms of time or in terms of optimal cost are obtained.

References

1. Guner Ali, R., Mehmet, S.: A Discrete Particle Swarm Optimization Algorithm For Uncapacitated Facility Location Problem. Hindawi Publishing Corporation Journal of Artificial Evolution And Applications 2008, Article ID 861512, 9 (2008), doi:10.1155/2008/861512
2. Kennedy, J., Eberhart, R.C., Shi, Y.: Swarm Intelligence. Morgan Kaufmann, San Francisco (2001)
3. Ghosh, D.: Neighborhood search heuristics for the uncapacitated facility location problem. European Journal of Operational Research 150(1), 150–162 (2003)

4. Eberhart, R.C., Kennedy, J.: New optimizer using particle swarm theory. In: Proceedings of the 6th International Symposium on Micro Machine and Human Science (MHS 1995), Nagoya, Japan, October 1995, pp. 39–43 (1995)
5. Beasley, J.E.: OR-Library (2005),
<http://people.brunel.ac.uk/mastjjb/jeb/info.html>

Design of Hybrid Genetic Algorithm with Preferential Local Search for Multiobjective Optimization Problems

J. Bhuvana and C. Aravindan

Department of Computer science and Engineering,
SSN College of Engineering, Chennai, India
{bhuvanaj, aravindanc}@ssn.edu.in

Abstract. Evolutionary algorithms are used to obtain the optimal solutions for varieties of engineering problems. The performance of evolutionary algorithms can be enhanced by integrating them with local search methods. The idea is to combine the best of both global and local optimization approaches to perform a better exploration of search space. This paper presents a preferential hybrid evolutionary algorithm, where the gradient descent method is integrated into the NSGA-II. This new algorithm has been verified on a set of multiobjective benchmark problems using four different performance metrics. The results show that the proposed algorithm brings out the optimal solutions with better diversity and closeness to the known optimal solutions than NSGA-II and also consumes less time than traditional hybrid algorithm.

Keywords: Hybrid GA, Gradient Descent, Preferential local search.

1 Introduction

Memetic or hybrid genetic algorithms combine the best of both global and local search techniques to improve the quality of the solutions. The ability of the local search to produce a local optimal solution when integrated with global search method using genetic algorithms will generate diverse set of pareto optimal solutions and also prevent premature convergence [5]. A hybrid GA with preferential local search has been proposed in this paper, which addresses the issues such as where to integrate local optimization and how solutions are favored to undergo the local search to reduce the computational time of any hybrid approach. Experimental results show that this approach out-performs the traditional hybrid algorithm in terms of metrics and computational time.

The proposed Hybrid GA with Preferential Local search is based on NSGA-II as the global search and Gradient Descent as the local search method. NSGA-II [8] is one of the most popular elitist genetic algorithms [2] uses ranking method based on non dominated sorting. Gradient Descent local search minimizes a function $f(x)$ where $x \in \mathbb{R}$, by computing a new solution using $x_{k+1} = x_k + \alpha_k d_k$; $k = 0, 1, \dots$, where the k is the step size [12] and d_k is the descent direction given by $-\Delta f(x_k)$.

Varieties of Hybrid algorithms have been reported in the literature [3]. Review about most commonly used hybrid architectures is available in [6]. Adaptive Local

Search (ALS) is applied by combining weighted fitness and restricted mutation in [1]. New weighted fitness is assigned to N_s superior individuals and applies restricted crossover and mutation. Local search based EMO in [11] is a hybrid algorithm with Achievement Scalarizing Functions (ASF) and local search method. Local search begins from a reference point, minimizes the associated ASF. Two hybridization schemes are discussed in [7] such as, GA with LS, and GA then LS. First hybrid scheme applies LS to solutions of each generation of GA where the second one applies LS to solutions obtained from GA. And also recommends the second approach as the preferable scheme.

2 Algorithm Description

The idea is to propose a preferential hybrid approach that identifies the optimal solutions by combining local search and global search methods. Whenever new individuals are generated, they undergo a limited local search. As generation progresses, when the individuals survive across generation further local search is applied on them. Thus the local search iteratively deepens on potential candidates thereby introducing elitism in local search.

In designing a preferential hybrid approach, this paper proposes a new method of combining the NSGA-II and the Gradient Descent local search method. The basic assumption here is the objective functions considered are differentiable. The random values populate the initial population and mating pool is created using binary tournament selection. SBX cross over with polynomial mutation are used to generate new offsprings. A limited local search is applied on the offsprings and will exploit the neighborhood space and prepare them to compete with their parents for survival. To emphasize the potential solutions, preferential local search is applied on the surviving parents. This approach is entirely different from traditional memetic or hybrid algorithms and addresses the issue of termination condition in a way. The pseudo code for the proposed hybrid algorithm with preferential Local search is given in Table 1. For local search, the weighted sum method is employed to convert the multiple objectives into a single objective function, where the weights are uniformly distributed to all the objectives.

Table 1. Pseudo code for Hybrid GA with Preferential Local search (HPLS)

- | |
|---|
| <ol style="list-style-type: none"> 1. Initialize population with random solutions. 2. Repeat until termination condition is met. 3. Evaluate the fitness. 4. Sort the population according to non domination. Assign ranks accordingly. 5. Perform the selection. 6. Apply preferential local search on elitist parent solutions 7. Apply crossover and Mutation to generate new offsprings. 8. Apply limited local search on all offsprings. 9. Update the population with values obtained from Local search and best parents |
|---|

3 Implementation and Experiments

The Hybrid NSGA-II with preferential local search (HPLS) was implemented in C, with a population size of 40 for 7000 generations. The crossover and mutation probability were set to 1 and $1/(\text{number of variables})$ respectively. Same initial population was fed for three approaches, HPLS, Hybrid NSGA-II with local search (HLS) and NSGA-II for a single run and 10 such runs were accomplished.

The average of performance metrics were computed and used for further statistical testing. The gradient for one descent step with a step size of 0.01 was computed. Six unconstrained bi-objective test problems such as SCH, BINH, ZDT1, ZDT2, ZDT3 and ZDT6 [4, 9] have been used for establishing the performance of the HPLS. The performances of HPLS, HLS and NSGA-II over the benchmark problems were compared on the basis of two major criteria such as, Closeness of obtained solutions to known Pareto optimal solutions and the diversity of solutions along the Pareto optimal front. Four different metrics such as, Error Ratio(ER), Generational Distance (GD), Spread and Hyper volume ratio (HVR) have been used to compare the performances of three approaches. Time taken by the algorithms have been observed and used for comparison.

Table 2. Performance Metrics

Problem	Algo.	ER	GD	Spread	HVR	Avg. Time(ms)
ZDT1	HPLS	0.0144	0.1511	0.8960	0.2399	117.898
	NSGA-II	0.0250	0.1555	0.9264	0.1822	206.915(HLS)
ZDT2	HPLS	0.0225	0.1503	0.9437	0.2309	138.455
	NSGA-II	0.1250	0.1544	0.9966	0.1747	287.227(HLS)
ZDT3	HPLS	0.0150	0.1566	0.0561	0.5736	137.1282
	NSGA-II	0.0250	0.1708	0.2681	0.4671	359.080(HLS)
ZDT6	HPLS	0.0175	0.1550	0.6260	0.4173	28.918
	NSGA-II	0.0250	0.1650	0.6480	0.3155	46.836(HLS)
SCH1	HPLS	0.0250	0.1548	0.3161	0.0644	14.659
	NSGA-II	0.2825	0.1631	0.5513	0.0338	18.327(HLS)
BINH	HPLS	0.0755	0.1497	0.0649	0.0100	15.1496
	NSGA-II	0.3875	0.1655	0.3068	0.0062	38.3964(HLS)

4 Results and Discussions

The performance metrics in Table 2 state that HPLS outperformed NSGA-II in all the metrics. When HPLS compared with HLS following observations were made, 51.5% improvement in the ER, about 2.2% enhancement in GD, 14.5% improvement in Spread and for HVR there was 18.2% increase in performances. Apart from performance metrics, HPLS has consumed less than 50% of time consumed by HLS for 10 fixed steps shown in Table 2. HPLS has yielded more number of non dominating optimal solutions than other two was inferred from Figure 1. The global search method when accompanied by preferential local search made a thorough exploration of the search space than global search alone or the traditional hybrid method. Paired t test [10] was applied over the four metrics computed across

generations. The results confirmed that there was a significant difference between these algorithms. This inference when associated with the obtained performance metrics revealed the fact that HPLS and HLS were different and the HPLS behaved better than the HLS and NSGA-II in terms of closeness to the known optimal solutions and also generated comparatively good spread of solutions. And the objective of the new approach, exploring the search space globally and also locally was henceforth achieved.

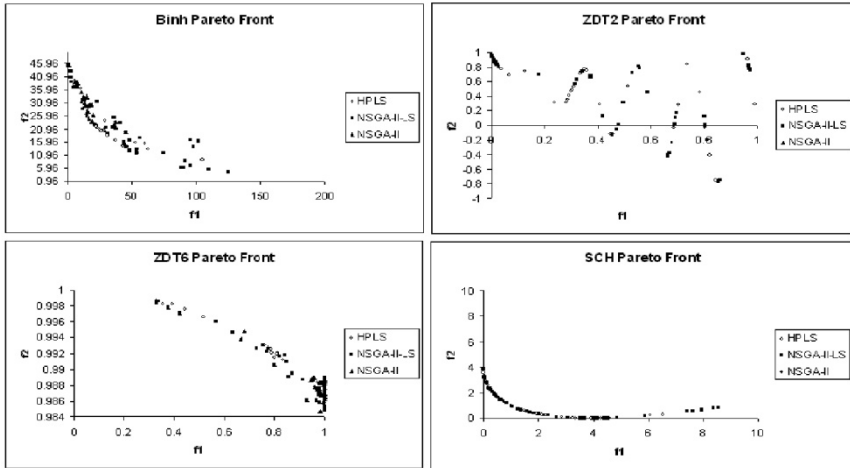


Fig. 1. Pareto Fronts of Binh, SCH, ZDT2 and ZDT6

5 Conclusion

This paper has introduced a new elitist approach with reduced computation time by applying preferential local search integrated with a multiobjective GA. Two local searches have been used, where the first one was preferentially applied over elitist parents. Parents surviving in the next generation reveal the fact that they are prominently good solutions and require to be preserved. A limited second local search applied on offsprings enable the optimization to exploit and explore the unexplored areas of search space. This work can be extended further by deciding the termination conditions for the local search applied. The experimental results show that HPLS not only taken evolution more closely towards the optimal solutions than the traditional hybrid approach but also yielded more number of additional non dominant solutions with less than 50% of computational time.

References

1. Ahn, C.W., et al.: A hybrid evolutionary algorithm for multiobjective optimization. In: Fourth International Conference on Bio-Inspired Computing (BICTA 2009), pp. 1–5 (2009)
2. Coello, C.A.C.: Twenty years of evolutionary multi-objective optimization: what has been done and what remains to be done. In: Computational Intelligence: Principles and Practice, ch. 4, pp. 73–88. IEEE Computational Intelligence Society, Los Alamitos (2006)

3. Coello, C.A.C.: Evolutionary Multiobjective Optimization: some current research trends and topics that remain to be explored, vol. 3(1), pp. 18–30. Higher Education Press and Springer Verlag (2009)
4. van Veldhuizen, D.A., et al.: Multiobjective evolutionary algorithm test suites. In: Proceedings of the 1999 ACM symposium on Applied computing, Texas, United States, pp. 351–357 (1999)
5. Elmihoub, T., et al.: Hybrid genetic algorithms - a review. *Engineering Letters* 13, 124–137 (2006)
6. Grosan, C., Abraham, A.: Hybrid evolutionary algorithms: Methodologies, architectures, and reviews. In: Grosan, C., Abraham, A. (eds.) *Hybrid Evolutionary Algorithms*. SCI, vol. 75, pp. 1–17. Springer, Heidelberg (2007)
7. Ken, H., et al.: Hybridization of genetic algorithm and local search in multiobjective function optimization: recommendation of GA then LS. In: *GECCO 2006: Proceedings of the 8th Annual conference on Genetic and Evolutionary Computation*, pp. 667–674. ACM, USA (2006)
8. Deb, K., et al.: A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II. In: Deb, K., Rudolph, G., Lutton, E., Merelo, J.J., Schoenauer, M., Schwefel, H.-P., Yao, X. (eds.) *PPSN 2000*. LNCS, vol. 1917, pp. 849–858. Springer, Heidelberg (2000)
9. Deb, K.: *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons(Asia) Pte Ltd., Singapore (2001)
10. Ross, S.M.: *Introduction to Probability and Statistics for Engineers and Scientists*. Academic Press, Singapore (2004)
11. Karthik, S., et al.: Local search based evolutionary multiobjective optimization algorithm for constrained and unconstrained problems. In: *CEC 2009: Proceedings of the Eleventh conference on Congress on Evolutionary Computation*, Norway, pp. 2919–2926 (2009)
12. Salomon, R.: Evolutionary algorithms and gradient search: similarities and differences. *IEEE Transactions on Evolutionary Computation* 2(2), 45–55 (1998)

Synergy of Multi-agent Coordination Technique and Optimization Techniques for Patient Scheduling

E. Grace Mary Kanaga¹ and M.L. Valarmathi²

¹ Department of Computer science and Engineering,
Karunya University, Karunya Nagar, Coimbatore-641 114,
Tamil Nadu, India
grace@karunya.edu

² Department of Computer science and Engineering,
Government College of Technology, Coimbatore-641 013,
Tamil Nadu, India
ml_valarmathi@rediffmail.com

Abstract. This paper discusses about the synergy of multi-agent coordination technique with different optimization techniques viz. integer programming optimization with Lagrangian Relaxation and a simple heuristic approach with experience based learning effect in case of patient scheduling problem. The objective is to achieve good reduction in waiting time of the patients in hospital while achieving better resource utilization. The proposed methods have been implemented in JADE and the results are compared with the traditional scheduling techniques. The performance of the proposed methods based on total weighted waiting time of the patient increases by 15% to 52% when compared to traditional scheduling techniques. The waiting time of the patient is further reduced by 5%, when experience based learning effect is incorporated. Optimization based on simple heuristic method shows that it gives near optimal solution with drastic reduction in the execution time ie. 90.84% when compared to integer programming approach.

Keywords: Multi-agents, Coordination technique, Optimization techniques, Patient scheduling.

1 Introduction

A high amount of complexity is faced by the patient scheduling system in hospitals and coordinating patients for having an optimum schedule is a difficult task. Depending on the available resources, the schedules may be more or less congested. In countries like India demand regularly exceeds the resource availability. According to the statistics, WHO (World Health Organization) says that India has one doctor per 2500 people. It is well below international standards. However, even in this situation, there is no effective scheduling is used to reduce the waiting time of the patients and to improve the resource utilization in the hospital. An efficient agent-based patient scheduling system with various optimization techniques is discussed in this paper.

2 Related Work

In many current scheduling approaches focus is only on single units which are centralized. These systems do not account for the distribution and dynamics of the patient scheduling problem and this has been discussed by Paulussen et al. [1]. Emergencies or complications lead to disturbances in the schedule and result in variable waiting times for other patients according to Bartelt et al. [2]. The main challenge of the patient scheduling problem is that it is distributed and it is a safety critical real time system.

The hospital scheduling problem is diverse and many papers have addressed different aspects of the hospital scheduling problem. Some of the major problems addressed in literature are, patient/medical appointment scheduling as seen in Vermuelen et al. [3], and Kaandorp and Koole [4]. Another type is the specialized case of Surgical Case Scheduling (SCS) found in Pham and Klinkert [5] / Operating Room (OR) scheduling as given by Becker and Navarro [6]. Marinagi et al. [7] and Kapamara et al. [8] discussed about patient scheduling in laboratories and scheduling for treatment of radiotherapy respectively.

3 Agent-Based Patient Scheduling

In the proposed Agent-based Patient Scheduling (APS) framework, each patient and resource is represented by an agent and a combinatorial auction is used to produce a schedule. This framework consists of the Patient Agent (PA), the Resource Agent (RA) and the Common Agent (CA). CA represents a general physician who assigns the initial treatment plan. Each patient agent, PA_i has a set of tasks $T_{1i}, T_{2i}, \dots, T_{Mi}$. These tasks represent the treatment plan prescribed for each patient. Each resource agent, RA_m conducts an auction at a decision point, t_c . The patient agents available at the time of auction participate in the auction. Each patient agent PA_i participates in auctions conducted by the resource agents for the tasks to be performed. The time slot is allocated to the PA by RA, such that the total weighted waiting time is minimum. Different optimization techniques such as Integer Programming Optimization with Lagrangean Relaxation (IPLR), a heuristic optimization and an Experience Based Learning (EBL) effect is incorporated in this APS frame work is discussed in the following sections.

3.1 Agent-Based Patient Scheduling with IP-LR

The Integer Programming (IP) formulation for agent-based patient scheduling problem is modeled as follows.

The objective function which minimizes total weighted waiting time is given by,

$$\sum_{i=1}^{N_c} \sum_{t=t_c}^{t_c+L} w_i T_i X_{iO_i t} \tag{1}$$

is subject to the following constraints:

Constraint 1: To ensure that each task of each patient can finish within an allotted time slot.

$$\sum_{t=t_c}^{t_c+L} X_{ijt} = 1, \forall i, j \tag{2}$$

Constraint 2: Task Precedence constraints of two tasks.

$$\sum_{t=t_c}^{t_c+L} tX_{ijt} + p_{ij+1} \leq \sum_{t=t_c}^{t_c+L} tX_{i,j+1,t}, \forall i, j \tag{3}$$

Constraint 3: Resource capacity constraints to ensure that the capacity of the resource is not violated in each timeslot.

$$\sum_{i=1}^{N_c} \sum_{j=K_i}^{O_i} X_{ijt} Y_{ijm} + \sum_{i=1}^{N_c} \sum_{j=K_i}^{O_i} \sum_{t=L+1}^{\min\{L,t+p_{ij}-1\}} X_{ijt'} Y_{ijm} \leq 1, \forall m, t \tag{4}$$

Constraint 4: Patient release date constraints to ensure that the first task cannot be completed before the patient has been in the hospital for at least the time equal to the processing of the first task.

$$\sum_{t=t_c}^{t_c+L} tX_{ijt} \geq p_{i1} + r_i, \forall i \tag{5}$$

where w_i is the weight based on priority assigned to patient i , T_i the tardiness (waiting time) of patient i , X_{ijt} is the integer variable equals 1 if task (i,j) is completed at time t , Y_{ijm} is the integer variable equals 1 if task (i,j) is processed on resource m , P_{ij} is the processing time of task j of patient i and i, j, m are the indices of the patient, task and resource respectively.

The above problem formulation is for a centralized approach [9]. Lagrangean Relaxation (LR) of the resource constraints is done to decompose the problem into patient level and resource level sub problems in order to solve the problem in a distributed environment.

3.2 Agent-Based Patient Scheduling with Experience-Based Learning

The phenomenon of “learning effect” is investigated in the context of scheduling [10] and it shows that the Experience Based Learning (EBL) gives better solution. Hence in this paper this EBL effect is incorporated in the previously explained IPLR method. In this approach each resource has a Learning Agent (LA) and this learning agent gains its experience whenever a patient is allocated to that particular resource. The time required to process a particular task is reduced as the experience gained by it increases. Based on EBL effect, the completion time of patient i scheduled in the v^{th} slot in a sequence using experience based learning is given by:

$$C_i = C_{i-1} + P_{ij} (E_{\beta}(v)) \tag{6}$$

where the processing time of patient i scheduled in the v th position in a sequence of resource j is given as follows:

$$P_{ij}(E_{\beta}(v)) = P_{ij} - b_i (\min\{E_{\beta}(v), g_i\})^{\alpha_i} \quad (7)$$

where P_{ij} is the normal processing time without experience and $i, v = 1, \dots, n$, and $j = 1, \dots, m$ and the experience of the resource at the start of execution of the v th patient is,

$$E_{\beta}(v) \triangleq E(v-1) + \beta_{[v]} e_{[v]} = \sum_{i=1}^{v-1} e_{[i]} + \beta_{[v]} e_{[v]} \quad (8)$$

where $\sum_{i=1}^{v-1} e_{[i]}$ denotes the experience already possessed by the resource and $\beta_{[v]} \in [0,1]$ is the amount of experience (percentage of $e_{[v]}$) provided to the resource by patient in the v th position at the start of its execution. The other variables, $b_i > 0$ denote the linear learning ratio of the patient i , and g_i is its learning threshold.

3.3 Agent-Based Patient Scheduling with EBL and Simple Heuristic Approach

Heuristic algorithms called suboptimal procedures or simply, heuristics are designed to find the best possible solutions with small computational efforts.

Heuristic algorithm

- Step 1: Obtain the list of patients for each of the resources.*
- Step 2: Set $j=1$, $y=2$ and $k=2$. Pick the first two patients from the list of patients of j^{th} resource and schedule them in order to minimize the weighted sum of total completion time and total tardiness.*
- Step 3: Check whether the schedule is satisfied for the precedence constraint.*
If so,
- a) Update the selected partial solution as the new current solution otherwise*
 - b) Repeat the process with the next lowest value of weighted sum of total completion time.*
- where the completion time of a patient i processed in v^{th} time slot is calculated using equation(6).*
- Step 4: Check whether y exceeds n_j*
If yes go to Step 6
Otherwise increment k and y by one.
- Step 5: Generate k candidate sequences by inserting the first patient in the remaining patient list into each slot of the current solution. Among these candidates select the best one with the least partial minimization of the weighted sum of total completion time and total tardiness and go to Step3.*
- Step 6: Update the values of i and k and repeat until $i=m$.*
- Step 7: Stop.*

4 Results and Discussion

The proposed methods are implemented in the JADE platform. In order to evaluate the performance of the proposed methods with other traditional scheduling techniques

such as FCFS, MS, LPT and SPT, data instances are created with 10 resources and varying number of patients. The arrival time and the weight of the patients are generated randomly. The weight of the patients varies from 1 to 5. Higher weight values are assigned to the patient in poor health state. The task sequence for each patient is allotted randomly.

4.1 Performance Metrics

There are several performance metrics used for scheduling problems. The objective of patient scheduling problem is to minimize the patient waiting time and improve the resource utilization in the hospitals. A good scheduling technique will effectively minimize the values of all these metrics.

- Total weighted Waiting Time ($\sum w_i T_i$)
- Maximum Waiting Time (T_{max})
- Total Completion time ($\sum C_i$)
- Total Waiting Time ($\sum T_i$)
- Maximum Completion Time (C_{max})
- Total Weighted Completion time ($\sum w_i C_i$)

4.2 Comparison of Results

The proposed methods have been tested with the data set for 50x10 patient scheduling problem to study their performance. The chief metric, weighted waiting time of the APS-IPLR method is compared with the traditional techniques and it is seen that APS-IPLR performance is similar to FCFS and better than traditional techniques MS, SPT, LPT. For 50 patients, the total weighted waiting time APS-IPLR is 1542 whereas in other algorithms it is 1600 to 3200 which varies from 5% - 52% of increase in performance.

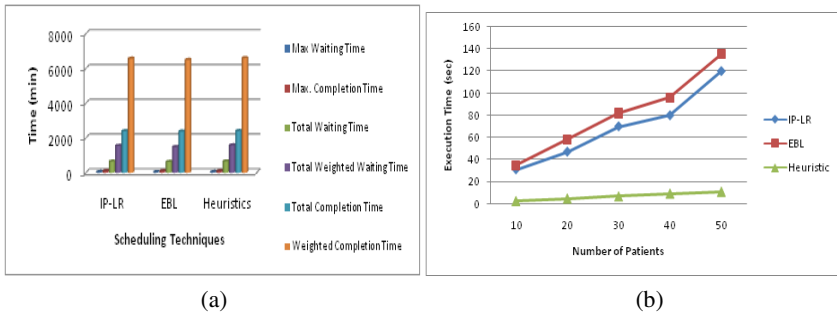


Fig. 1. (a) Comparison of performance metrics (b) Comparison of execution time

Perusal of Fig. 1(a) shows that the waiting time as well as the completion time of the APS with EBL effect is considerably reduced when compare to IP-LR without EBL. The weighted waiting time of the patients has reduced up to 4.47% in APS-EBL. The waiting time as well as the completion time of the heuristics approach is slightly high when compare to IP-LR. Fig. 1 (b) shows the comparison of the execution time of the proposed methods. The average reduction of execution time is 90.84 % when compared to IP-LR. The above discussions show that the simple heuristic

based patient scheduling gives near optimal solutions in terms of weighted waiting time and completion time with an emphatic decrease in the execution time.

5 Conclusion

The agent-based solution for the resource and task allocation problem in the healthcare domain, namely patient scheduling problem is discussed in this paper. The results shows that agent-based patient scheduling with various optimization techniques gives better solution than traditional approaches and also handles dynamic and distributed situations in a robust manner. A schedule that minimizes the patient waiting time may even aid in saving the person's life.

This agent based approaches with optimization techniques can be extended for any complex resource allocation problems. Further the system may be tested with other optimization techniques such as simulated annealing, bee colony optimization approaches.

References

1. Paulussen, T.O., Jennings, N.R., Decker, K.S., Heinzl, A.: Distributed patient scheduling in hospital. In: International Joint Conference on Artificial Intelligence, vol. 18, pp. 1224–1232 (2003)
2. Bartelt, A., Lamersdorf, W., Paulussen, T.O., Heinzl, A.: Agent Oriented Specification for Patient-Scheduling Systems in Hospitals. Dagstuhl Article (2002)
3. Vermeulen, I., Bohte, S., Somefun, K., La Poutré, H.: Multi-agent Pareto appointment exchanging in hospital patient scheduling. *Service Oriented Computing and Applications* 1(3), 185–196 (2007)
4. Kaandorp, G., Koole, G.: Optimal outpatient appointment scheduling. *Health Care Management Science* 10(3), 217–229 (2007)
5. Pham, D.N., Klinkert, A.: Surgical case scheduling as a generalized job shop scheduling problem. *European Journal of Operational Research* 185(3), 1011–1025 (2008)
6. Becker, M., Navarro, M., Krempels, K.H.A., Panchenko: Agent based scheduling of operation theatres. In: EU-LAT eHealth Workshop on Agent Based Scheduling of Operation Theatres (2003)
7. Marinagi, C.C., Spyropoulos, C.D., Papatheodorou, C., Kokkotos, S.: Continual planning and scheduling for managing patient tests in hospital laboratories. *Artificial Intelligence in Medicine* 20(2), 139–154 (2000)
8. Kapamara, T., Sheibani, K., Hass, O.C.L., Reeves, C.R., Petrovic, D.: A review of scheduling problems in radiotherapy. In: Proceedings of 18th International Conference on Systems Engineering, pp. 207–211 (2006)
9. Liu, N., Abdelrahman, M.A., Ramaswamy, S.: A Complete Multi-agent Framework for Robust and adaptable dynamic job Shop Scheduling. *IEEE Trans. On Systems Man And Cybernetics Part C Applications And Reviews* 37(5), 904–916 (2007)
10. Janiak, A., Rudek, R.: Experience-Based Approach to Scheduling Problems With the Learning Effect. *IEEE Transactions on Systems, man, and Cybernetics—part a: systems and human* 39(2), 344–357 (2009)

Naive Bayes Approach for Website Classification

R. Rajalakshmi and C. Aravindan

Department of Computer Science and Engineering
SSN College of Engineering, Chennai, India
{rajalaxmi, aravindanc}@ssn.edu.in

Abstract. World Wide Web has become the largest repository of information because of its connectivity and scalability. With the increase in number of web users and the websites, the need for website classification gains attraction. The website classification based on URLs alone plays an important role, since the contents of web pages need not be fetched for classification. In this paper, a soft computing approach is proposed for classification of websites based on features extracted from URLs alone. The Open Directory Project dataset was considered and the proposed system classified the websites into various categories using Naive Bayes approach. The performance of the system was evaluated and Precision, Recall and F-measure values of 0.7, 0.88 and 0.76 were achieved by this approach.

Keywords: Website classification, URL, Bayes classifier.

1 Introduction

Web page classification is the process of assigning a web page to one or more predefined category labels. The classification of web pages is necessary and it is normally performed based on the content of the website. If the web sites are classified based on URLs alone, the web pages need not be fetched and analyzed. The URL based website classification can be used to identify the abnormal traffic generated from an organization by observing packets. So systems are designed to automate the classification process based on URLs.

In the proposed system, the classification of websites was done based on their URLs alone using Naïve Bayes approach. The websites were classified into one of the categories viz., Arts, Business, Computers, Games, Health, Home, News, Recreation and Reference. The performance of the system was evaluated using ODP dataset and Precision, Recall and F-measure of 0.7, 0.88 and 0.76 were achieved.

2 Related Works

Qi, X. and Davison [1] surveyed the existing automated classification systems and compared the approaches in terms of features and classifiers used. But, all these systems require the content be downloaded before classification. This method of classification after downloading the web page content may not be useful if the objective is to

block objectionable content or when speed is of crucial importance. Min Yen Khan [2] proposed an approach for categorizing web pages without web content using the ILP98 WebKB dataset. For this SVM based classification, the features viz., URL text, anchor text, title text and page text were considered and they reported an average F-measure of 0.43 for the following 4 categories: Course, Faculty, Project and Student. In the approach proposed by Min-Yen [3], the URLs were segmented into meaningful chunks and features such as URL component length, content, orthography, token sequence and precedence were considered. Then Maximum-entropy learning was applied to classify the web sites and an F-measure of 0.62 was achieved. Lim Wern Han et al.[4] proposed a framework for classifying the websites considering the features from URL, web page title and metadata information from web pages. They reported an average F-measure of 0.78 for 6 categories viz., Business, Economy, Entertainment, Government, Health, News and Sports. But they used additional features that are not part of the URL, and considered only 6 categories among which 3 are subcategories. In the proposed system, Naïve Bayes approach for classification of websites was explored using the features extracted only from the URLs.

3 URL Classification

The system was designed to classify the given URL into one of the 9 categories, viz., Arts, Business, Computers, Games, Health, Home, News, Recreation and Reference. A total of 8,55,939 URLs of 9 categories were extracted from the ODP dataset [5]. Among this 7,70,309 URLs were chosen randomly for forming the dictionary. The URLs from each category were parsed into tokens. The common words such as "www", "http", "html" etc. were removed from the list of tokens and only the unique tokens were used for constructing the dictionary. For 9 categories, a total of 9 dictionaries were used for training the Bayes classifier.

By applying Bayes theorem, the probability of given URL to belong to Category C_i can be computed and the URLs can be classified using Naïve Bayes Classifier. To obtain the most probable hypothesis, given the training data, Maximum a posteriori hypothesis h_{MAP} was used.

$$h_{MAP} = \arg \max P(D|h) * P(h) \quad (1)$$

For classification, an URL was separated into tokens and each token was checked with all the dictionaries in the training set. In each dictionary, the occurrence of these tokens was checked and the number of equal matches and the partial matches were counted. The value of each token was computed as shown in Equation (2). The value of the URL depends on its individual tokens, and the dictionary size of the corresponding category.

$$\text{valueToken}_i = \text{eqmatch} + (0.5 * \text{partmatch}) \quad (2)$$

The likelihood probability of an URL, given the Category was estimated by Eqn. (3), in which DictSize_{C_i} is the total number of words in the dictionary for the category

C_i , and i represents category number that varies from 1 to 9. The token count denoted by t , varies from 0 to n , where n represents the maximum number of tokens in a URL.

$$P(URL | C_i) = \frac{\sum_{t=0}^n valueToken_t}{DictSizeC_i} \quad (3)$$

To estimate the probability of the given URL to belong to Category C_i , prior probability and likelihood probability were used. The prior probability was assumed for every category according to ODP dataset. To obtain the prior probability of each category, total number of URLs in each category was divided by total number of URLs in the dataset. The probability of given URL to be classified under the Category C_i was computed using equation (4).

$$P(C_i | URL) = P(URL|C_i) * P(C_i) \quad (4)$$

Here, $P(C_i | URL)$ is the posterior probability of given URL to be in the Category C_i , and $P(URL|C_i)$ is the likelihood probability that is calculated using Equation (3) and $P(C_i)$ is the prior probability of Category C_i . By Bayes theorem, final decision of classification was made. The category for which the URL has the maximum value was adjudged as the class of the URL. The experiments showed acceptable results for this approach of classification.

4 Experiments and Performance Evaluation

For experimental purposes, 9 categories of URLs from ODP dataset were used as shown in Table 1. A set of 9 experiments were performed by taking 80% for training and remaining for testing and cross validation was done. Finally 90% of training set was used and tested with remaining 10% URLs in each category. The system was implemented in Java. To evaluate the performance of the classifier, Precision, Recall and F-measure values were calculated and shown in Table 2. The average value of Precision, Recall and F-measure of 0.7, 0.88, and 0.76 were achieved.

Table 1. Number of URLs in Training and Testing Dataset

Category	Number of URLs	Training Set	Testing Set
Arts	227337	204597	22740
Business	227000	204300	22700
Computers	109201	98390	10921
Games	51207	46080	5127
Health	57515	51750	5755
Home	25368	22824	2544
News	8235	7407	828
Recreation	94565	85104	9461
Reference	55521	49967	5554
Total	8,55,939	7,70,309	85,630

Table 2. Performance measures

	Arts	Business	Comp	Games	Health	Home	News	Recn	Refer
Precision	0.80	0.97	0.65	0.69	0.76	0.80	0.53	0.66	0.44
Recall	0.82	0.57	0.92	0.90	0.97	0.97	0.93	0.93	0.88
F-measure	0.81	0.72	0.76	0.79	0.86	0.88	0.67	0.77	0.59

5 Conclusion

The proposed system classifies websites purely based on the URLs. With this approach of classification the Precision, Recall and F-measure values of 0.7, 0.88 and 0.76 were achieved. This system helps in monitoring the browsing behavior of users inside an organization and can assist the administrators to block unnecessary sites. This classifier can be further extended to identify the abnormal traffic generated from an organization and the classification accuracy can still be improved by using more features.

References

1. Qi, X., Davison, B.D.: Web page classification: Features and algorithms. *ACM Comput. Surv.* 41(2), 1–31 (2009)
2. Kan, M.-Y.: Web Page Categorization without the Web Page. In: *Proceedings of the 13th International World Wide Web Conference*, pp. 262–263. ACM, New York (2004)
3. Kan, M.-Y., Thi, H.O.N.: Fast Webpage Classification using URL Features. In: *Proceedings of CIKM 2005, Germany* (2005)
4. Han, L.W., Alhashmi, S.M.: Joint Web-Feature (JFEAT): A Novel Web Page Classification framework. *Communications of IBMA*, Article ID 73408 (2010)
5. Open Directory Project, <http://www.dmoz.org>

Method to Improve the Efficiency of the Software by the Effective Selection of the Test Cases from Test Suite Using Data Mining Techniques

Lilly Raamesh¹ and G.V. Uma²

¹ Research Scholar, Anna University, Chennai 25

lillyraamesh@yahoo.co.in

² Professor & Head/IST, CSE, Anna University, Chennai-25

Abstract. This paper proposes a new technique for the efficient selection of test cases from the large suite. The main aim of selection of test cases is that the fewer test cases may reduce the testing time and project cost. And also test case selection may increase a test suite's rate of fault detection. An increased rate of fault detection during testing provides earlier feedback on the system under test, allowing debugging to begin earlier, and supporting faster strategic decisions about release schedules. Further, an improved rate of fault detection can increase the likelihood that if the testing period is cut short, test cases that offer the greatest fault detection ability in the available testing time will have been executed.

Keywords: software testing, test case, test suite, data mining, clustering.

1 Introduction

In test automation, software is used to control the execution of tests, the comparison of actual outcomes to predicted outcomes, the setting up of test preconditions, and other test control and test reporting functions. Test automation involves automating a manual process already in place that uses a formalized testing process. Ideally, testing is completely automatic after the specification has been given. The basic objective of test cases is to validate the testing coverage of the application. An optimized set of test cases brings standardization and minimizes the ad-hoc approach in testing.

2 Software Testing

Software testing is the process of validating and verifying a software program. Testing can never identify all the defects within software. Instead, it provides a comparison that compares the state and behavior of the product against principles or mechanisms by which someone might recognize a problem.

A test case has components that describe an input, action or event and an expected response, to determine an application is working correctly.

A test suite otherwise known as a validation suite is a collection of test cases that are intended to be used as input to a software program to show that it has some specified set of behaviors.

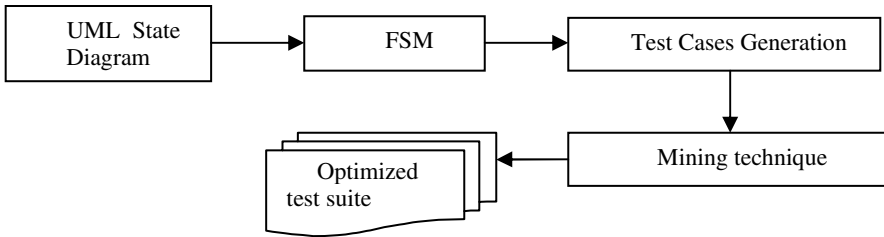
3 Data Mining

Data mining, the extraction of hidden predictive information from large databases, is a powerful new technology with great potential. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions. The automated, prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of decision support systems. Data mining tools can answer business questions that traditionally were too time consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because it lies outside their expectations.

4 System Architecture

The diagram explains the system architecture

- Converting the UML state diagrams into FSM (Finite State Machine)
- Generate test cases from the mined FSM diagrams using available techniques.
- Reduce the test suite size.



5 Converting the UML State Diagrams into FSM

The Unified Modeling Language (UML) is a standard in industry for modelling with several diagrams that express static and dynamic aspects of a system. As one of the behavioural models, state diagram is often used to model the life cycle of certain object.

A FSM (Finite State Machines) A is a quintuple (Q, L, δ, q_0, q) , where Q is a finite set of states of A , L is a finite set of transition labels of A , $\delta : Q \times L \rightarrow Q$ is the transition function relating two states by the transition going between them.

Xholon tool is used for transformation mechanism from state diagrams to FSM models. Xholon is a flexible open source tool for multi-paradigm (UML 2, ABM, SBML, NN, GP, PSys, CA, ...) modeling, simulation, design, execution, and transformation.

6 Test Case Generation from State Machine Diagrams

The method given in [3] is used to create test cases from state machine diagrams which is as follows:

There are three main steps in test case generation, in the first step a predicate is selected on a transition from a UML state machine diagram. In the next step, the selected predicate is transformed into a predicate function. In the third step, test data are generated corresponding to the transformed predicate function.

6.1 Predicate Selection and Transformation

For selecting a predicate, a traversal of the state diagram is performed using depth first (DFS) traversal or breadth first (BFS) traversal to see that every transition is considered for predicate selection. DFS traversal is used here. During traversal, conditional predicates on each of the transitions are looked. Corresponding to each conditional predicate, test data are generated.

Let I_0 consists of all variables that affect a predicate q in the path P of a state machine diagram, then two points named ON and OFF for a given border satisfying the boundary-testing criterion are created. The relational expressions of the predicates are transformed into a function F called predicate function.

6.2 Test Data Generation

The basic search Procedure used for finding the minimum of a predicate function F is the alternating variable method. This method is based on minimising F with respect to each input variable in turn. An initial set of inputs can be randomly generated by instantiating the data variables. Each input data variable x_i is increased/ decreased in steps of S_{x_i} , while keeping all other data variables unchanged. Here, S_{x_i} refers to a unit step of the variable x_i . The exact value of unit step can be defined conveniently. For example, unit step of 1 is normally used for integer values. Unit step can easily be defined for many other types of data such as float, double, array, and pointer and so on.

The refinement is done by reducing the size of the step and comparing the value of F with the previous value. Also, the distance between the data points is minimised by reducing the step size. For each Conditional predicate in the state machine diagram, the test data is generated. The generated test data are stored in a file. A test executor can use these test cases later for automatic testing.

The above said procedure produces a test suite that is of somewhat smaller size. But the size is further reduced by using mining techniques.

7 Mining Techniques for Test Suite Optimisation

This paper proposes a clustering Approach. Clustering is the process of grouping the data into classes or clusters so that object within a cluster has high similarity in comparison to another, but is dissimilar to object in other clusters. It doesn't require the class label information about the data set because it is inherently a data driven approach. It is the process of grouping or abstract object into classes of similar object.

Among all the mining techniques, clustering is the most effective technique, which can be used for test case mining. Clustering analysis helps constant meaningful partitioning of a large set of object based on a “divide and conquer” methodology, which decomposes a large scale system into smaller components to simplify design and implementation. As a data mining task, data clustering identifies cluster or densely populated regions, according to some distance measurement, in a large, multidimensional data. Given a large set of multidimensional data points, the data space is usually not uniformly occupied by the data points. Data clustering identifies the sparse and the crowded places, and hence discovers the overall distributions patterns of the data set.

algorithm TestSuiteOptimisation

```

begin
currentClusterIndex := 0; Sort(clusters,ascending);
while there exists ri such that marked[i] == FALSE do
if currentClusterIndex == k then currentClusterIndex := 0;
list := all tj ∈ clusters[currentClusterIndex];
if Card(list) == 1 then
test := tℓlist;
if there exists rj ∈ requirements where cv[test,rj] == TRUE and
marked[j] == FALSE then nextTest := test;
else
currentClusterIndex := currentClusterIndex + 1;
continue;
endif
else nextTest := SelectTest(list);
endif
if nextTest ≠ 0 then
RS := RS ∪ {nextTest};
foreach rj ∈ requirements where cv[nextTest,rj] == TRUE do
marked[i] := TRUE;
endfor
endif
currentClusterIndex := currentClusterIndex + 1;
endwhile
return RS;
end TestSuiteOptimisation

```

8 Conclusion and Future Work

The proposed approach helps to generate optimised number of test cases which results in manageable size test suite. So that the test suite can be run many times as software is modified and evolved. The automation of the process can be concentrated in the future.

References

1. Wang, J., Mong, W.H., Lee, L., Sheng, C.: A Partition-Based Approach to Graph Mining National University of Singapore, Singapore 117543 (1989), wangjunm, whsu, leeml shengcha @comp.nus.edu.sg
2. Zhang, S., Yang, J., Cheedella, Monkey, V.: Approximate Graph Mining Based on Spanning Trees EECS Dept. Case Western Reserve University (1995), jiong.yang@case.edu, venumadhav@gmail.com
3. Samuel, P., Mall, R., Bothra, A.K.: Automatic test case generation using unified Modeling language (UML) state diagrams Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur 721302, West Bengal, India (1999), philips@cusat.ac.in
4. Wang, X., Guo, L., Miao, H.: An Approach to Transforming UML Model to FSM Model for Automatic Testing School of Computer Engineering and Science Shanghai University China _whitecn, glory (1994), hkmiao@sh
5. Lin, J.-W., Huang, C.-Y.: Huang Analysis of test suite reduction with enhanced tie-breaking techniques *Department of Computer Science, National Tsing Hua University, No. 101, Section 2, Kuang Fu Road, Hsinchu 300(1996), Taiwanu.edu.cn
6. Bader, J., Chaudhuri, A., Rothberg, J., Chant, J.: Gaining confidence in high-throughput protein. *Interaction Networks, Nature Biotechnology* 22(1), 78–85 (2004)
7. Koyuturk, M., Grama, A., Szpankowski, W.: An efficient algorithm for detecting frequent subgraphs in bioioicalnetworks. *Bionformatics* 20, 200–207 (2004)
8. Kuramochi, M., Karypis, G.: Frequent subgraph discovery. In: Proc. of ICDE (2001)
9. Nijssen, S., Kok, J.: A quickstart in frequent structure mining can make a difference. In: Proc of KDD (2004)
10. Yan, X., Han, J.: CloseGraph: Mining Closed Frequent Graph Patterns. In: Proc. of SIGKDD (2003)
11. Yan, X., Han, J.: gSpan: graph-based substructure pattern mining. In: Proc. of ICDM (2002)
12. Bertolino, A.: Software Testing: Guide to the software engineering body of knowledge. *IEEE Software* 16, 35–44 (1999)
13. Quan, Z., Bernhard, Gioranni: Automated Software Testing and Analysis: Techniques, Practices and Tools. In: Proc. of Intl Conf. on System Sciences, HICSS 2007, p. 260 (2007)

A Hybrid Intelligent Path Planning Approach to Cooperative Robots

K. Prasad¹, Vinodh P. Vijayan², and Biju Paul²

¹ Professor, CSE, Valia Koonambakulathamma college of Engg & Tech, Parippally
ksprasad@yahoo.com

² Assistant Professor, Rajagiri School of Engineering and Technology, Cochin-39
vinodhvjayan@yahoo.com, biju_paul@rajagiritech.ac.in

Abstract. The main problem in a distributed cooperative mobile robots when they navigate to a target is that fixing the degree of freedom in movement for an individual robot without breaking the communication link with other robots. In this paper an agent based approach to intelligently coordinate the group members in a cooperative robots are done. Initial steps of optimized link state routing algorithm is used to get the information of other members. And a fuzzy based decision can be taken regarding the movement without breaking the communication link. So this technique ensure , each and every time all the robots are in communication with at least one robot in the group, at the same time system provides maximum degree of freedom for movement.

Keywords: Type2 Fuzzy, distributed cooperative robot, Multi Agent, OLSR.

1 Introduction

Cooperation among the robots are the major design issues in a distributed cooperative mobile robots system. The distributed cooperative means, there is no master slave relation all the robots are autonomous and capable of making decision. But the individual decision of a robot is depend on the other robot in the group, hence it is called cooperative. So they travel to destination as team. Here the communication among the robot is need to be well maintained. And communication is established using some kind of wireless technique in physical layer, hence there will be a specific range for the communication. The problem that may occur is when any of the robot move in different direction as a result of obstacle avoidance etc may go out of wireless range and the word cooperative become meaning less. Hence it is need to be calculate the degree of freedom for robot movement, means the range in meters a robot can move with out breaking the communication with other robot in the group or by maintaining the communication with at least one robot in the group. Once this is calculated individual robot can take turning decision by keeping this distance in consideration. The use of Optimized Link State Routing Protocol (OLSR)[5,6] in this context will provide the next hop or neighbor robot information to the current robot. So once the distance to different neighbor robots are known, a robot can calculate how much it can travel in each direction with out breaking the wireless link. Here we specifically use a communication agent in each robot, which maintain the communication

between robots. Other functions like sensing , collecting link state information, maintain wireless communication, final actuator to robots legs are also designed as agents. Hence a multi agent approach to the system provide a higher level of intelligence.

2 System Diagram

Figure 1 shows the general functional block diagram that map with in a robot. The sensor ring with fuzzy based collision avoidance is used for dynamic obstacle avoidance in target path. Sensor ring is used to identify the obstacle direction, size etc. The wireless LAN maintain necessary connectivity between the robots. The Link information block runs a optimized link state algorithm and finds neighbors using hello packet, an echo packets can be used to find the transmission time. Once the transmission time is obtained, the distance to particular neighbor can be calculated and update in the table. Based on the distance calculated to each neighbor and the prior knowledge of range of wireless system, the degree of freedom for movement of a particular robot with respect to each neighbors can be calculated.

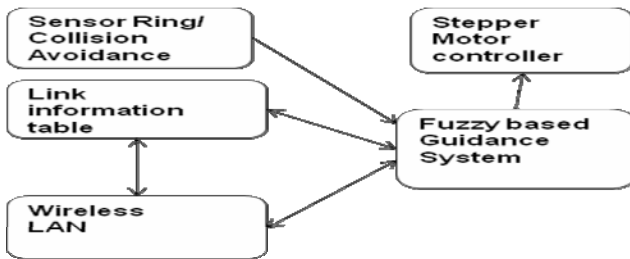


Fig. 1. System block diagram

3 Tools and Techniques Used

3.1 Optimized Link State Routing Protocol (OLSR)

The OLSR[5,6,7,8] is an IP routing protocol optimized for mobile ad-hoc networks. Here OLSR is executed and the neighbor robot information is collected as link state information. Assume let robot R1 and robot R2 is L meter apart and maximum range of available wireless is R meter, then the distance for which R1 is free to move,

$$D_r = (R/2-L/2) \text{ meters.}$$

We consider only half of available wireless range because other robot itself may move away. Next issue is how frequent the link information need to be updated. Simple solution is the time taken to travel D_r , and if the speed or all robot are similar then it is very straight to calculate.

3.2 Fuzzy Logic Controller as Decision Making System

Decision making can be done through Fuzzy [3] logic. we can use either type-1 fuzzy or type-2 fuzzy according to the requirement.

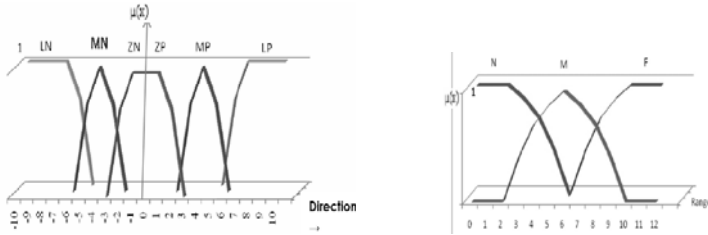


Fig. 2. Membership function for ' Direction and Range'

Figure 2 shows Type-2 fuzzy membership[1,2] function for direction and range. A range has the subsets: Far (F), Medium (M), and Near (N). A direction has the subsets: Large Negative (LN), Medium Negative (MN), Zero Negative (ZN), Zero Positive (ZP), Medium Positive (MP), and Large Positive (LP).

4 System Architecture: Implementation Details

Figure 3 show the multi agent architecture of the system, the simplified diagram representing the multi-agent system using a interface agent. The agents basically interact with the other components of the system by manipulating information on the interface agent. The information on the interface agent may represent facts, assumptions, and deductions made by the system during the course of solving problem.

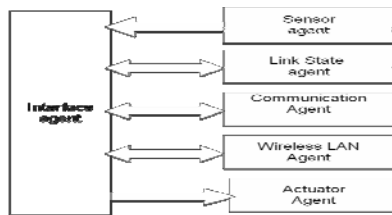


Fig. 3. Multi-agent System

5 Experimentation and Result

Figure 4 shows a working Environment and the dotted Circle with radius Dr_1 and Dr_2 shows the range in which the robot is free to move, The Dr_1 and Dr_2 is calculated for a particular robot from its two different neighbor. Figure 5 show the performance measures for calculation of link updating interval. Which shows that the link interval calculation need to be more frequent as speed of the robot increases.

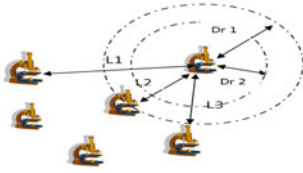


Fig. 4. Working Environment

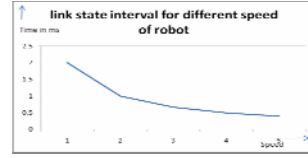


Fig. 5. Link state interval for various speed

6 Conclusion

Cooperative robots can be well organized using optimized link state routing technique and type 2 fuzzy based decision making. This paper basically addressed the problem of degree of freedom for movement of a robot when they are in a group. So using the proposed technique an individual robot will get maximum freedom for moving with in the group without breaking communication link.

References

1. Vijayan, V.P., John, D., Thomas, M., Maliackal, N.V., Vargheese, S.S.: Multi Agent Path Planning Approach to Dynamic Free Flight Environment. *International Journal of Recent Trends in Engineering* 1(1) (May 2009)
2. Poornaselvan, K.J., Kumar, G., Vinodh, T., Vijayan, P.: Agent Based Ground Flight Control Using Type-2 Fuzzy Logic And Hybrid Ant Colony Optimization To A Dynamic Environment. In: *First International Conference on Emerging Trends in Engineering and Technology 2008* (2008)
3. Sotel, M.A., Narnjo, J., Gonzalez, C., Garcia, R., de Pedro, T.: Using Fuzzy Logic in Automated Vehicle Control. In: *Intelligent Transportation Systems*. IEEE, Los Alamitos (2007)
4. Coupland, S.: Type-2 fuzzy control of a mobile Robot. Centre for Computational Intelligence De Montfort University, U.K
5. RFC_3626
6. Abolhasan, M., Hagelstein, B., Wang, J.C.-P.: Real-world performance of current proactive multi-hop mesh protocols (2009), <http://ro.uow.edu.au/infopapers/736/>
7. Chandra, M., Roy, A.: Extensions to OSPF to Support Mobile Ad Hoc Networking. RFC 5820 (March 2010)
8. Ogier, R., Spagnolo, P.: MANET Extension of OSPF using CDS Flooding. RFC 5614 (August 2009)

Towards Evaluating Resilience of SIP Server under Low Rate DoS Attack

Abhishek Kumar¹, P. Shanthi Thilagam¹, Alwyn R. Pais¹, Vishwas Sharma²,
and Kunal M. Sadalkar¹

¹ Dept. of Computer Engineering, NITK-Surathkal, India-575025
{aksinghabsk, santhisocrates, alwyn.pais,
officialmails.kunal}@gmail.com

² Dept. of SERC, IISc. Bangalore, India-560012
vishwas@mmsl.iisc.ernet.in

Abstract. Low rate Denial-of Service, DoS, attack recently emerged as the greatest threat to enterprise VoIP systems. Such attacks are difficult to detect and capable of discovering vulnerabilities in protocols with low rate traffic and it noticeably affects the performance of Session Initiation Protocol, SIP, communication. In this paper, we deeply analysis the resilience of SIP server against certain low rate DoS attacks. For this purpose we define performance metrics of SIP server under attack and non-attack scenarios. The performance degradation under attacks gives a measure of resilience of the SIP server. In order to generate normal SIP traffic and the attacks, we defined our own XML scenarios and implemented them using a popular open source tool known as SIPp. The system under evaluation was an open source SIP server.

1 Introduction

Voice over Internet Protocol (VoIP) is a technology that is reshaping the future of telephony. While enterprise VoIP offers low cost and various functionality it's also opens the door for external threats [1]. Most VoIP services uses the SIP infrastructure, because its simplicity and wide range of features, which makes its service vulnerable. To provide a better VoIP services we need to understand the behavior of VoIP server under different attack rates [2] and its countermeasure. The attack rate is comparable with that of normal traffic and hence it is not flooding the target to cause saturation. This qualifies the attacks to be termed "low rate". These attack strategies are novel approach to launch flooding DoS attack without sending high rate traffic to the victim. These attacks are mixed approaches between flooding and vulnerability attacks by which attackers get advantages after reducing traffic rate.

We use a freely available open source Asterisk SIP server version 1.6.20 [3] & [4]. We populated 8000 unique username and passwords in user account and directory data of server. We automate a registration flooding low rate DoS attack scenario and populate a CSV file of 8000 valid and invalid username and passwords to generate legitimate and illegitimate traffic to SIP application server. Fig. 1 (A) shows the notional diagram of DoS attack target (SIP Registrar) of SIP application server.

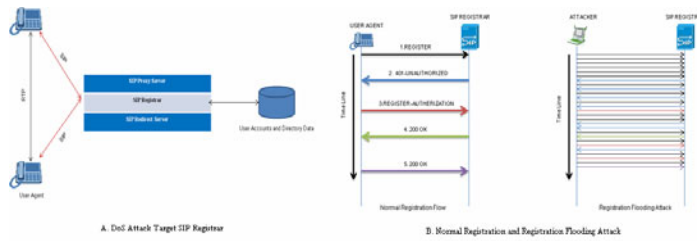


Fig. 1. A. DoS Attack Target SIP Registrar, B. Normal Registration and Registration Flooding Attack

2 Performance Metrics and Experimental Configuration

2.1 Performance Metrics

Our all metrics are SIP based to evaluate the resilience of VoIP SIP server. If any metrics are degraded, it will result to un-resilient SIP server. In our study we define following metrics:

- **Database Lookup Efficiency (R12).** The ratio of total number of first REGISTER request for database lookup by SIP server to total number of responded second “401-Unauthorized or 404- Not found” messages. This will measure the searching efficiency of SIP server for large number of user database in attack and no attack scenario.
- **SIP URI Binding Efficiency (R34).** The ratio of total number of third REGISTER request with MD5 digest to total number of responded fourth “200 OK” messages. This measure defines the digest computation efficiency of SIP server in attack and no attack scenario.
- **Successful Registration Acceptance Rate (SRAR).** Total number of successful URI binding per second. This measure defines the successful acceptance rate in attack and non attack mode. This metric determine the utilization of resources of SIP server.
- **Registration Drop Rate (RDR).** Total number of rejected URI binding per second. This measure defines the registration drop rate in attack and non attack scenario. This metric determine the loss of potential resources of SIP server under flooding attack.

2.2 Configuration Module

In our experiment in Fig. 2, we used SIPp tool [5] with Transport Layer Security (TLS) support as an attack generation and analysis module and an open source version of ASTERISK 1.6 as a SIP server. For attack generation module and client configuration module, we added 8000 illegitimate and legitimate user’s credentials in our CSV file respectively, which is an input of created XML attack scenario. We compiled and run a C code to add 8000 users in SIP database to configure SIP.conf file.

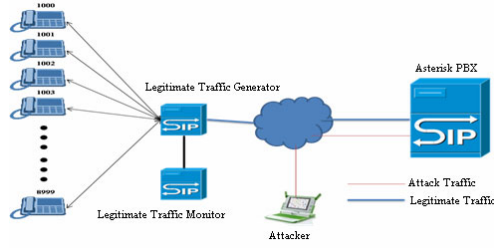


Fig. 2. Experimental Testbed

3 Resilience Result of SIP Server

According to our SIP server specification and experiment, SIP server is capable to process 105 legitimate REGISTER packet/sec for 8000 configured phones in no attack scenario. The results of our experiment show that REGISTER flooding attack significantly degrades the input statistics and derived statistics R12, R34, SRAR and RDR as shown in Fig.3 (A), (B), (C), (D), (E), (F), (G) and (H).

From Fig. 3 (B) & (C), the differences of successful call are almost more than 600 in attack to no attack scenario for each legitimate traffic rate. An attack rate of zero and for legitimate traffic rate 100 REGISTER packet/sec on X-axis shows expected efficiency. In Fig. 3 (E) and (F), R12 and R34 is 100%, for 100 REGISTER packets/sec in no attack scenario. But those metrics significantly decreases for different legitimate traffic rate as well as the attack rate increases. For attack rate 200 packet/sec, the database lookup efficiency and SIP URI binding efficiency of SIP server decreases to more than 94% and 98% respectively, that causes to occur 480 and 160 clients denied by registrar per second respectively. For higher traffic rate, the

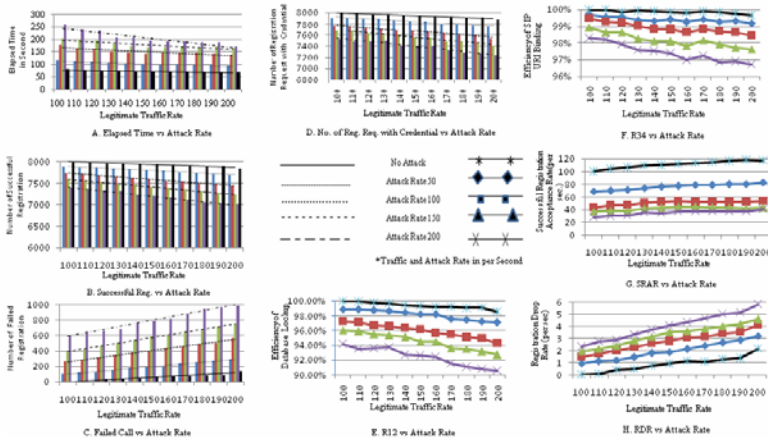


Fig. 3. A. Elapsed Time vs Attack Rate, B. Successful Reg. vs Attack Rate, C. Failed Call vs Attack Rate, D. No. of Reg. Req. with Credential vs Attack Rate, E. R12 vs Attack Rate, F. R34 vs Attack Rate, G. SRAR vs Attack Rate, H. RDR vs Attack Rate.

difference in value of R12 and R34 at attack rate 200 packets/sec changes significantly. Note that the absolute value of 1% loss cause 80 clients (out of 8000 clients) denied to service per second.

In Fig. 3(G), it is clear that, in no attack scenario the SRAR rate is 100 registration per second for 100 REGISTER packet/sec traffic, because for 8000 successful SIP URI binding the server takes 80 seconds. But it decreases to 28 registrations per second, when attack rate is 200 for same benign traffic rate. In Fig. 3(G), we can easily see that the differences in SRAR from no attack scenario to highest attack rate scenario are almost constant (75 registrations/sec) for each traffic rate. From Fig. 4(H), we see that the RDR value increases 0 to 2.5 in no attack scenario to highest attack rate scenario. That means, in each second 2.5 REGISTER requests are discarded by SIP server when attack rate is 200. This drop rate significantly degrades the performance of SIP server, when requests come to server for long time.

4 Conclusion and Future Work

SIP server is not capable of defending itself against low rate DoS attack also. All performance metrics change significantly for low attack rates and server becomes less resilient and more vulnerable to low rate DoS attack. Our study analyzes the performance of different protocol supporting server of VoIP infrastructure under low rate DoS attack. In future we proposed to build a detection tool to detect low rate DoS attack on SIP server and define a protocol state machine to mitigate it.

Acknowledgments. We would like to express our appreciation to Prof. N. Balakrishnan, Dept. of Supercomputer Education and Research Center, Indian Institute of Science Bangalore for valuable support of this research.

References

1. Collier, M.D.: Enterprise Telecom Security Threats, http://download.securelogix.com/library/Enterprise_Telecom_Security_Threats_Draft_10-12-04.pdf
2. Sengar, H.: Overloading Vulnerability of VoIP Networks. In: DSN 2009, pp. 419–428 (2009)
3. Asterisk, The Open Source Telephony Projects, <http://www.asterisk.org>
4. VoIP Wiki- A Reference Guide to all things VoIP, <http://www.voip-info.org/wiki/view>
5. Gayraud, R., et al.: SIPp, <http://sipp.sourceforge.net>

Performance Analysis of a Multi Window Stereo Algorithm on Small Scale Distributed Systems: A Message Passing Environment

Vijay S. Rajpurohit¹ and M.M. Manohara Pai²

¹ Gogte Institute of Technology, Belgaum, India
vijaysr2k@yahoo.com

² Manipal Institute of Technology, Manipal, India
mm.pai@rediffmail.com

Abstract. Stereo vision systems determine the depth from two or more images which are taken at the same time, but from slightly different viewpoints. A novel approach for depth map generation is proposed for a multi-window stereo algorithm on a cluster computing setup using Message Passing Instructions (MPI) to overcome the speed limitations.

Keywords: Stereo vision, Depth map, Symmetric stereo with Multiple Windowing algorithm, Parallel Depth map Generation, Cluster computing setup, Message Passing Instructions.

1 Introduction

Stereo vision refers to the process, which transforms the information of two plane images (2-D) into a description of the 3D scene and recovers depth information in terms of exact distance (Depth map). The Stereo vision in mobile Robot is attained by equipping the Robot with two stereo cameras similar to Human Visual System (HVS). To overcome the speed limitations of vision-based algorithms several hardware software implementations were presented[2][3][4][5][6]. The existing hardware architectures are difficult to implement in real-life environment and not suitable low cost commercial applications due to their high equipment cost and complex configurations. In this work, a novel method for depth map generation is presented for an existing state of the art algorithm SMW[1] to overcome its speed limitations.

2 Depth Map Generation: A Parallel Implementation

The sequential SMW(Symmetric Stereo With Multiple Windowing) algorithm is adaptive and uses multiple windowing approach. For each pixel a correlation is performed with nine different windows. The SMW algorithm has the computational complexity of $O(n^2)$.

2.1 Parallel SMW Algorithm

The parallel algorithm is designed to work with a cluster of computers connected in parallel. Each node in the cluster processes a part of the left-right image pair to generate partial depth map. The server receives the sensor data of right and left images. The received data is divided into equal number of segments which is equal to the number of processors in the cluster (Fig. 1) including the server. Each slave processor has the copy of SMW algorithm implementation. The server sends each pair of image segments to a unique processor in the cluster. Each processor executes SMW algorithm and calculates the partial depth map to each image segment. Partial depth maps are sent to the server and are combined to form the complete depth map of the image pair.

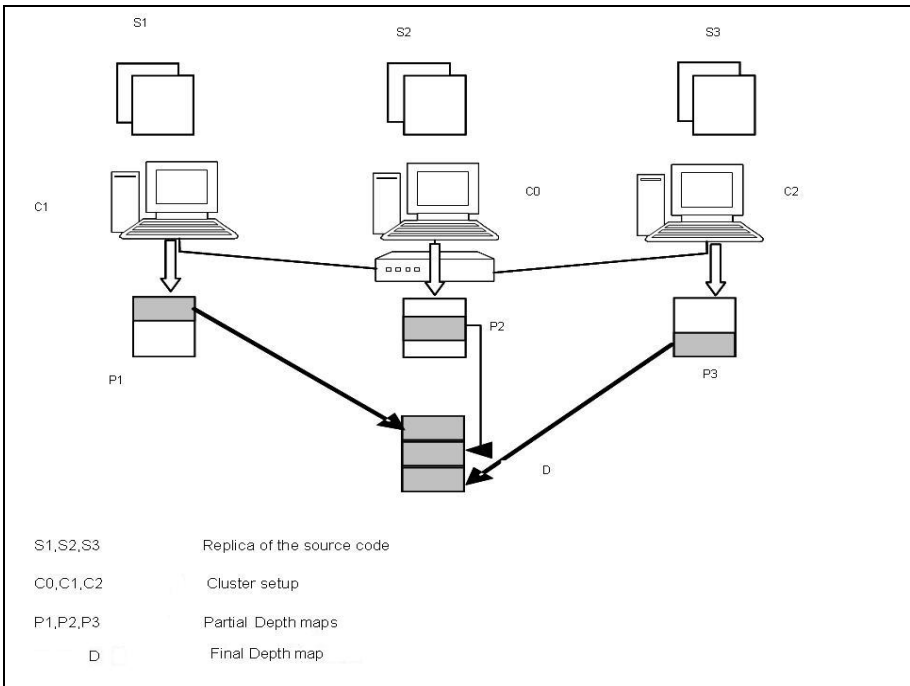


Fig. 1. Cluster computing setup for depth map generation

3 Analysis of the Algorithm

Based on the Amdahl’s law analysis, the speedup achievable by parallel SMW algorithm for a fixed size problem increases initially with the increase in the number of processors and becomes stable (Fig. 2). As the number of processors increases, the parallel overhead increases which reduces the overall performance of the algorithm.

The expected execution time is given by

$$[m^2 / p + n[\log p]\lambda + [\log p](k/\beta)] . \tag{1}$$

Where λ (latency) represent the time needed to initiate a message. β (bandwidth) represent the number of data items that can be sent down a channel in one unit of time. k indicates the data items. n indicates the message length.

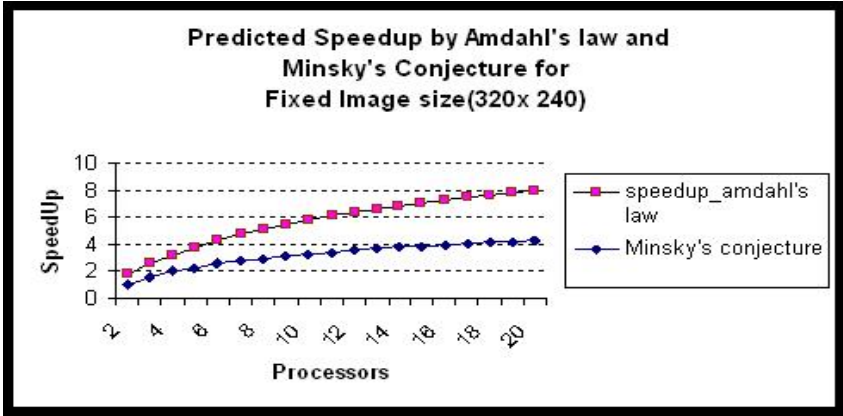


Fig. 2. Predicted speedup of depth map generation

The algorithm is tested for various synthetic and real life stereo images (Fig 3) and a speedup in the band from 1.48 to 1.68 and efficiency from 0.74 to 0.84 on two node cluster. By adding one more node i.e. on three node cluster, speedup varied in the band from 2.28 to 2.58 and efficiency from 0.77 to 0.86. Based on the results it can be concluded that the SMW algorithm can be implemented in parallel to improve speedup, yet the practical speedup is less than the speedup predicted by Amdahl's law.

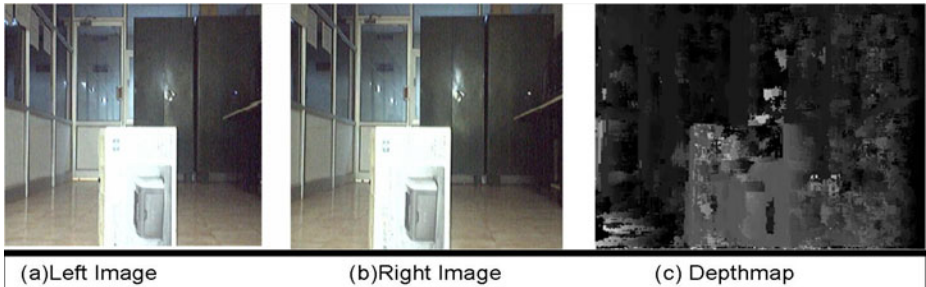


Fig. 3. Depth map creation using stereo vision

4 Conclusion

A parallel multi-window stereo algorithm (SMW) for the depth map generation, on cluster computing setup is developed to overcome the slow response time of the existing multi-window algorithms. Based on the results it can be concluded that the

performance of the algorithm initially improves with increase in number of processors but the observed improvement is less than the improvement expected by the Amdahl's law in the cluster computing environment.

References

1. Fusiello, A., Roberto, V., Trucco, E.: Symmetric stereo with multiple windowing. *International Journal of Pattern Recognition and Artificial Intelligence* 14(8), 1053–1066 (2000)
2. Koschan, A., Rodehorst, V.: Towards Real-Time Stereo Employing Parallel Algorithms For Edge-Based And Dense Stereo Matching. In: *Proceedings of the IEEE Workshop on Computer Architectures for Machine Perception (CAMP 1995)*, Como, Italy, pp. 234–241 (1995)
3. Rosselot, D., Hall, E.L.: Processing real-time stereo video for an autonomous Robot using disparity maps and sensor fusion. In: *Proceedings of SPIE Intelligent Robots and Computer Vision XXII: Algorithms, Techniques and Active Vision*, vol. 5608, pp. 70–78 (2004)
4. Laine, A.F., Roman, G.C.: A Parallel Algorithm for Incremental Stereo Matching on SIMD Machines. In: *Proceedings of the 10th ICPR*, Atlantic City, New Jersey, USA, vol. II, pp. 484–490 (1990)
5. Szeliski, R., Zabih, R.: An experimental comparison of stereo algorithms. *International Journal of Computer Vision* 32(1), 45–61 (1999)

Ant Colony Algorithm in MANET-Review and Alternate Approach for Further Modification

Jyoti Jain, Roopam Gupta, and T.K. Bandhopadhyay

Rajeev Gandhi Technical University, Bhopal, India
jyotijain.phd@gmail.com

Abstract. Mobile ad hoc network(MANET) is the latest application of telecommunication. This is one of the most innovative and challenging area of wireless networking. Ant Colony Algorithm has been used in Mobile Network since long because of isomorphism between them. Pheromone graph and stigmergic architecture of ant colony algorithm are comparable with structure & constraints of communication network. In this paper, A literature survey about the application of ACO is given. In this paper we review ANT algorithm and different approaches proposed by researchers for the improvement of routing performance. In this proposed work, ACO will be used in case of link failure. Path will be discovered by reactive routing, and maintained by periodically generating HELLO messages by all the nodes in the link. All nodes in the link will also find an alternate route for next to next node proactively. By using this method, throughput, and end to end delay parameters can be improved probably the on the cost of increase in the overhead. Overhead will increases in proactive route finding at the same time number of route failure reduces so the bits required in alternate route finding will reduce.

Keywords: MANET, Routing, ACO, AODV, Pheromone.

1 Introduction

Mobile ad hoc network (MANET) is the latest application of telecommunication. This is one of the most innovative and challenging area of wireless networking. MANET is an autonomous collection of mobile users that communicate over relatively band width constrained wireless links. Due to movement of the nodes, network topology may change rapidly and unpredictably over time. In this decentralized network, discovering the route and delivering of data becomes complicated. Various protocols[1] for routing are developed for MANET over past few years. Characteristics of mobile ad hoc network and ant colony algorithm are common. Both of them are self configured, self built and distributed. The concept of ant colony can be utilized to enhance the performance of Routing protocol[2]. In this paper part I give the details of ant colony routing algorithm. part II will discuss about the work done by different researcher. In part III, new concept of application of ant algorithm in MANET routing protocol is discussed. In part IV conclusion is given. Some of the desirable characteristics such as scalability and robustness are exhibited in simple biological systems of insects like ant colony.

Part I

The Ant Colony Optimization Metaheuristic (ACO)

The ACO metaheuristic is a family of multi-agent algorithms to solve combinatorial optimization problems. The representation of the combinatorial problem exploited by the ant-like agents is split in two parts, forward ant and backward ant as shown in figure 1 and figure 2. Each agent is an autonomous construction process making use of a stochastic policy and aimed at building a single solution to the problem at hand, possibly in computationally light.

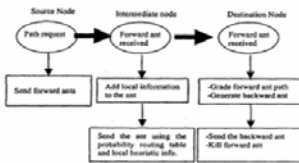


Fig. 1. Working of forward ant

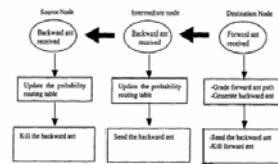


Fig. 2. Working of backward ant

These ant agents are the instrument used to repeatedly sample the solution set according to the bias implemented by pheromone variables. The colony's management tasks involve the generation of ant agents, the scheduling of optimization procedures other than the ants (e.g. local optimizers) and whose results can be used in turn by the ants, and the management of the pheromone, in the sense of modifying the pheromone values.

Part II

Recently, some ACO-based routing algorithms have emerged for the employment in MANET. Some of the work done in this field by the researchers is discussed here. Schoonderwoerd et al.[3] designed ABC (Ant-Based Control) for circuit-switched telephone networks. In this method ACO were able to optimize performance in the network by balancing the load in the network. The AntNet algorithm introduced by Di Caro and Dorigo [4] in 1998 for adaptive routing in packet switching wired networks. AntNet, was highly adaptive to network and traffic changes, uses lightweight mobile agents (called ants) for active path sampling, is robust to agent failures, provides multipath routing, and automatically takes care of data load spreading. Mobile Ants Based Routing (MABR)[5] by M. Heissenbüttel, & Braun is introduced routing algorithm for MANET's inspired by social insects. a short hello message was used to announce its presence and position to update routing table from the corresponding node in highly dynamic networks. The advantages include the ability to react and deal quickly with local and global changes. Ant Colony Based Routing (ARA) [6] by Mesut Günes, gives route discovery mechanism similar to other algorithms such as AODV and DSR. Liu Zhenyu developed EARA-CG (Emergent Ad hoc Routing Algorithm) and EARA-QoS(Emergent Ad hoc Routing Algorithm)[7], In these

algorithm, the principle of swarm intelligence is used to evolutionally maintain routing information. This algorithm provides Congestion control. In ADRA by X. Zheng et al.[8] 2004, The ants deposit simulated pheromones as a function of their distance from their source node, the quality of the link, the congestion encountered on their journey. The ADRA system is shown to result in fewer call failures.

Part III

Researchers have done lot of work in improving routing. Most of the work is done at the level of Monitoring and Admission control. This algorithm is of flexible nature and can be used for further modification to improve routing protocol. Most of the researchers developed a technique to find a new path from the source node in case of route failure. This increases end to end delay during the transmission. Work on alternate route finding from the nearer of the faulty node is still untouched. In this proposed work, In place of selecting a new path from source node, ACO can be used in finding new path from a precursor of the failed node. In this proposed work path will be discovered by reactive routing, and maintained by periodically generating HELLO messages by nodes in the link. All nodes in the link will also find an alternate route

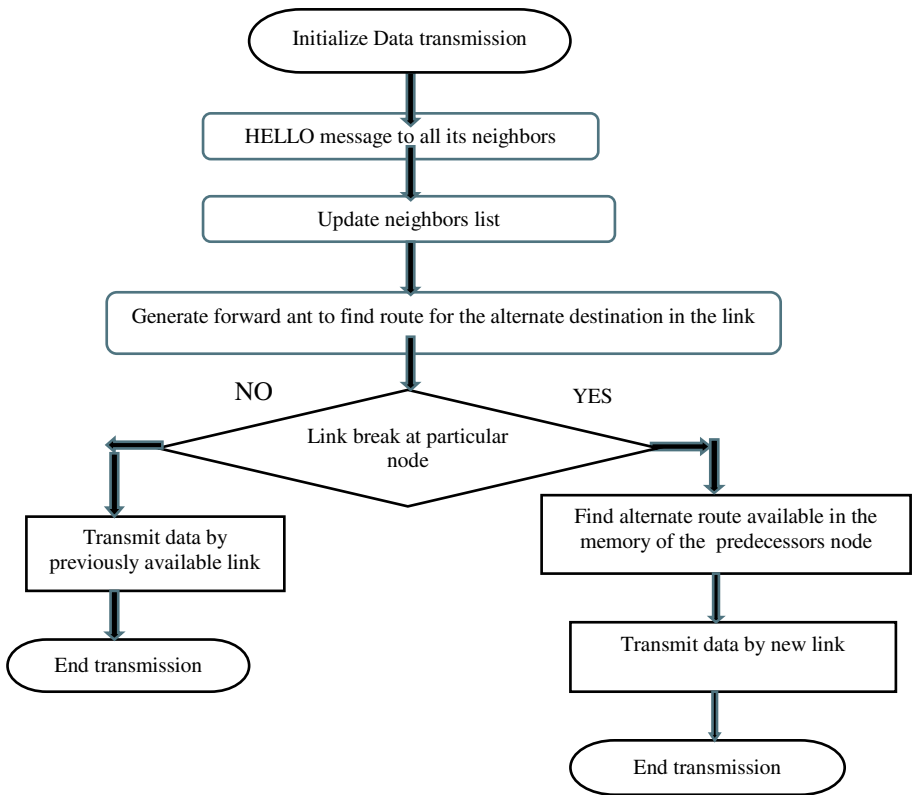


Fig. 3. Algorithm for proposed work

for next to next node proactively. In case of route failure alternate route in the memory of these nodes can be used. Following flow chart presenting proposed algorithm.

Ant colony algorithm can be used to modified maintenance procedure of AODV protocol. End to end delay will be reduced and hence throughput of the network will increased. By local repairing the path in case of link failure the overhead during the communication will also reduce which helps in maintaining the scalability. In place of sending error message to the source & destination during route failure, fresh route request can be generated by node just previous to the lost node. By reducing overhead better efficiency and scalability can be achieved. Overhead can be reduced by reducing number of bits in control packets like forward ant backward ant and Hello packets. Some speeding techniques can be applied for the algorithm from the closer point to the optimum solution.

2 Conclusion

In this paper we analyzed basic ant colony algorithm and its application in the field of Mobile ad-hoc network to improve the qualitative characteristic of routing process. Flexibility of the ant colony algorithm provides numerous approaches to modify mobile ad-hoc network. This flexibility can also be used at the maintenance level. Alternate route finding in case of route failure can be done from node nearer (just before) to the faulty node. The use of location information at the nodes is used as a heuristic parameter. This resulted a significant reduction of time needed to establish a route from a source to a destination which is important for a reactive routing algorithm. Application of ant colony algorithm at the maintenance level can provide a new scope in the optimization of routing protocol.

References

1. Perkins, C.E., Royer, E.M.: Ad-hoc on-demand distance vector routing. In: Proc. Of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (1999)
2. Dorigo, M., Gambardella, L.: Ant colony system: a cooperative learn approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation* 1(1), 53–66 (1997)
3. Schoonderwoerd, R., Holland, O., Bruten, J., Rothkrantz, L.: Ant-Based Load Balancing In Telecommunications Networks. *Adaptive Behavior* (1996)
4. Di Caro, G., Dorigo, M.: AntNet: distributed stigmergetic control for comm. networks. *Journal of Artificial Intelligence Research* 9, 317–365 (1998)
5. Heissenbuttle, M., Braun, T.: Ants-Based Routing in Large Scale Mobile Ad-Hoc Networks. In: *Kommunikation in Verteilten Systemen, KiVS* (2003)
6. Gunes, M., Sorges, U., Bouazizi, I.: ARA-The Ant-Colony Based Routing Algorithm for MANETs. In: *International workshop, WAHN 2002, British Columbia, Canada, August18-21* (2002)
7. Liu, Z., Kwiatkowska, M.Z.: Costas Constantinou:A Biologically Inspired QoS Routing Algorithm for MANET. In: *19th International Conference AINA 2005*, pp. 426–431 (2005)
8. Zheng, X., Guo, W., Liu, R.: An ant-based distributed routing algorithm for ad-hoc networks. In: *ICCCAS*, pp. 412–417. *IEEE, Washington, DC* (2004)

RDCLRP-Route Discovery by Cross Layer Routing Protocol for Manet Using Fuzzy Logic

Mehajabeen Fatima, Roopam Gupta, and T.K. Bandhopadhyay

R.G.P.V., Bhopal, M.P., India

mehajabeen.fatima@gmail.com, roopam_1710@yahoo.co.in,
bandotushar_bando@hotmail.com

Abstract. MANET links have dependency on node mobility and node battery power causes degradation of network's performance and it could be alleviated by allowing cross layer interaction in the protocol stack. In this paper Route Discovery by Cross Layer Routing Protocol (RDCLRP) is proposed for route discovery and route maintenance and AODV is modified. Instead of periodic hello message of AODV, non periodic hello warning message is used for route maintenance and route discovery in RDCLRP. HELLO warning message is used to give warning when link is going to break before link breakage when it touches the threshold and it is broadcasted whenever it is required and not periodically. With the help of Hello warning message new route can also be discovered before link breakage. In this paper only HELLO warning message interval is optimized with the help of fuzzy logic. This reduces overhead and power consumption, in turn it reduces delay in search of new route. The result illustrates the improvement of optimized RDCLRP over the basic AODV. Network performance such as throughput, delay, jitter, overhead and battery power consumption can be improved significantly with this method.

Keywords: MANET, AODV, RDCLRP, Cross layer approach, Hello warning message, Hello warning message interval (*HI*).

1 Introduction

A Mobile Ad-Hoc Network (MANET) is a collection of wireless mobile nodes forming a temporary network without using any centralized access point, infrastructure, or centralized administration. Some or possibly all of these nodes are mobile. This network can be deployed rapidly and flexibly. The person-to-person, person-to-machine or machine-to-person can communicate instantaneously, immediately and easily with MANET.

Cross-layer design allows exchange of information of a layer with any other possibly non-adjacent layer in the protocol stack to improve the network performance. This helps in reduction of delays, overhead, congestion, it saves bandwidth e.t.c.

In Part 2 RDCLRP-new proposal for route maintenance and route discovery is given. In this proposal AODV is modified using Cross layer approach and Fuzzy Logic. In Part 3 results and graphs are given. Then conclusion is drawn.

2 RDCLRP

In basic AODV the path is searched when it is required to send the packets. The path is searched by RREQ and RREP messages, after searching of path, route has to be maintained. Route is maintained by periodic hello messages. Every path which participate in transmission of packets acts as active path or active route and every node which participate in transmission of packets acts as active node send hello message shows its presence in the transmission range and if a node does not send the hello message for long time then it is assumed that it is die. Then a RERR message is sent back to source. Then source will reinitiate the route discovery procedure. This increases delay and it is not good for real time traffic. Hello message sends after every one second in AODV it increases overhead and traffic inturn increases possibility of congestion, consumes bandwidth unnecessarily and most important consumes battery power[2]. So in [3], it is proposed that instead of this periodic hello message, non periodic hello message can be send as hello warning message. If speed, transmission range varies then a node will go out of transmission range in different times, so Hello warning message(HWM) should be given adaptively. So a RDCLRP is proposed in which adaptive HWM is used. This HWM in RDCLRP is send only once whenever it is required and its interval is calculated on the basis of speed, transmission range, battery power using fuzzy Inference system. This saves bandwidth and battery power since transmission of packet consumes maximum of energy[5]. In **RDCLRP**, Hello Warning Message Interval(HI) is made adaptive w.r.t. speed, transmission range, battery power. When a node is going out of transmission range or if its battery power is going to zero called critical node. Then critical node will send a HWM to warn the neighbour nodes that the link is going to break after sometime. So warning has to be given before link breakage. Neighbour nodes listen the warning message, so neighbour nodes reply to critical node. Critical node search the route locally and try to find out the route before link breakage. So instead of sending hello message periodically, hello sends as HWM only once whenever it is required. This HWM broadcast before link breakage. The node which reply first will participate in new route. The critical node will reply service replicate message consist of new node id to its previous node and old route will replace by new route before link breakage. So a new route is discovered through HWM locally. This reduces battery power consumption, delay and overhead simultaneously.

HWM depends on prediction of position of node and remaining battery power. Position of a node depends on transmitted power, this transmitted power is present at Physical layer[4]. The value of transmitted power and corresponding transmission range will call on network layer from physical layer. The position of a node also depends on speed of a node and this speed is available at physical layer and will be call on network layer - cross layer approach.

The proposed protocol is based on sharing of MAC & physical layer information at Network layer[3]. These parameters are transmit power, transmission range, full charge battery capacity, remaining battery power of node at time t , speed of node at time t , direction of node at time t .

In this paper Fuzzy inference system is used to make Hello warning message interval (HI) adaptive. Low, medium and high values are taken for speed, transmission

range, battery power for making of fuzzification rules. Membership functions for different parameters consider different ranges. Rules for adaptive *HI* are given in Table 1.

Table 1.

S	B_p Low(L)			B_p Medium(M)			B_p High(H)		
	T_r			T_r			T_r		
	L	M	H	L	M	H	L	M	H
L	L	L	L	H	M	H	H	H	H
M	L	L	L	M	M	H	M	M	H
H	L	L	L	L	L	M	L	L	M

3 Results and Graphs

To analyze RDCLRP performance, fifty times simulation was run on Qualnet network simulator. Scenario area is 1500x1500 sqm, number of nodes is 30. The communication distance of nodes is 300m. Nodes move every second. So, the speeds of nodes from 0mps to 18mps are taken. The traffic model has 4000 packets. The packet length is 512bytes. The transporting protocol is UDP. The MAC protocol is 802.11 and the channel rate is 11Mbps. Mobility model is random way point and the work frequency is 2.4GHz. The simulation period is 400s. *HI* is calculated by fuzzy logic. In this paper calculations are made for *HI* for high transmission range and high battery power. Here, basic AODV and RDCLRP are compared for throughput, delay, overhead and number of hello messages. Fig. 1 to 4 drawn for $T_r = H, B_p = H$.

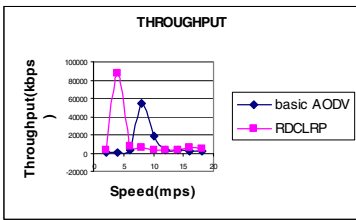


Fig. 1. Throughput Vs. speed

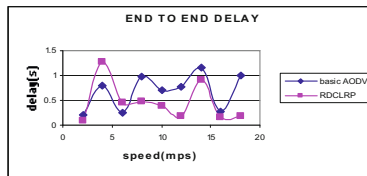


Fig. 2. End to end delay Vs. speed

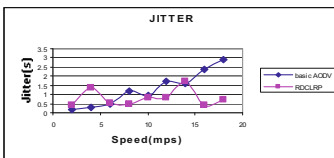


Fig. 3. Jitter Vs. speed

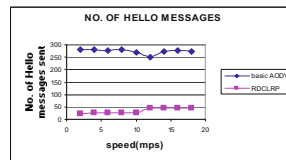


Fig. 4. No. of Hello messages Vs. Speed

Fig. 1 shows that throughput has remarkable improvement specially at lower speeds. This shows that if a node is moving slowly then there is no need to send Hello packet again and again. This saves battery power. Fig. 2, 3 shows reduction in delay and jitter which is best suited to real time traffic. The hello messages or control packets reduces in RDCLRP inturn reduces overhead, power consumption, undesirable traffic again in turn it reduces congestion. Less overhead is favourable for non real time traffic. This saves battery power consumption.

4 Conclusion

From above graphs it is clear that delay, jitter and overhead are reduced in RDCLRP in comparison of basic AODV. The no. of hello messages reduced considerably in RDCLRP comparative to basic AODV. This reduces power consumption and congestion indirectly. And this improves throughput as shown in graphs. RDCLRP reduces control overhead considerably as very less number of hello messages are sent and they are sent as warning message which discovers new route without any increase in overhead i.e. without addition of new control packets. This reduces delay also. Concluded that RDCLRP can be used for both real time and non real time traffic.

References

1. Stojmenovic, I., Lin, X.: Power-aware Localized Routing in Wireless Networks. IEEE, Los Alamitos (2000), ISBN: 0-7695-0574-0/2000
2. Zhang, X., Gao, X., Shi, D.: Lifetime-aware Leisure Degree Adaptive Routing Protocol for Mobile Ad hoc Networks. In: Proceedings of the Third International Conference on Wireless and Mobile Communications (ICWMC 2007). IEEE, Los Alamitos (2007), ISBN: 0-7695-2796-5/07
3. Fatima, M., Gupta, R., Bandhopadhyay, T.K.: Route maintenance improvement By Warning Hello In AODV Of MANET Using Cross Layer Approach. In: IEEE International Conference on advance Computing and communication techniques, Panipat, Haryana, India, October 30 (2010)
4. Cano, J.-C., Kim, D.: Investigating Performance of Power-aware Routing Protocols for Mobile Ad Hoc Networks. In: Proceedings of the International Mobility and Wireless Access Workshop (MobiWac 2002). IEEE, Los Alamitos (2002), ISBN: 0-7695-1843-5/02
5. Marie feeney, L., Nilson, M.: Investigating the energy consumption of a wireless network interface in an adhoc networking environment. In: Proceedings of IEEE INFOCOM, Anchorage, AK (2001)

A New Classification Algorithm with GLCCM for the Altered Fingerprints

R. Josphineleela¹ and M. Ramakrishnan²

¹ Research scholar, Sathyabama University, Chennai
ilanleela@yahoo.com

² Department of Information Technology, velammal Engineering College, Chennai
ramkrishod@gmail.com

Abstract. Fingerprints have always been the most practical and positive means of Identification. Offenders, being well aware of this, have been coming up with ways to escape identification by that means. Erasing left over fingerprints, using gloves, fingerprint forgery; are certain examples of methods tried by them, over the years. Failing to prevent themselves, they moved to an extent of mutilating their finger skin pattern, to remain unidentified. This article is based upon obliteration of finger ridge patterns .In this article, we propose a new classification algorithm GLCCM (**Gray Level Correlation Co-occurrence Matrix**) algorithm for altered fingerprints classification. It is based on the fact that altered fingerprint image is composed of regular texture regions that can be successfully represents by co-occurrence matrix. So, we first extract the features based on certain characteristics and then we use these features for classifying altered fingerprints.

Keywords: altered fingerprints, enhancement, classification and co-occurrence matrix.

1 Introduction

Fingerprints are used to Identify Objects, Compare Objects Remotely and Test an Object for Changes. Since fingerprints are smaller, they are very useful as stand-ins for remote objects. The primary purpose of fingerprint alteration is to evade identification using techniques that vary from abrading, cutting, accidents etc. In this paper, we propose new classification algorithm GLCM(Gray level co-occurrence matrix) for identifying criminals by purposely altering their fingerprints .Gray level co-occurrence matrix for altered fingerprint classification consists of Pre-processing, Feature Extraction, Classification.

Motivation: The motivation behind the work is growing need to identify a person for security. The fingerprint is one of the popular biometric methods used to authenticate human being. The proposed enhancement method provides reliable and better performance than the existing technique.

Contribution: In this paper we used altered and natural Fingerprint enhancement using adaptive wiener 2 filter coefficient with the help of MATLAB codes. Organization: This paper is organized into the following. Pre-processing, Feature Extraction, Co-occurrence matrix (CM), classification, Result finally the Conclusions.

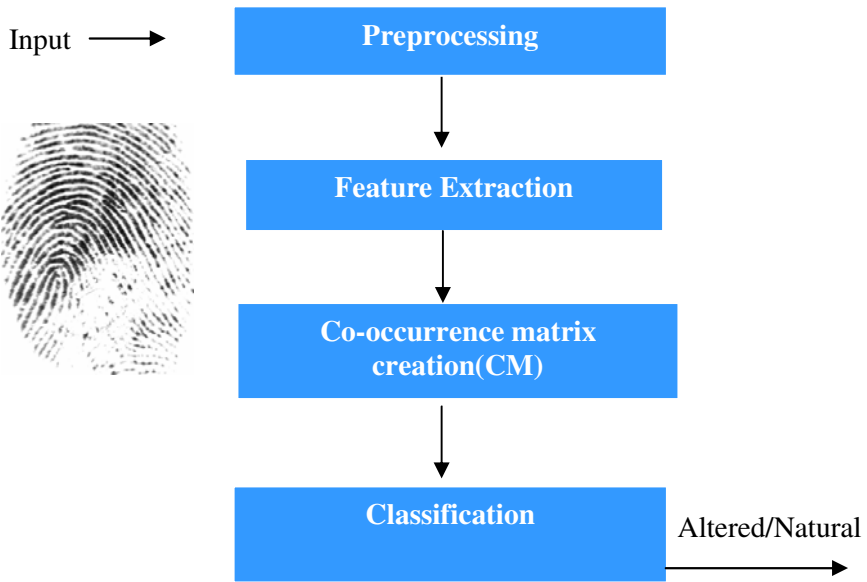


Fig. 1. Block diagram of Altered Fingerprints Classification

2 Preprocessing

Preprocessing consists of histogram equalization, Image enhancement, binarization

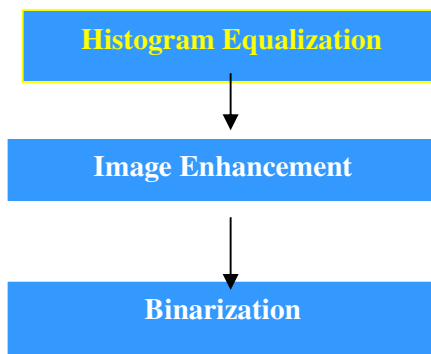


Fig. 2. Block diagram of Pre-processing

2.1 Histogram Equalization

Because there are many noises in original fingerprint image, an image enhancement algorithm such as histogram equalization is usually applied to reduce the influence of the noise in fingerprint image and to emphasize the ridges structures of damaged fingerprint patterns.

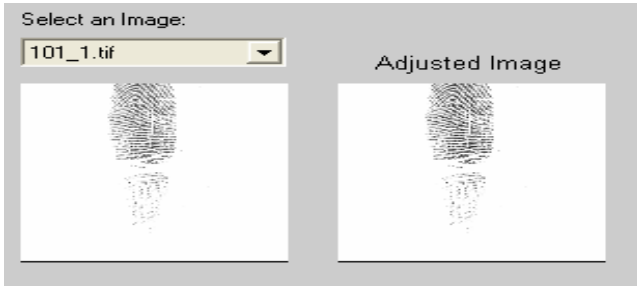


Fig. 3. Altered fingerprint noise reduction experiment samples

2.2 Image Enhancement

To decrease the noise and improve the definition of ridges against the valleys. For this enhancement the adaptive filter is applied separately to each block of the image

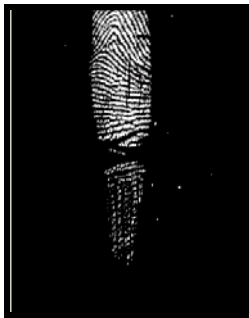


Fig. 4. Altered fingerprint enhancement experiment samples

2.3 Binarization

Converts the enhanced image into binary image adaptive threshold is used, pixel values above the threshold are assigned to 1 and pixel values below the threshold are assigned to 0.

3 Feature Extraction

In the feature extraction phase, the feature region is divided into 8*8 blocks and their directional values are estimated. The directional field is calculated using Gabor

filter and the location of the pixels; orientation and ridge type are stored in the co-occurrence matrix.

4 Co-occurrence Matrix Creation

For Co-occurrence matrix creation the following algorithm is used GLCMA (Gray level co-occurrence matrix algorithm) is used.

```

GLCMA(ip,jp)
{
if (ridge type==2)or(ridge type==4)
{
//termination minutiae has been found
store termination minutiae(it,jt);
}
else if(ridge type==1) or (ridge type==3)
{
//bifurcation minutiae may exist
if( point is valid(it,jt))
store bifurcation minutiae(it,jt);
}
else
{
delete the false pixel position
}
//Creation of co-occurrence matrix
} //end of algorithm

```

5 Classification

For the classification task, the GLCCM (Gray level Correlation coefficient Co-occurrence matrix) algorithm is used. The co-occurrence matrix of the input image is compared with the template image matrix,

x is the template gray level image,

\bar{x} is the average gray level in the template image

y is the source image section

\bar{y} is the average gray level in the source image

N is the number of pixels in the section image

($N = \text{template image size} = \text{columns} * \text{rows}$)

The value cor is between -1 and $+1$, with larger values representing a stronger relationship between the two images.

$$cor = \frac{\sum_{i=0}^{N-1} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=0}^{N-1} (x_i - \bar{x})^2 \cdot \sum_{i=0}^{N-1} (y_i - \bar{y})^2}} \tag{1}$$

6 Result and Discussion

In this paper, we have proposed altered fingerprint classification based on the co-occurrence matrix approach. The above analysis, with experiment result support, shows that the proposed design is of better performance and security. The features extracted from the matrices can well characterize the regular texture of altered fingerprint images and Natural fingerprint images.

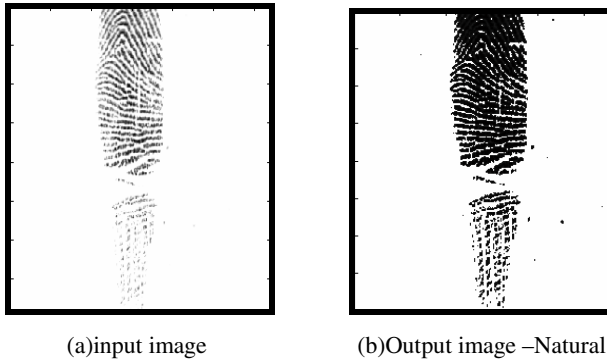
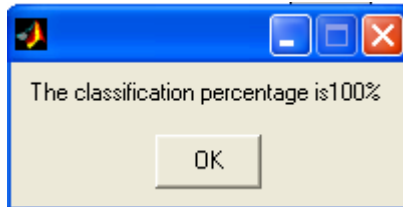


Fig. 5. Sample Output of fingerprint image

7 Conclusion

In this paper we have prompted **GLCCM** algorithm for altered fingerprints classification. Experiments were conducted on both real world altered fingerprints. The proposed method **GLCCM** gives better classification.



References

1. Jain, A.K., Feng, J.: Latent fingerprint Matching. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (2010)
2. Jain, A.K., Ross, A.: Detecting Altered Fingerprints. In: *ICPR 2010, Istanbul, Turkey* (2010)
3. Bhuyan, M.H., et al.: An effective method for fingerprint classification. *International Arab Journal of e-technology* 1(3) (January 2010)
4. Ne'ma, B., Ali, H.: Multipurpose code generation using fingerprint images. *International Arab Journal of Information Technology* 6(4) (October 2009)
5. Ravi, J., et al.: Fingerprint Recognition Using Minutia Score Matching. *International Arab Journal of e-Technology* 2(35-42), 89 (2009)
6. Yazdi, M., Gheysari, K.: A new approach for the fingerprint classification based on gray-level co-occurrence matrix. *World Academy of Science Engineering and Technology* 30 (July 2008)
7. Wang, X., Long, H., Su, X.: Method of Image enhancement based on differential evolution Algorithm. In: *International Conference on Measuring Technology and Mechatronics Automation* (2010)
8. Sihalath, K., Choomchuay, S.: Fingerprint image enhancement with second derivative Gaussian filter and directional wavelet transform. In: *2010 Second International Conference on Computer Engineering and Applications* (2010)
9. Kukula, E.P., Blomeke, C.R., Modi, S.K., Elliott, T.J.: Effect of Human Interaction on Fingerprint Matching Performance, Image Quality, and Minutiae Count. In: *International Conference on Information Technology and Applications*, pp. 771–776 (2008)
10. Ji, L., Yi, Z.: Fingerprint Orientation field Estimation using Ridge Protection. *The Journal of the Pattern Recognition* 41, 1491–1503 (2008)
11. <http://bias.csr.unibo.it/fvc2000/download.asp>
12. <http://bias.csr.unibo.it/fvc2002/download.asp>
13. <http://bias.csr.unibo.it/fvc2004/download.asp>
14. <http://www.neurotechnologija.com/download.html>

Footprint Based Recognition System

V.D. Ambeth Kumar¹ and M. Ramakrishan²

¹ Research Scholar, Sathyabama University, Chennai, India
ambeth_20in@yahoo.co.in

² Professor and Head, Velemmal Engineering College, Chennai, India
ramkrishod@gmail.com

Abstract. This paper proposes a new method of personal recognition based on footprints. In this basis method, an input pair of raw footprints is normalized, both in direction and in position. The foot image is segmented and its key points are located. The foot image is aligned and cropped according to the key points. The footprint image is enhanced and resized. Sequential Modified Haar transform is applied to the resized footprint image to obtain Modified Haar Energy (MHE) feature. The sequential modified Haar wavelet can map integer-valued signals onto integer-valued signals abandoning the property of perfect reconstruction. The MHE feature is compared with the feature vectors stored in database using Euclidean Distance. The accuracy of the MHE feature and Haar energy feature under different decomposition levels and combinations are compared. More than 88.37% accuracy achieved from the proposed MHE feature.

Keywords: Footprint, Sequential Haar Transform network, Heel Shape.

1 Introduction

Innumerable automated biometrics based identification and verification systems have been developed [1]. The biometrics features derived from fingerprints [2][3] faces [2], palm print[4], irises, retinas, a speaker's voice, and perhaps a variety of other characteristics. The systems are now used in a wide range of environments, such as law enforcement, social welfare, banking, and various security applications. Practical methods for automated biometrics based identification have not yet been developed for use with unconstrained subjects. The main Problem in automatic personal identification is how to verify the sampled feature against the registered feature with high reliability. Some systems for personal identification use fundamental biometric features derived from fingerprints and irises. To acquire the feature, subjects must input their biometrics to a sensor. Signatures [5] and speakers' voices [6] are also features useful for verification; however, obtaining these features requires the subjects' cooperation. Automatic face recognition based on vision technique can be a promising method since it can work without any help user [7] but changing elimination, occlusion and hair style change are still very difficult problem. Footprint based recognition is another emerging method which does not need any operation from user. A high recognition rate by verifying raw footprints directly is difficult to obtain, because people stand in various positions with different distances and angles between the two feet. To achieve robustness

in matching an input pair of footprints with those of registered footprints, the input pair of footprints must be normalized in position and direction. Such normalization might remove useful information for recognition, so geometric information of the footprint prior to normalization into an evaluation function for personal recognition decision is included. In this paper, we propose a footprint-based personal recognition method and test its reliability. Footprint texture features are usually extracted using transform-based method such as Fourier Transform [8] and Discrete Cosine Transform [9]. Besides that, Wavelet Transform [10] is also used to extract the texture features of the footprint. In this work, a Sequential modified Haar Wavelet is proposed to find the Modified Haar Energy (MHE) feature. Fig 1,2 shows the proposed footprint identification using Sequential modified Haar Transform.

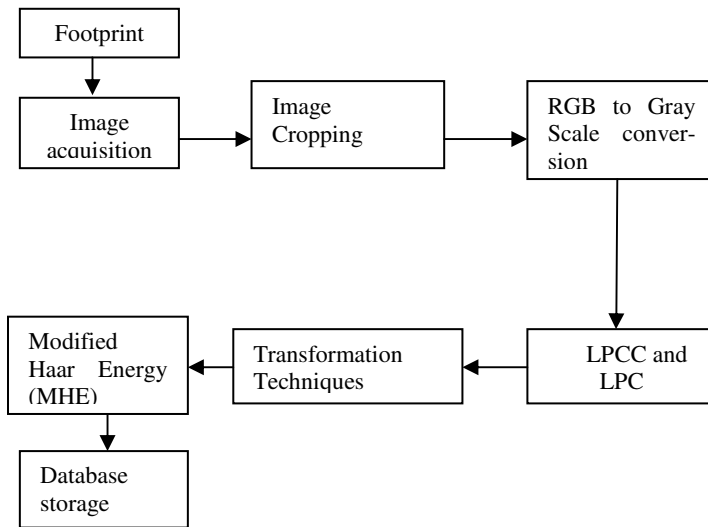


Fig. 1. Footprint identification using Sequential Modified Haar Wavelet

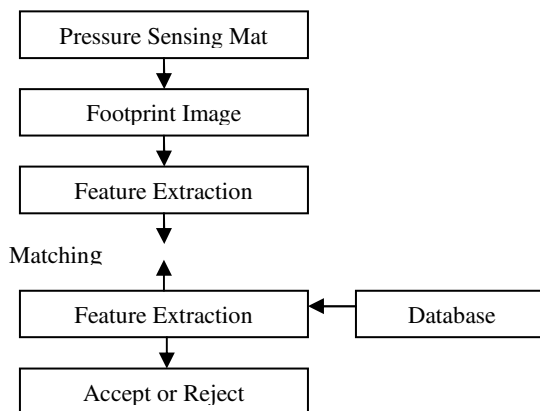


Fig. 2. Footprint identification using Feature Extraction

In this work, hundred images of left leg of 30 people are acquired using a pressure sensor. The foot image is aligned according to the keypoints and it is cropped. The energy features of the footprint are extracted using Sequential modified Haar Wavelet. The Modified Haar Energy (MHE) is represented using feature vector and compared using Euclidean Distance with the threshold values stored in the database. Fig 2 shows the process of footprint identification using feature extraction. The footprint image captured using pressure sensor is extracted for its features. The already existing features of footprint images stored in database are compared. Then, the footprint image is accepted or rejected based on the comparison.

2 Proposed Method

To obtain a footprint image, we use a pressure-sensing mat [11]. The pressure distribution of a footprint is represented as a monochrome image $f(x, y): -X/2 < x \leq X/2, -Y/2 < y \leq Y/2$ (Where X and Y are widths of the mat). Details of the sensor and method for acquiring the footprint are described in the next section. In this section, we describe a normalization procedure and a recognition method.

2.1 Normalization

2.1.1 Image Acquisition

Footprints are normalized both in direction and in position, so that the subject does not need to always stand in the same position as when first recorded. The normalization procedure is as follows:

Firstly, the position of the center of mass of the whole footprint image is calculated and moved to the center of the frame. The image of the footprint remains otherwise unaltered.

$$\left(\frac{\sum_x x \sum_y f(x, y)}{\sum_{x,y} f(x, y)}, \frac{\sum_y y \sum_x f(x, y)}{\sum_{x,y} f(x, y)} \right)$$

Secondly, a scanning line method is employed which gives a number for each segment of the footprint. When the scanning line meets a segment, a number is given to the segment. Each segment is copied into two new frames, the left footprint image L(x,y) or the right footprint image R(x,y), by considering the sign of the x-coordinate of the center of mass of each segment. For example, if the x-coordinate of the center of mass of a segment has a negative sign, the segment is assumed to belong to the left foot.

Thirdly, the center of mass of the left footprint, including all segments of the left foot, is calculated and moved to the center of the frame. The left footprint remains otherwise unaltered. The center of mass of the left footprint is indicated by “+” and that of the right footprint by “x”, which is shown in the diagram Fig.3.

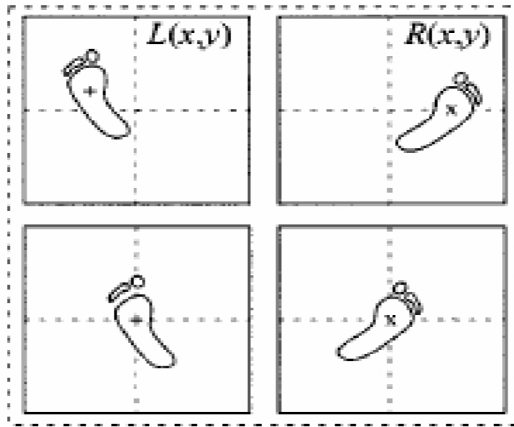


Fig. 3. Calculation of the center of foot

Fourthly, by using an orthogonal transformation (x, y) is normalized in direction. First, we obtain the covariance matrix summation of $L(x, y)$

$$H_L = \begin{pmatrix} \sum_{x,y} L(x, y)x^2 & \sum_{x,y} L(x, y)xy \\ \sum_{x,y} L(x, y)xy & \sum_{x,y} L(x, y)y^2 \end{pmatrix}$$

The principal components (the eigenvectors ordered by highest to lowest corresponding eigen value), $eL1$ and $eL2$, are extracted. Assuming that the footprint is an oval shape, $eL1$ and $eL1$ show the major and minor axes, respectively. The diagonal matrix $uL=(eL1eL2)$ is defined from two eigenvectors $eL1$ and $eL2$. Image $L(x,y)$ is rotated so that the first eigenvector $eL1$ fits the vertical direction by the orthogonal transformation using uL . The same procedures, from Step 3 to Step 4 (Shown in Fig. 3, Fig. 4), are carried out for the right footprint image $R(x,y)$.

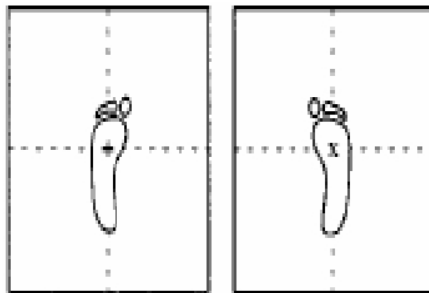


Fig. 4. Normalization in direction

Finally, the footprints of both feet are normalized in position. Both left and right footprints are integrated into a reconstructed image, $I(x,y)$, with a fixed distance, T (Shown in Fig. 5). The center of mass of the two feet is the center of the frame.

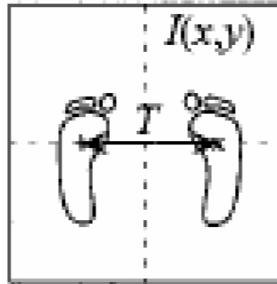


Fig. 5. Reconstructed whole footprint image

2.1.2 Image Cropping

After acquiring the foot image, the next step is to crop the image. Cropping can be done by keypoints determination and extracting the image. The foot image is cropped to extract the heel shape as the intensity is highest in the heel portion. Fig. 6 shows the extracted footprint image using the proposed method.



Fig. 6. Footprint image

2.1.3 Conversion to Grayscale Format

The footprint image is acquired in RGB format. It is converted into grayscale intensity format before image enhancement. Fig. 7 shows the footprint image in grayscale format.

Footprint texture features are usually extracted using transform-based method such as Fourier Transform [8] and Discrete Cosine Transform [9]. While using Discrete Cosine Transform, some of the points are missed leading to an inaccurate inference. Also Fourier Transform involves floating-valued signals to integer-valued signals, thus less accuracy. Besides that, Wavelet Transform [10] is also used to extract the texture features of the footprint. In this work, a Sequential modified Haar Wavelet is proposed to find the Modified Haar Energy (MHE) feature [12].

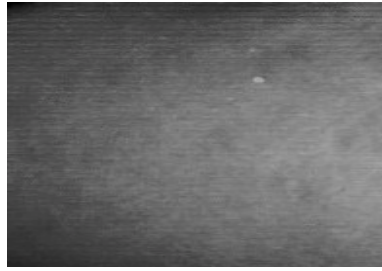


Fig. 7. Footprint image in Grayscale Intensity format

The Haar wavelet coefficients are represented using decimal numbers. It required eight bytes for storing each of the Haar coefficients. The division of the subtraction results in y-axis operation will also reduce the difference between two adjacent pixels. Due to these reasons, sequential Haar wavelet that maps an integer-valued pixel onto another Integer-valued pixel is suggested. Sequential Haar coefficient requires only two bytes to store each of the extracted coefficients. The cancellation of the division in subtraction results avoids the usage of decimal numbers while preserving the difference between two adjacent pixels. Fig.8. shows the decomposition using sequential Haar wavelet.

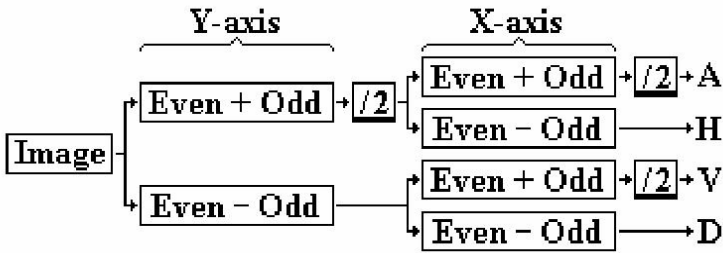


Fig. 8. Sequential Haar Transform

where X or Y-axis represents the direction of addition and subtraction. “/2” is the rounding of the “divide by 2” results towards the negative infinity. Fig 9 shows the location of horizontal details, vertical details and diagonal details for every decomposition levels.

For every image of the detail coefficients, the image is further divided into 4×4 blocks. The Modified Haar Energy for each of the block is calculated using (1).

$$MHE_{i,j,k} = \sum_{p=1}^P \sum_{q=1}^Q (C_{p,q})^2 \quad (1)$$

where i is the level of decomposition, j is Horizontal, Vertical or Diagonal details, k is the block number from 1 to 16, $P \times Q$ is the size of the block. The MHE energy feature for every detail coefficients are arranged as in (2).

$$MHE_{ij} = [MHE_{ij,1}, MHE_{ij,2}, \dots, HE_{ij,16}] \quad (2)$$

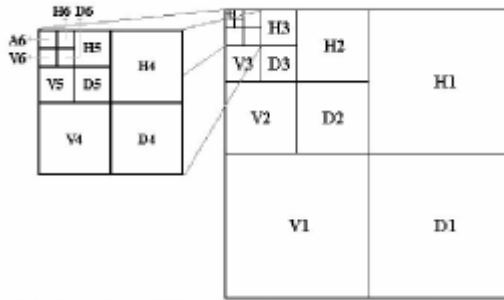


Fig. 9. Location of details in Every Decomposition Levels

2.1.4 Storing

Lastly, the Modified Haar Energy (MHE) or the threshold value of the footprint image is stored in the database for future references.

B. Recognition Method

2.2 Recognition Method

Recognition Method is also called as *Testing*. The normalization procedure removes the geometric information of input footprint (raw image). In Testing, normalization is done for the footprint image and then it is compared with the one stored in the database.

2.3 Experiments and Results

For experiment we took image samples of 30 different persons in 10 different angles using a pressure sensor. We used the block to get the conclusion.

2.3.1 Training

We took image samples of 30 different people and cropped heel portion to find the MHE of all blocks. We took a number of MHE out of which we select the minimum.

Let $MHE_1, MHE_2, MHE_3, \dots$ etc be the Modified Haar Energy for n blocks.

Using the formula,

$MHE = \text{Minimum}(MHE_1, \dots)$

For each person, the MHE is found in this way.

$V_{Person_1} = \text{Minimum}(MHE_1, \dots)$
$V_{Person_2} = \text{Minimum}(MHE_2, \dots)$
.....
$V_{Person_{30}} = \text{Minimum}(MHE_{30}, \dots)$

These are stored in the database.

2.3.2 Testing

The footprint to be tested was taken in 10 different angles and heel part was cropped so as to get the MHE of the heel shape using:

$$MHE = \text{minimum}(MHE_1, \dots)$$

This test MHE was compared with the MHE of different persons stored in the database.

$$\text{Result} = \text{Test}_{MHE} - \text{VPerson}_{iMHE}$$

Where $i=1$ to 30

If result is 0, the person with same footprint is found. That person is the master of the test footprint sample.

Comparison with other methods

The accuracy of Recognition of Footprint images using Sequential Modified Haar Transform is compared with the other methods for recognition. The following table shows the comparison of accuracy of recognition for Sequential Modified Haar Transform with Discrete Cosine Transform and Fourier Transform.

Table 1. Comparison of accuracy (performance) of the proposed method with other wavelet based method

Transform Types	Recognition Accuracy (%)	Computation time in ms (Recognition)
DCT [Ref.6]	81.64	152
FT [Ref.5]	84.43	138
SHT [Proposed]	88.37	128

The corresponding graph for the accuracy rate of recognition footprints using above mentioned techniques in as follows

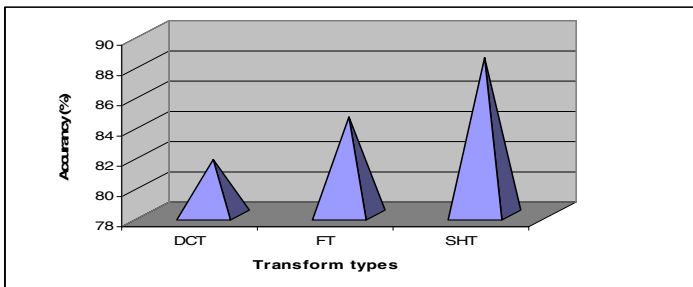


Fig. 10. Comparison of existing transforms (DCT, FT) Accuracy in % with proposed one (SMHET) The corresponding graph for the computation time (ms) of recognition footprints using above mentioned techniques in as follows

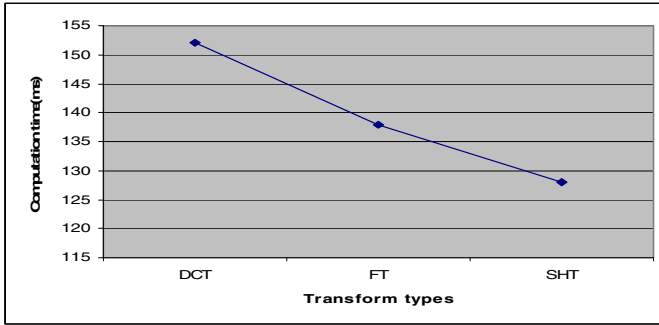


Fig. 11. Comparison of existing transforms (DCT, FT) Computation time (ms) with proposed one (SMHET)

3 Conclusion

The heel portion of the leg is cropped as it is having more intensity at this portion. This cropping is done using built-in function. The heel portion is divided into blocks using Sequential Modified Haar Transform. Minimum MHE is selected from all the calculated MHEs. Comparing the MHE of test image with all person's MHEs. If both MHEs match then the master of the footprint is found. This is an efficient method as it is more accurate. A high accuracy rate of 88.37% is achieved using Sequential Modified Haar Transform. Future enhancement of this project can be finding the age, weight and BP a person.

References

1. Special issue on automated biometrics. Proc. IEEE 85, 1348–1491 (1997)
2. Jain, L.C., et al. (eds.): Intelligent Biometric Techniques in Fingerprint and Face Recognition. CRC Press, Boca Raton (1999)
3. Masood, H., Mumtaz, M., Khan, S.A.: Wavelet based plamprint authentication system. In: IEEE Trans. Biometric and Security Tech., pp. 1–7 (2008)
4. Li, W., Zhang, D., Xu, Z.: Plamprint identification by recognition and artificial intelligence 16(4), 417–432
5. Plamondon, R., Lorette, G.: Automatic signature verification and writer identification—The state of the art. Pattern Recog. 22, 107–131 (1989)
6. Bimbot, F., Hutter, H.P., Jaboulet, C., Koolwaaij Lindberg, J., Pierrot, J.B.: Speaker Verification in the Telephone network: Research activities in the CAVE project. In: Proc. 5th Int. Conf. Speech Communication and Technology, Rhodes, Greece (September 1997)
7. Sukthar, R., Stockenton, R.: Argus the digital doorman. IEEE, Intelligent System 2 (2001)
8. Li, W., Zhang, D., Xu, Z.: Palmprint Identification By Fourier transform. Intl. Journal of Pattern Recognition and Artificial Intelligence 16(4), 417–432 (2002)
9. Jing, X.-Y., Zhang, D.: Face and Palmprint Recognition Approach Based on Discriminant DCT Feature Extraction. In: IEEE Trans.on Systems, Man, and Cybernetics –Part B: Cybernetics, December 2004, vol. 34(6), pp. 2405–2415 (2004)

10. Wu, X.-Q., Wang, K.-Q., Zhang, D.: Wavelet Based Palm print Recognition. In: Proceedings of the First Intl. Conference on Machine Learning and Cybernetics, November 2002, pp. 1253–1257 (2002)
11. Nakajima, K., Mizukami, Y., Tanaka, K., Tamura, T.: Foot-Based Personal Recognition. IEEE Tr. On Biomedical Engineering 47(11) (2000)
12. Ambeth Kumar, V.D., Ramakrishan, M.: Footprint Recognition using Modified Sequential Haar Energy Transform (MSHET). Intl. Journal of Computer science Issue 7(3/5) (2010)

An Efficient Approach for Data Replication in Distributed Database Systems

Arun Kumar Yadav¹, Ajay Agarwal², and S. Rahmatkar³

¹ Dept. of Computer Science, BMAS Engineering College, Agra,(U.P.), India
arun26977@rediffmail.com

² Dept. of MCA, Krishna Institute of Engg. & Tech., Ghaziabad (U.P.), India

³ Dept. of Computer Science, Nagpur Institute of Technology, Nagpur, India

Abstract. To increase the availability of data and fault tolerance in distributed database, it is better to add a backup server for each primary server in the system. So, the primary server and backup server must be connected to each other. To connect these computers to each other when they are at a long distance, it is necessary to use a leased line which needs to be charged as data is transferred. When packets are transferred between primary and backup server, more money need to be paid for charging this line. So if number of transferred packets between these computers can be reduced, the company can economize in its expenditures. Moreover, the number of transactions which should be performed in the backup server reduces.

In order to achieve this, a new method is introduced which will reduce the number of transferred packets between primary and backup server. In this method, the replicated data of primary server on other computers is used as its backup.

Keywords: Distributed Database, Remote backup, Data replication, Load balancing, filtering.

1 Introduction

One of the non-functional requirements in distributed databases is performance [1]. A technique to improve this requirement is data replication [2]. Data replication maintains multiple copies of data, called replicas, on separate servers. So the requests to these servers are answered locally [3]. Data Replication improves performance by the following: i) reducing latency, since users can access replicated data, so it avoids remote network access; and ii) increasing throughput, since data are available on multiple computers and can be accessed simultaneously. When data is replicated on more than one computer, it is necessary to ensure its consistence. There are some protocols such as single lock, distributed lock, primary copy, majority protocol, biased protocol, and quorum consensus protocol [4] which are responsible to ensure the data consistence.

Another non-functional requirement in distributed system is fault tolerance. So it ensures that the system can work accurately even in case of occurrence of faults [5, 6]. We can achieve this ability by performing transaction processing at one server, called the primary server, and having a remote backup server where all data of the

primary server are replicated. When a primary server fails, its remote backup server is responsible to answer the requests until the primary server come back to stable state [4]. Figure 1 shows the architecture of a remote backup system.

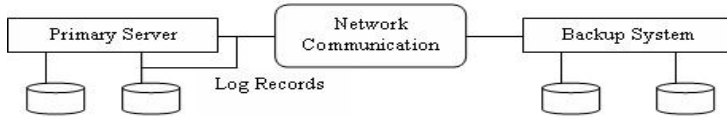


Fig. 1. Remote Backup System

In this paper we are introducing a technique which uses backup server data to recover the primary server from failure. The rest of the sections are organized as follows: Section 2 expresses related work. Section 3 describes how to use replicated data in backup process. Section 4 includes implementation remarks. In section 5, we have evaluated our technique and At last in section 6, some constrains in our technique and conclusion of the paper is presented.

2 Related Works

One of the advantages with distributed databases is high availability of data; that is, the database must function almost all the time. In particular, since failures are more likely in large distributed systems, a distributed database must continue functioning even when there are various types of failures. The ability to continue functioning even during failures is referred to as robustness. Whenever a transaction writes an object it also updates the version number in this way:

- If data object A is replicated on n different sites, then a lock-request message must be sent to more than one-half of the n sites in which a is stored. The transaction does not operate on A until it has successfully obtained a lock on a majority of the replicas of a.
- Read operations look at all replicas on which a lock has been obtained, and read the value from the replica that has the highest version number.

Writes read all the replicas just like reads to find the highest version number. The new version number is one more than the highest version number. The write operation writes all the replicas on which it has obtained locks, and sets the version number at all the replicas to the new version number. The read one, write all available scheme can be used if there is never any network partitioning, but it can result in inconsistencies in the event of network partitions [4].

The Database Mirroring method is one of the best methods for backup data. In the simplest operation of database mirroring, there are two major server-side components, the *principal server instance* (Principal) and the *mirror server instance* (Mirror) [7].

The basic idea behind database mirroring is that synchronized versions of the database are maintained on the principal and mirror. If the principal database becomes unavailable, then the client application will smoothly switch over to the mirror database, and operation will continue as normal. The mirroring architecture of database is shown in figure 2 [7].

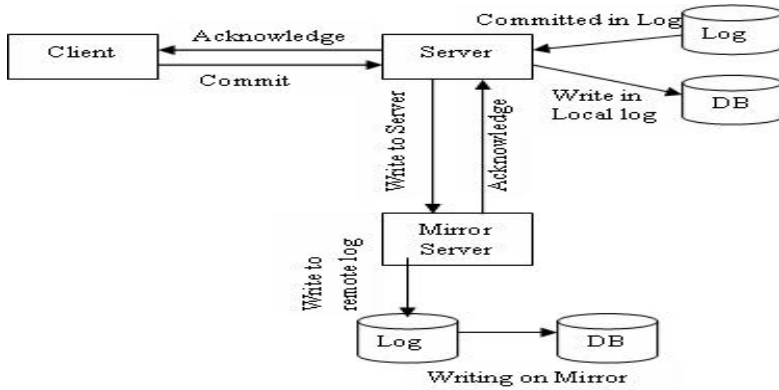


Fig. 2. Mirroring Architecture of Database

3 Data Replication to Reduce Backup Cost

To apply our method in the distributed database, some operations in the primary server, backup server, and guest computer must be changed or added. In the following sections we describe these necessary changes.

3.1 Changes Required in the Primary Server

As we mentioned, the primary server transfer the transactions on data which is not replicated in the guest computers. So the primary server and backup server should know which data item is replicated on which guest computer. Hence when the primary server replicates a data item to a guest computer, assigns a name to that data and send this name along data item to the guest computer. After that, the primary server sends a packet to backup server which includes name of replicated data and the guest on which is replicated.

If a guest computer requests a data item which is named already, the primary server does not name it again. The primary server sends data item with its name to the guest. It also resends data item name and new guest computer which requests that data to the backup server.

3.2 Changes Required in the Backup Server

As we mentioned, some data on backup server in not updated. Instead, it receives from the primary server name of this data and the guest computer which has this data. So when the primary server fails, backup server has to get these data from appropriate guest and updates its old data by new data. If a data item is replicated to more than one guest, backup server should receive that data item form a guest which has the last version of that data.

3.3 Changes Required in the Guest Computers

When a guest computer requests a data form the primary server to replicate, it receives name of that data along with data. The guest computer should save that name along data. So in the failure time, if the backup server requests a data item by its name, the guest computer know which data corresponds with that name, and transfers the requested data to the backup server.

4 Implementation Remarks

As we have mentioned, the primary sever should name data item which is being replicated on the guest computers. To name data items, primary server can use a counter; a data item is named according to the counter. After naming a data item, the counter is increased by one. To assign this name to data items, for example in a relational database, it is better to add a column named *data_name* to each table in database. So name of each record is inserted in this column. All records in a data item have the same name and the value of *data_name* attribute of these records is equal to each other.

There is a database named *Student* in the primary server with some records in it is shown in Table 1.

Table 1. Student Table

Row No.	Name of Student	Stud_Number	Data_Name
1	Nikhil	1001	
2	Bharat	1002	
3	Jayesh	1003	
4	Ankur	1004	
5	Nititn	1005	

To clarify the issue, suppose a primary server, its backup server, and two guest computers are connected to each other. Suppose guest 1 requests records number 1, 3, and 5 to replicate from the primary server. The primary server images these records as a data item and names it *Q1*. So the value of *data_name* attribute for these records is equal to *Q1* and this name is sent to the guest 1 and backup server. This process is shown in figure 3.a, suppose guest 2 request records number 3, 2, 4. Records number 3 is member of data item *Q1*, so the primary server send all records in data item *Q1* to the guest 2. After that it assign a name *Q2* to records number 2 and 4, and send them to the guest 2. This process is shown in figure 3.b.

So if a guest requests from the primary server some records which are subset of a data item *Q1*, all records in that data item are sent to the guest.

Now suppose the primary server fails. The backup server knows that data item *Q2* is replicated on guest 2, and requests this data item from guest 2. Data item *Q1* is replicated in both guest 1 and 2. The primary requests this data item from the guest which has the last version of that data item. So the backup server requests this data item from guest 2 and updates its data. This process is shown in figure 4.

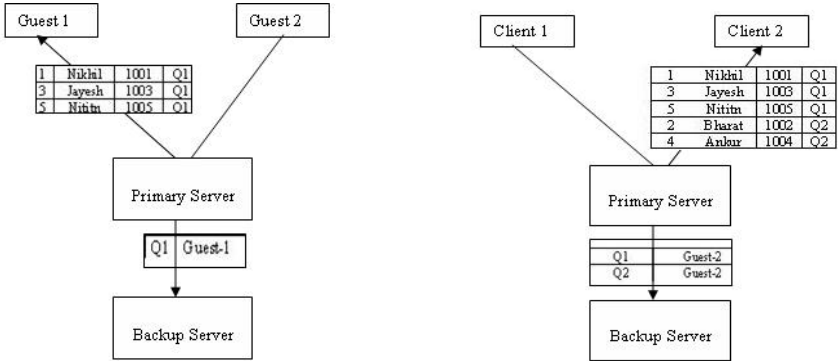


Fig. 3. Replicating Data to Client 1 and Client 2 from Primary Server

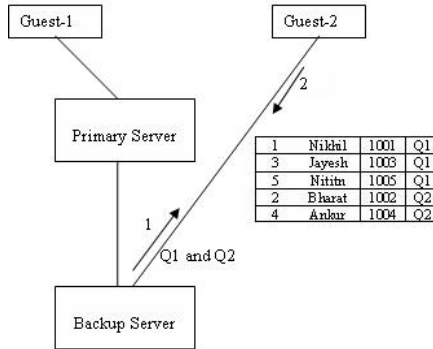


Fig. 4. Transferring Data to the Backup Server

5 Evaluations

In this section, we are evaluating our technique. As mentioned in the previous section, when the primary server data is replicated on some guest computers, this replicated data can be used to recover system from a failure state.

Firstly we are measuring the network traffic between primary and backup server in mirroring method. In this method primary server sends the transactions on data to the backup server in specific interval time. After that we will apply our method to the system and will measure the network traffic which is imposed to the system. Finally we can compare the network traffic in our method and mirroring method.

To evaluate our method we are using SQL Server 2005 database system. Our primary server has 200000 records. We replicate 500 records to a guest computer and perform 50 update transactions to these records. The network traffic to update backup server in the mirroring method is equal to 250 Kbyte. If we apply our method to the system, this value is equal to 55 Kbyte. We continue to increase replicated data and perform the same 50 update transaction to them and monitor network traffic in our

method and mirroring method. The results are shown in table 2 and figure 5 shows this result as a diagram. We find that the amount of replicated data effects in performance in our method, and improves it.

Table 2. Network Traffic and replicated data comparison in mirroring and our method

Network Traffic Between Primary Server and Backup Server (Mbyte)		Number of Records Replicated
Proposed Method	Mirroring Method	
0.055	0.25	500
0.102	0.25	1000
0.508	3.125	5000
1.016	5.75	10000
3.555	16.812	30000
5.078	30.687	50000
10.148	55.312	100000

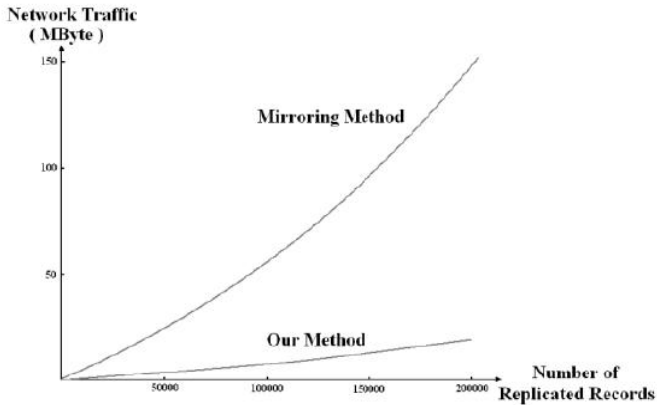


Fig. 5. Diagram of relationship between network traffic and replicated data in the mirroring and our method

6 Method Constrains and Conclusion

There are some constraints in our method. First, if the primary server and one of the guest computers which has data which is replicated only on that guest computer get fail concurrently, the backup computer cannot recover all data, and as a result data may be inconsistent. As you know this condition takes place very seldom. Second, our method imposes some over load in the primary server. So the primary server should have a suitable hardware to continue its work. Third, the recovering system from failure needs more time than previous methods. Because the backup computer should gets updated information from the guest computers, if the number of guest computers is high, this process may needs very much time.

This paper presents a new method to decrease network traffic between primary server and backup server. In this method replicated data is used for backup process. Transactions on data which is replicated on the guest computers are not transfer from primary server to backup server. Instead, at the time of primary server fails, this replicated data are transferred from the guest computers to the backup server.

Reduction of network traffic between primary server and backup server causes fewer transactions execute in the backup server. In addition, the company has to pay less money for charging the link.

References

1. Sleit, A., AlMobaideen, W., Al-Areqi, S., Yahya, A.: A Dynamic Object Fragmentation and Replication Algorithm in Distributed Database Systems. *American Journal of Applied Sciences* 4(8), 613–618 (2008)
2. Loukopoulo, T., Ahmad, I.: Static and adaptive distributed data replication using genetic algorithms. *Journal of Parallel and Distributed Computing* 64(11), 1270–1285 (2004)
3. Abdul-Wahid, S., Andonie, R., Lemley, J., Schwing, J., Widger, J.: Adaptive Distributed Database Replication Through Colonies of Pogo Ants. In: *IEEE International Parallel and Distributed Processing Symposium, IPDPS 2007*, pp. 358–365 (2008)
4. Silberschats, A., Korth, H.F., Sudarshan, S.: *Database System Concepts*. McGraw-Hill, New York (2006)
5. Gashi, I., Popov, P., Strigini, L.: Fault tolerance via diversity for off-the-shelf products: A study with SQL database servers. *IEEE Transactions on Dependable and Secure Computing* 4(4), 280–294 (2007)
6. Wang, Mueller, F., Engelmann, C., Scott, S.: A job pause service under lam for transparent fault tolerance. In: *IEEE International Parallel and Distributed Processing Symposium*, pp. 1–10 (2007)
7. Rizzo, T., Machanic, A., Skinner, J., Davidson, L., Dewson, R., Narkiewicz, J., Sack, J., Walters, R.: *Pro SQL Server 2005*, ch. 15. A Press (2006)

Unicast Quality of Service Routing in Mobile Ad Hoc Networks Based on Neuro-fuzzy Agents

V.R. Budyal¹, S.S. Manvi², and S.G. Hiremath³

¹ Basaveshwar Engineering College, Bagalkot, India
vrbudyal@yahoo.co.in

² Reva Institute of Technology and Management, Bangalore, India
agentsun2002@yahoo.com

³ G. M. Institute of Technology, Davanagere, India
Department of Electronics and Communication Engineering

Abstract. This position paper presents Quality of Service (QoS) routing model in Mobile Ad hoc Networks (MANETs) by using software agents that employ fuzzy logic and neural networks for intelligent routing. It uses Dynamic Source Routing (DSR) in MANETs to find various paths and attributes. Fuzzy static agents decide whether each node on the path satisfies QoS requirement for multimedia application. The static neuro-fuzzy agents are used for training and learning to optimize the input and output fuzzy membership functions according to user requirement, and Q-learning (reinforcement learning) static agent is employed for fuzzy inference instead of experts experience. Mobile agents are used to maintain and repair the paths.

Keywords: mobile ad hoc networks, routing, neuro-fuzzy, software agents.

1 Introduction

A Mobile Ad hoc network (MANET) is a self configuring network consisting of mobile devices connected by wireless links. Each device in a MANET is free to move independently in any direction. Each node acts as host and a router. The QoS routing protocol is also needed in a stand-alone multi-hop mobile network for real-time applications (like voice, video, etc.). QoS routing requires not only finding a route from a source to destination, but a route that satisfies the end-to-end QoS requirement, often given in terms of bandwidth or delay [1].

Recently software agents are considered to be one of the important paradigms for providing flexible and adaptable services in intelligent communication networks. Agents are autonomous programs activated on an agent platform of host. The agents use their own knowledge base to achieve the specified goals without disturbing the activities of the host [2]-[4].

The problem addressed in the paper is to route the packets between source and destination by path discovered using DSR (based on DSR [5]) to satisfy QoS parameters. Our contributions in the paper are as follows: building the neuro-fuzzy agents for optimizing the input and output fuzzy membership functions using feed forward back

propagation learning, designing fuzzy inference based on Q-learning agent, finding the QoS satisfied path from source to destination with multi-metric uncertain input parameters using fuzzy agents and maintaining the QoS path whenever there is failure either in link or node using mobile agents.

2 Proposed Work

An agent based model comprising of User Agency, DSR Agency and QoS Agency for QoS routing in MANETs at each node is shown in Fig.1. The knowledge base facilitates inter-agent communication.

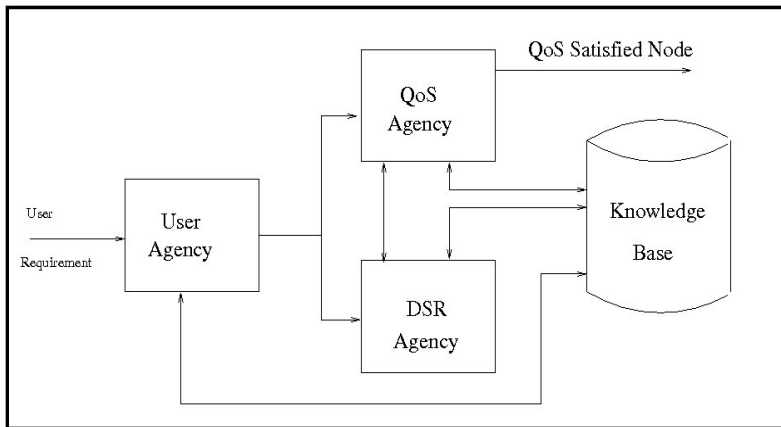


Fig. 1. Components of Agency

2.1 User Agency

User Agency consists of User Manager Agent and User Agent. *User Manager Agent* is a static agent, which is directly accessible to an application, and creates a *User Agent* whenever there is request from the user for multimedia data transmission with specific QoS requirement and it also, triggers QoS Agency and DSR Agency. User Agent holds the requirement of the user for particular application.

2.2 DSR Agency

DSR Agency comprises of static DSR Manager Agent, DSR Agent, Node Dis-joint Multipath Routing and Stability agent and Maintenance mobile Agent. *DSR Manager Agent* is a static agent. It accepts destination node address from User Manager Agent and creates and co-ordinates with DSR Agent, NDMR and Stability Agent and Maintenance Agent, to obtain stable NDMR path. DSR Agent is static agent. It sends Route REQuest (RREQ) to its neighbours till destination is found. Route REPLY (RREP) carries information of available bandwidth and packet loss rate at each intermediate node which is saved in DSR Manager Agent.

Node Dis-joint Multipath Routing and Stability agent is a static agent that performs the following tasks at the node it is residing. It gets all multipaths from DSR Manager Agent. NDMR and Stability Agent compute to find node dis-joint paths. The other task performed by NDMR and Stability Agent is to compute stable path from number of NDMR paths. DSR Manager Agent decides higher stability path and sends message to QoS Agency. *Maintenance Agent* is a mobile agent that migrates from node to node periodically on the QoS satisfied route. In case of either mobility or violation of QoS satisfaction of intermediate node, application may use either new path from DSR Manager Agent, which satisfies QoS, or local recovery patch up paths are used to, further continue the communication.

2.3 QoS Agency

QoS Agency comprises of QoS Manager Agent, Neuro-fuzzy Agent and Fuzzy QoS agent. *QoS Manager Agent* is a static agent creates Neuro-fuzzy Agent, Fuzzy QoS agent and Q-learning Agent. Neuro-fuzzy Agent updates optimized fuzzy membership functions in QoS Manager Agent. Q-learning Agent infers the rules of the fuzzy if-then rules. Fuzzy QoS Agent maps available bandwidth and packet loss on to fuzzy logic and obtains the crisp value. *Neuro-fuzzy Agent* is a static Agent that performs the following tasks at the node it is residing. It generates the training data set and learns by feed forward backward propagation, which consists of five layers. Neuro-fuzzy Agent computes output at each layer till the last layer which is compared with desired value and obtains error. To distribute the error in the hidden layer, Neuro-fuzzy Agent back propagates the error at each layer and updates the weights.

The same procedure is repeated till error is within the acceptable limits. QoS Manager Agent holds the optimized input and output fuzzy membership functions. *Q-Learning Agent* is a static agent that performs the tasks like evaluating the consequences for the rule, computing the value of Q with number of iterations until the value Q stabilizes. Consequence computed by Q-learning Agent is updated in Fuzzy QoS Agent. *Fuzzy QoS Agent* is static agent which performs the following tasks: (1) It fuzzifies the input variables on optimized membership functions. (2) Applies if-then rules to obtain fuzzy inference for each rule where fuzzy inference is obtained from Q-learning Agent. (3) Fuzzy QoS agent computes defuzzified value using centroid method to decide QoS satisfaction node.

An algorithm to depict the functioning of the model is as follows.

Algorithm. Functioning of QoS path discovery

1. : Begin
2. : Accept application route request and QoS requirement;
3. : Compute Stable NDMR paths from DSR multipaths;
4. : Optimize input and output fuzzy membership functions based on QoS Requirement using neuro-fuzzy agents;
5. : Learn fuzzy inference by Q-learning agent;
6. : Check QoS satisfaction using fuzzy logic at each node for the complete stable NDMR path selected;
7. : IF {node satisfies QoS}
8. : THEN {check QoS for next intermediate node};

9. : ELSE {accept next higher stable NDMR path, go to step 6};
10. : Start sending the multimedia data packets on QoS satisfied path;
11. : End.

3 Conclusion

Proposed model optimizes input and output fuzzy membership functions and inference based on neuro-fuzzy agents and Q-learning agents respectively. For QoS multimedia applications requirement, proposed scheme takes appropriate fuzzy decisions and maintains the path when there is a path failure using mobile agents. Using this approach, we can guarantee the QoS routing for MANETs. We are planning to simulate the model using ns-2 and evaluate the performance of model in terms of application rejection ratio, agents overhead and route discovery time.

References

1. Kannan, T., Pushpavalli, M., Natarajan, A.M.: Fortification of QoS routing in MANETs using Proactive Protocols. In: IEEE International Conference on Wireless Communication and Sensor Computing (ICWCSC 2010), India (2010)
2. Budyal, V., Manvi, S.S., Hiremath, S.G., Kala, K.M.: Intelligent Agent based QoS Enabled Dis-joint Multipath Routing in Manets. In: 17th IEEE International Conference on Advanced Computing and Communication, India (2009)
3. Wang, Z., Seitz, J.: An Agent based Service Discovery Architecture for mobile environments. In: 1st EuroAsian conference on Information and Communication Technology, UK, pp. 350–357. Springer, Heidelberg (2002)
4. Manvi, S.S., Vekataram, P.: International Journal on Computer Communications 27(15), 1493–1508 (2004)
5. Zhong, Y., Yuan, D.: Dynamic Source routing Protocol for Wireless Ad hoc Networks in special scenario using Location Information. In: International Conference on Communication Technology ICCT, vol. 2, pp. 1287–1290 (2003)

Performance Comparison of Routing Protocols in Wireless Sensor Networks

Geetika Ganda¹, Prachi², and Shaily Mittal³

¹ M-tech Student, ² Assistant Professor, ³ Assistant Professor
Dept. of Computer Science, Dept. of Information Technology
IITM University, Gurgaon

geetikaganda@gmail.com, prachi@itmindia.edu, shally@itmindia.edu

Abstract. This paper aims to compare performance of some routing protocols for Wireless Sensor Networks (WSNs). A Wireless Sensor Networks (WSN) is a set of hundreds or thousands of micro sensor nodes that have capabilities of sensing, establishing wireless communication between each other and doing computational and processing operations. The efficiency of sensor networks strongly depends on the routing protocol used. Routing protocols are needed to send data between sensor nodes and the base station. In this paper, we analyzed three different types of routing protocols: Fisheye, LANMAR, LAR1. Sensor networks were simulated using Qualnet simulator. Several simulations were conducted to analyze the performance of these protocols on the basis of performance metrics such as hop count, throughput, end-to-end delay.

Keywords: WSNs, LANMAR, Fisheye, LAR1.

1 Introduction

Routing is a function in the network layer which determines the path from a source to destination for the traffic flow. WSNs routing protocols are broadly divided into two categories [2] i.e reactive(on-demand) and proactive(table-driven). In Table-driven routing protocols, each node maintains one or more tables containing routing information to every other node in the network. All the nodes update these tables so that a consistent and up-to-date network is maintained. In contrast to table-driven routing protocols, all up-to-date routes are not maintained at every node; instead the routes are created as and when required. Various reactive protocols are DSR, AODV, ABR, TORA etc. LSR, OLSR, DSDV, LAR1, Fisheye, LANMAR, are proactive protocols. In this paper we have used Fisheye, LANMAR and LAR1 i.e. unicast routing protocols for their performance comparison. We have taken these protocols as these are not evaluated earlier for such comparison.

An earlier protocol performance comparison was carried out by Guangyu Pei *et al* in [10], who conducted experiments with Ad hoc On-Demand Vector routing (AODV), Fisheye, Dynamic MANET On-demand (DYMO), Source Tree Adaptive Routing (STAR) protocol, Routing Information Protocol (RIP), Bellman Ford, LandMark Ad hoc Routing protocol (LANMAR) and Location Aided Routing protocol (LAR). This simulation experiment showed that AODV, Dymo and Bellman ford protocols are having higher end to end delays than others, indicating that the speed of simulation in

large scale networks will be affected, whereas LANMAR and RIP shows the considerable amount of delay in scaled up environment.

Performance comparison of AODV, DSR, FSR and LANMAR is presented by M. Gerla *et al* in [11]. According to their simulation results LANMAR outperforms FSR under all delay and throughput measures. In the last few years, there are several researches have evaluated the performance of routing protocols for mobile Ad-Hoc network as a function of mobility rate and pause time using ns2(network simulator 2)[9]. There are lesser evaluations available using Qualnet simulator [1] which is commercially available and faster than ns2 [3]. We are using Qualnet simulator for comparison evaluation of *LANMAR, LAR1 and Fisheye*.

DSR and DSDV were simulated and compared to a newly developed Cluster-based Routing Protocol (CBRP) by Mingliang, Tay and Long [12]. The simulations were performed with pause times from 0 to 600 seconds and with 25 to 150 mobile nodes. Their results shows CBRP performed much better with a delivery ratio always greater than 90 percent and a lower routing overhead than DSR in larger networks.

An earlier protocol performance comparison was carried out by authors in [12], who conducted experiments with Destination Sequence Distance Vector (DSDV), Temporally-ordered routing algorithm (TORA) along with DSR and AODV. The simulations were quite different for they used a constant network size of 50 nodes, 10 to 30 traffic sources, seven different pause times and various movement patterns on ns2 simulator. The rest of the paper is organized as follows: Section 2 describes three concerned protocols in detail i.e. Fisheye, LANMAR and LAR1. Section 3 describes the simulation environment, parameters evaluated and simulation results. Lastly work is concluded in section 4.

2 Preliminaries

2.1 Fisheye

Fisheye technique proposed by Kleinrock and Stevens [4] to reduce the size of information required to represent graphical data. The eye of a fish captures with high detail the pixels near the focal point. The detail decreases as the distance from the focal point increases. Fisheye State Routing (FSR) [4] generates accurate routing decisions by taking advantage of the global network information. Fisheye Routing determines routing decisions using a table-driven routing mechanism similar to link state.

2.2 LANMAR

The Landmark Ad-hoc Routing Protocol (LANMAR) [5] combines the features of FSR and landmark routing. LANMAR assumes that the large scale ad hoc network is grouped into logical subnets in which the members have a commonality of interests and are likely to move as a "group". LANMAR uses the notion of landmarks to keep track of such logical subnets[6]. The route to a landmark is propagated throughout the network using a Distance Vector mechanism [11]. The routing update exchange of LANMAR routing can be explained as follows. Each node periodically exchanges topology information with its immediate neighbors. In each update, the node sends entries within its Fisheye scope [4]. Updates from each source are sequentially numbered. To the update, the source also piggybacks a distance vector of all landmarks.

As a result, each node has detailed topology information about nodes within its Fish-eye scope and has a distance and routing vector to all landmarks.

2.3 LAR1

The goal of Location-Aided Routing (LAR)[7] is to reduce the routing overhead by the use of location information. LAR protocol uses the GPS (Global Positioning System) to get location information of mobile hosts. In the LAR routing technique,[8] route request and route reply packets similar to DSR and AODV are being proposed.

3 Performance Evaluation

We carried out simulations on Qualnet simulator. We designed the network using Random waypoint model with different number of nodes. We compiled the results using 5 simulations and the application traffic between the randomly chosen source and destination is CBR traffic[7]. The metrics used to measure the performance of protocols are average end to end delay, average TTL based hop count and throughput.

3.1 Simulation Results

3.1.1 Average End to End Delay

End-to-end delay indicates duration for a packet to travel from the CBR source to the application layer of the destination. According to results obtained in figure 1 LANMAR shows minimum end to end delay of 0.015 s and almost remains constant irrespective of increase in no. of nodes. Similar is the case with LAR1 protocol which shows a slight higher delay in comparison with LANMAR. FSR shows worst performance with highest end to end delay of 0.023s.

3.1.2 TTL Based Average Hop Count

Hop count is the number of hops a packet took to reach its destination. The results for TTL based hop count in figure 1 shows three protocols have a constant hop count irrespective of the no. of nodes. With the increase of no. of nodes it remains constant. LANMAR have highest hop count of 64 hops while FSR and LAR1 require less number of hops of 19 hops and 1 hop respectively. The plotted graph shows that hop count is independent of no. of nodes. Hence, according to the results down, it is clear that these three protocols are independent of scalability.

3.1.3 Throughput

Throughput is the average rate of successful message delivery over a communication channel. This data may be delivered over a physical or logical link, or pass through a certain network node. It is usually measured in bits per second. The results of throughput in figure 1 shows that LAR1 performs the best with the highest throughput in spite of increased no. of nodes while FSR drops down to approx. near zero.

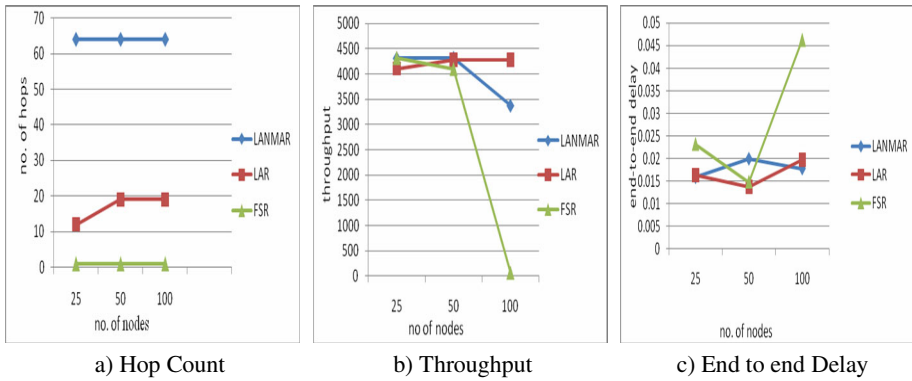


Fig. 1. Simulation results of FSR, LANMAR AND LAR1

4 Conclusion

In this paper, a performance comparison of three different routing protocols i.e. FSR, LANMAR, and LAR1 for wireless sensor network is presented. Three performance metrics used to compare protocols are average end to end delay; average TTL based hop count and throughput. LANMAR performs best in measuring end to end delay and FSR performs best in TTL based hop count. LAR1 performs best in case of throughput. In future, this work may be extended for analyzing the behavior of these protocols in heterogeneous networks with many more metrics for evaluation.

References

1. The Qualnet simulator, <http://www.Scalable-Networks.com>
2. Acs, G., Buttyabv, L.: A taxonomy of routing protocols for wireless sensor networks. BUTE Telecommunication Department (January 2007)
3. Jorg, D.O.: Performance comparison of MANET routing Protocols in different network sizes. Institute of computer Science and applied mathematics computer Networks and distributed systems University of Berne, switzerland
4. Sun, A.C.: Design and Implementation of Fisheye Routing protocol for Ad Hoc Networks. Dept of Electrical and CSE, Massachusetts institute of Technology (May 2000)
5. Lee, Y.Z., Chen, J., Hong, X., Xu, K., Breyer, T., Gerla, M.: Experimental Evaluation of LANMAR, a scalable Ad- Hoc routing protocol. In: IEEE Communications Society/WCNC 2005. University of California, Los Angles (2005)
6. Pei, G., Gerla, M., Hong, X.: LANMAR: Landmark Routing for Large Scale Wireless Ad-Hoc Networks with Group Mobility. In: Proceedings of IEEE/ACM MobiHOC 2000, Boston, MA, August 2000, pp. 11–18 (2000)
7. Arshad, J., Azad, M.A.: Performance Evaluation of Secure on-Demand Routing Protocols for Mobile Ad-hoc Networks. In: 15th IEEE International Conference on Network Protocols, Beijing, China (2007)
8. Kurkowski, S., Navidi, W., Camp, T.: Discovering Variables that Affect MANET Protocol Performance. In: Proceedings of IEEE Global Telecommunications Conference, GLOBE-COM 2007, Air Force Inst. of Technol., Wright Patterson, November 26-30. IEEE, Los Alamitos (2007)

9. Boomarani Malany, A., Sarma Dhulipala, V.R., Chandrasekaran, M.: Throughput and Delay Comparison of MANET Routing Protocols. *Int. J. Open Problems Compt. Math.* 2(3) (September 2009), ISSN 1998-6262
10. Pei, G., Gerla, M., Hong, X.: Landmark Routing for Large Scale Wireless Ad Hoc Networks with Group Mobility. CSE Department University of California
11. Gerla, M., Hong, X., Pei, G.: Landmark Routing for Large Ad Hoc Wireless Networks. In: *Proceeding of IEEE GLOBECOM 2000*, San Francisco, CA (November 2000)
12. Mingliang, J., Tay, Y., Long, P.: A cluster-based routing protocol for mobile ad hoc networks (1999/2002), <http://www.comp.nus.edu.sg/~tayyc/cbrp/hon>
13. Broch, J., Maltz, D.A., Johnson, D.B., Hu, Y.-C., Jetcheva, J.: A performance comparison of multi-hop wireless ad hoc network routing protocols. In: *Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM 1998)*, October 1998, pp. 85–97 (1998)

Security and Trust Management in MANET

Akash Singh, Manish Maheshwari, Nikhil, and Neeraj Kumar

School of Computer Science Engineering
Shri Mata Vaishno Devi University (J&K), India
akash72@gmail.com, manish5@live.in, nnikhil.roy@gmail.com,
nehra04@yahoo.co.in

Abstract. A mobile adhoc network (MANET) is a network which does not have any centralized control. Security and trust management are paramount concern for these networks for efficient data transfer among the participating nodes. In this paper, we propose an efficient security and trust management based algorithm for MANET. The proposed algorithm consists of three steps: initialization, data transmission, and detection. The time based nonce is generated at different time interval which gives effectiveness to the proposed approach in the sense that it is not easy to detect the generated nonce. The proposed approach is quite effective with the earlier approaches to detect the security threat in MANET.

Keywords: MANET, Security, Trust management.

1 Introduction

A mobile ad hoc network (MANET) is a collection of wireless mobile hosts forming a temporary network without the aid of any centralized administration or standard support services regularly available in wide-area networks to which the hosts may normally be connected [1]. MANET are vulnerable to a powerful attack known as wormhole attack. In a wormhole attack [2], an attacker introduces two transceivers into a wireless network and connects them with a high quality, low-latency link. Routing messages received by one wormhole endpoint are retransmitted at the other endpoint. Attackers can exploit wormholes to build bogus route information, selectively drop packets, and create routing loops to waste the energy of network.

Several cryptographic approaches have been proposed in literature to mitigate security attacks but many of these proposals may not be suitable for real-time applications where time is a real challenge/constraint since encryption and decryption of message requires time which introduces the factor of latency. Moreover, the devices used in encryption and decryption have limited resources in terms of resources such as computation power and memory. Keeping in the mind of above challenges and constraint, we propose a latency aware real time cryptographic communication algorithm for MANET. Rest of the paper is organized as follows: Section 2 discusses the related work, Section 3 describes the proposed approach (algorithm), and Section 4 provides conclusions.

2 Related Work

Many solutions for security in MANET have been proposed over years [3]. These solutions try to minimize the overhead and maximize the security of the system. Generalize architectures for intrusion detection is given in [4] in which all nodes participate in the monitoring of the data transmission. To prevent resource consumption attacks, LHAP [5] implements lightweight hop-by-hop authentication. Using LHAP, a node joining MANET only need to perform some inexpensive authentication operations to bootstrap a trust relationship with its neighbors. It then switches to a very lightweight protocol for subsequent traffic authentications.

Many Group Key Agreement protocols [6] have been proposed in the literature, most being derived from the two-party Diffie–Hellman (DH) [7] key agreement protocol. Some have no formal proofs while some are secure against passive adversaries only. Boyd et al. [8] proposed an efficient constant-round protocol where the bulk of the computation is done by one participant (the current group leader), thus making it highly efficient for heterogeneous ad hoc networks. It is provably secure in the Random oracle model [9], but lacks forward secrecy.

3 Proposed Approach

The proposed scheme consists of following three steps: Initialization, data transmission, and detection. We assume that there are no malicious nodes in the system. After time synchronization, each node sends packet to its neighbour to know its identity. This identity is multiplied with identity of node modulus P and stored in table. Here P is a large prime number. Let n is node id and n_1 is id of neighbour then $table[] = (n * n_1) \bmod P$. Nonce generation algorithm is started in each node. In this algorithm according to clock pulse, i is generated. The value of i and i^{-1} is stored in a hash table along with the time stamp. This hash table will store the value only for a particular window of time say n seconds. On every node, for a particular time value, same nonce value is generated. Nonce corresponding to the latest time stamp present in hash table is fetched say i . Let a_1 be the identity of given node. Then $b_1 \equiv a_1^i \bmod P$ where P is large prime number. The value of b_1 and the time stamp at which i is calculated is appended with the data encrypted in DES and transmitted as $D = b_1.time_stamp.E(K, data)$.

When this packet is received by the neighbor, the neighbor extracts the value of time stamp and b_1 from the packet. The value of i^{-1} corresponding to the extracted time stamp is fetched from the hash table. If the time stamp is not present in hash table, it means that the duration of time stamp has expired and the packet is dropped. If time stamp is present, packet is valid. The calculation $(a_2^{i^{-1}} * b_1) \bmod P$ is performed. This is equivalent to $(a_2^{i^{-1}} * a_1^i) \bmod P = a_2 * a_1 \bmod P$ as i and i^{-1} are inverse of each other and a_2 is the node id. This value is checked with value present in $table[]$. If value is found, the data is from authorized node. The proposed algorithm is described in the following section:

clock () : this is the clock of the node which keeps the time of the node and maintains the node in synchronization.

EXTENDED EUCLID (P, i) : This is extended Euclidean algorithm used to find inverse of i modulo P.

Step 1: Initialization

time = tsync, clock(time), P = large prime number, i = 0, neigh = 0

repeat

n = neighbour id;

table[i] = node_id * n mod P; i ++

until(neighbour ≠ ∅)

repeat

i = nonce; time_stamp = clock();

inv_i = EXTENDED EUCLID(P, i);

hash_table(time_stamp, inv_i, i);

forever

Step 2: Data Transmission

a1 = node_id;

repeat

t = latest time stamp present in hash

ii = hash_table.t.i;

b1 = a1ⁱⁱ mod P;

enc_data = b1.time_stamp.E(K, data);

forever

Step 3: Detection

a2 = node_id;

repeat

bb1 = enc_data -> b1; time_st = enc_data -> time_stamp

if(hash_table.time_st ∈ hash_table)

check = (a2^{inv_i} * b1) mod P

if(check ∈ table[])

data transmission routine is called and data is forwarded

else

data is discarded

else

data is discarded

forever

4 Comparisons and Discussion

Table 1. Relative comparison of existent schemes

Approach	Time Sync	Random Nonce	Centralized Authority	Key Compromise	GPS
Packet Leashes [10]	Yes	No	No	N.A.	Maybe
Directional Antenna[11]	No	No	No	N.A	Maybe
Group Key [6]	Yes	Yes	Yes	Possible	No
Proposed	Yes	Yes	No	Very Difficult	No

5 Conclusions

The proposed algorithm is nonce based and is resilient to various known attacks in MANET such as man in the middle attack, and passive eavesdropping, active interference. By applying nonce concept in proposed algorithm it enhances its capability for addressing real time application. Source consumption get reduced by calculating the identity of each neighbor and is stored in the hash table for a particular period of time. Authentication process get faster as only time stamping factor has to be matched with the hash table and the encrypted data is just forwarded to destination node which increases the packet delivery fraction in MANET.

References

1. Raimundo, J., Macêdo, A., Assis Silva, F.M.: The mobile groups approach for the coordination of mobile agents. *J.Parallel Distributed Computing* 65, 275–288 (2005)
2. Eriksson, J., Krishnamurthy, S., Faloutsos, M.: Truelink: A practical countermeasure to the wormhole attack. In: *ICNP* (2006)
3. Buttyan, L., Hubaux, J.-P.: Stimulating cooperation in self-organising mobile ad hoc networks. In: *Mobile Networks and Applications* (2003)
4. Kumar, N., Patel, R.B.: MASLKE: mobile agent based secure location aware key establishment in sensor networks. In: *Proc. ICON 2008*, New Delhi, India (2008)
5. Zhang, Y., Lee, W.: Instruction Detection in wireless adhoc networks. In: *Proceedings of ACM Mobicom*, Boston, USA (2000)
6. Kim, Y., Perrig, A., Tsudik, G.: Group key agreement efficient in communication. *IEEE Transactions on Computers* 53(7), 905–921 (2004)
7. Rescorla, E.: Diffie-Hellman Key Agreement Method, RFC 2631, IETF Network Working Group, <http://www.ietf.org/rfc/rfc2631.txt>
8. Boyd, C., Nieto, J.M.G.: Round-Optimal Contributory Conference Key Agreement. In: Desmedt, Y.G. (ed.) *PKC 2003*. LNCS, vol. 2567, pp. 161–174. Springer, Heidelberg (2003)
9. Bennett, C.H., Gill, J.: Relative to a Random Oracle A , $P^A \neq NP^A \neq co-NP^A$ with Probability 1. *SIAM Journal on Computing* 10(1), 96–113
10. Hu, Y., Perrig, A., Johnson, D.: Packet Leashes: A Defense against Wormhole Attacks in Wireless Ad Hoc Networks. In: *Proceedings of INFOCOM* (2004)
11. Hu, L., Evans, D.: Using Directional Antennas to Prevent Wormhole Attacks. In: *Proceedings of the 11th Network and Distributed System Security Symposium* (2003)

Analytical Parametric Evaluation of Dynamic Load Balancing Algorithms in Distributed Systems

Mayuri A. Mehta¹ and Devesh C. Jinwala²

¹ Department of Computer Engineering, Sarvajani College of Engineering and Technology, Surat, India

mayuri133@yahoo.com

² Department of Computer Engineering, S. V. National Institute of Technology, Surat, India
dcj@svnit.ac.in

Abstract. With ever increasing network traffic, distributed systems can provide higher performance using a typical dynamic load balancing (DLB) algorithm. Dynamic algorithm employs up to date load information of the nodes to make load distribution decisions and therefore, they have potential to outperform static strategies. In this paper, we illustrate the analytical comparative study of existing dynamic algorithms and result gives a thorough overview of various dynamic algorithms, helping designers in choosing the most appropriate approach for a variety of distributed systems. Moreover, researchers can use it as a catalog of available DLB schemes to come up with new design.

Keywords: Load balancing, dynamic load balancing, distributed system.

1 Introduction

In a typical distributed system setting, nodes are of heterogeneous nature and also jobs arrive at the different nodes in a random fashion. This causes imbalance in workload that is harmful to the system performance in terms of mean response time of jobs and resource utilization. A load balancing mechanism can be used to remove such load imbalance. Load balancing (LB) is a technique to spread work between two or more computers, network links, CPUs, hard drives, or other resources, in order to obtain optimal resource utilization, throughput, and/or response time [1].

The two major categories for load-balancing algorithm are: static and dynamic. In static load balancing (SLB), distribution of workload is predetermined and carried out using priory knowledge of the system [2]. Dynamic load balancing (DLB) makes use of current system state information to make more effective load distribution decisions. It has been shown in literature that DLB algorithm performs better than SLB algorithm in a variety of system conditions. Several studies have been conducted on a variety of DLB algorithms. In [3], a novel neural network approach that realized the selection and location policies was described. Fuzzy based DLB algorithms were presented to improve the response time in [1][4]. The approach that considered the node heterogeneity and random communication delays was discussed in [5]. In [6],

three different DLB approaches were discussed. The DLB algorithms suitable to grid environment were illustrated in [7]. DLB schemes for multi-class jobs were presented in [8]. As per our observation, none of these existing attempts exhaustively and comparatively evaluate the majority of the DLB algorithms. Hence, our survey of the dynamic algorithms discussed further is more complete and has a wider coverage.

The rest of the paper is organized as follows: In section 2, we list the various DLB algorithms with brief descriptions. Section 3 covers parameterized comparison of these algorithms. Finally conclusion and some future work are specified in section 4.

2 Various Dynamic Load Balancing Algorithms

A New Fuzzy Approach (ANFA) [1] took into consideration the crisp inputs and effectively tolerates uncertainty and inconsistency in state information using fuzzy logic to reduce response time. *Fuzzy-Based Algorithm (FBA)* [4] that consisted of a fuzzy rule base, a fuzzy inference engine, fuzzification, and defuzzification improved overall system performance. *A Novel Neural Network Approach (NNA)* [3] used Winner-Take-All (WTA) model to realize the selection and location policies of a typical DLB algorithm and considered all delays due to network communication.

Centralized One-Shot LB Policy (COSLBP) [5] considered only one LB instant that was time at which load balancing was executed [9]. Real-time experiments conducted by authors had shown that for a given initial load and average processing rates, there existed an optimal LB gain and an optimal balancing instant associated with the one-shot LB policy, that together minimized the average overall completion time (AOCT). *A Distributed, Adaptive Sender-Initiated Policy (DASIP)* [5] which was adapted to varying system parameters such as load variability, delay, and variable runtime processing speed, minimized the average completion time per task.

In *Primary Approach (PA)* [6], overloaded primary node first tried to find out light weighted node in the same cluster and if suitable node was not found then nearby cluster was searched for the same. In *Centralized Approach (CA)* [6], centralized node in each cluster accommodated the overload of a primary node. Whenever an overloaded node did not find lightly loaded primary node, overload was transferred to centralized node. To overcome the limitations of centralized approach, there was a *Modified Approach with Supporting Nodes (MASNs)* [6]. Centralized node was split into several Supporting Nodes (SNs). Now overloaded primary node would interrupt the SN. This SN used interrupt service routine (ISR) to schedule the process.

In *Perfect Information Algorithm (PIA)* [7], when a job arrived, a processor computed the job's finish time on all buddy processors. A job was relocated on the buddy processor if it could finish the job earlier than this processor. In *Estimated Load Information Scheduling Algorithm (ELISA)* [7], the processor would migrate the jobs to lightly loaded processor, if its queue length was greater than the average queue length in its buddy set. In *Modified ELISA (MELISA)* [7], at each estimation instant, processor calculated the average load in its buddy set. Now processor would make decision of job migration taking into consideration the node's heterogeneity if its load was greater than the average load in its buddy set. *Load Balancing on Arrival Algorithm (LBAA)* [7] balanced load by transferring a job on its arrival time rather than waiting for the next transfer instant as in case of ELISA and MELISA.

Dynamic Global Optimal Scheme (DGOS) and *Dynamic Noncooperative Scheme with Communication (DNCOOPC)* [8] were extended from static schemes and intended for multi-user (multi-class) jobs. The objective of DGOS was to minimize the expected response time of all jobs over the entire system. The goal of DNCOOPC was to minimize the expected response time of the individual users.

DLB Algorithm based on Distributed Database System (DLBDDS) [10] was aimed to provide a reasonable request-response time and transaction throughput. Due to considering the communication cost and employing the collected information, this algorithm was much more efficient and had better performance.

3 Analysis

Here we present comparison of these algorithms based on certain imperative issues such as handling resource and/or network heterogeneity, associated overhead, scalability and delays due to the underlying network or processing delays at the processors. A proper DLB algorithm tolerates heterogeneity in multitude of resources, incurs minimum overhead, integrates scalability and considers delays imposed by communication networks. Table 1 demonstrates the comparison of above mentioned DLB algorithms based on these issues.

Table 1. Comparison of dynamic load balancing algorithms

Algorithm/Approach	Tolerates Heterogeneity?	Associated Overhead	Considers Delays?	Scalability Incorporated?	Performance Metric(s) Used
ANFA	Yes	More	No	Yes	RT
FBA	Yes	Less	Yes	Yes	RT, T, TT
COSLBP	Yes	Less	Yes	No	AOCT
DASIP	Yes	More	Yes	Yes	AOCT, ACTT
PA	Yes	Less	No	Yes	ET
MASNs	Yes	More	No	Moderately	ET
PIA	Yes	More	No	Yes	RT
ELISA	Yes	Less	No	Yes	RT
MELISA and LBA	Yes	More	Yes	Yes	RT,JMC,ET,RU
DGOS and DNCOOPC	Yes	Less to more	No	Yes	RT
DLBDDS	Yes	Average	No	Yes	RT, T
NNNA	No	More	Yes	Yes	Mean RT

The various metrics viz. Response Time (RT), Throughput (T), Turnaround Time (TT), Average Overall Completion Time (AOCT), Average Completion Time per Task (ACTT), Execution Time (ET), Resource Utilization (RU), and Job Migration Cost (JMC) that are used to measure performance of DLB approaches are also different for various DLB algorithms and are shown in Table 1.

4 Conclusion and Future Work

Various DLB algorithms have been proposed in literature. We provide an overview of the existing DLB algorithms in distributed systems. It reinforces the belief that dynamic algorithms potentially do better for performance improvements. The principal objective of our study on DLB algorithms is to travel around the existing DLB approaches and to be familiar with the existing work to design a new competent dynamic algorithm for a heterogeneous distributed system. As part of our future work we intend to design a new efficient DLB and compare it with the existing ones.

References

1. Karimi, A., Zarafshan, F., Jantan, A.B., Ramli, A.R., Saripan, M.B.: A New Fuzzy Approach for Dynamic Load Balancing Algorithm. *Int. J. Computer Science and Information Security* 6, 1–5 (2009)
2. Waraich, S.S.: Classification of Dynamic Load Balancing Strategies in the Network of Workstations. In: *5th Proceedings of the International Conference on Information Technology: New Generations*, pp. 1263–1265. IEEE Computer Society, Washington (2008)
3. El-Abd, A.E., El-Bendary, M.I.: A Neural Network Approach for Dynamic Load Balancing in Homogeneous Distributed System. In: *30th Proceedings of IEEE International Conference on System Sciences*, pp. 628–634. IEEE Computer Society, Washington (1997)
4. A Fuzzy-Based Dynamic Load-Balancing Algorithm, <http://jitas.im.cpu.edu.tw/2004-2/4.pdf>
5. Dhakal, S., Hayat, M.M., Pezoa, J.E., Yang, C., Bader, D.A.: Dynamic Load Balancing in Distributed Systems in the Presence of Delays: A Regeneration – Theory Approach. *IEEE Transactions on Parallel and Distributed Systems* 18, 485–497 (2007)
6. Jain, P., Gupta, D.: An Algorithm for Dynamic Load Balancing in Distributed Systems with Multiple Supporting Nodes by Exploiting the Interrupt Service. *Int. J. Recent Trends in Engineering* 1, 232–236 (2009)
7. Shah, R., Veeravalli, B., Misra, M.: On the Design of Adaptive and Decentralized Load Balancing Algorithms with Load Estimation for Computational Grid Environment. *IEEE Transactions on Parallel and Distributed Systems* 18, 1675–1686 (2007)
8. Penmasta, S., Chronopoulos, A.T.: Dynamic Multi-User Load Balancing in Distributed Systems. In: *21st Proceedings of IEEE International Parallel and Distributed Processing Symposium*, pp. 1–10. IEEE Press, California (2007)
9. Dhakal, S., Paskaleva, B.S., Hayat, M.M., Schamiloglu, E., Abdalla, C.T.: Dynamical Discrete-Time Load Balancing in Distributed Systems in the Presence of Time Delays. In: *42nd Proceedings of IEEE Conference on Decision and Control*, pp. 5128–5134. IEEE Press, Piskataway (2003)
10. Feng, Y., Li, E., Wu, H., Zhang, Y.: A Dynamic Load Balancing Algorithm Based on Distributed Database System. In: *4th Proceedings of International Conference and Exhibition on High-Performance Computing in the Asia-Pacific Region*, pp. 949–951. IEEE Computer Society, Beijing (2000)

Wavelet Based Electrocardiogram Compression at Different Quantization Levels

A. Kumar* and Ranjeet

Indian Institute of Information Tehncology Design and Manufacturing,
Jabalpur, MP-482005, India
anilkdee@gmail.com, ranjeet281@gmail.com

Abstract. In this paper, a wavelet based electrocardiogram (ECG) data compression technique is reviewed. The method employs the discrete wavelet transform (DWT), thresholding, Huffman encoding followed by different quantization levels. A comparative study of performance at the different quantization levels and thresholding is made in terms of Signal-to-noise ratio (*SNR*), Percent root mean square difference (*PRD*) and Mean square error (*MSE*). The simulation results illustrates that good compression ratio can be achieved at lower quantization levels, while at higher quantization levels, all fidelity measuring parameters are enhanced.

Keywords: DWT, Thresholding, Quantization and Huffman Encoding.

1 Introduction

An electrocardiogram (ECG) is the graphical representation of electrical impulses due to ionic activity in the cardiac muscles of human heart. It is an important physiological signal which is exploited to diagnose heart diseases because every arrhythmia in ECG signals can be relevant to a heart disease [1]. ECG signals are recorded from patients for both monitoring and diagnostic purposes. Therefore, the storage of computerized is become necessary. However, the storage has limitation which has made ECG data compression as an important issue of research in biomedical signal processing. In addition to these, there are many advantages of ECG compression such as transmission speed of real-time ECG signal is enhanced and is also economical.

Several efficient methods [2-9] are available in literature which involve in compression schemes of ECG signal without losing and preserving the relevant clinical information for the accurate detection and classification. These schemes were classified into three categories [3]: dedicated techniques such as AZTEC, FAN, CORTES, and turning point. These techniques were based on the detection and elimination of redundancies on direct analysis of the original signal, and gives minimum distortion. In second category, all transform based techniques come and here, compression is achieved based on spectral and energy distribution of the signal. Other hand, the last technique is based on feature and parameter extraction in which some parameters

* Corresponding author.

such as measurement of the probability distribution of the original signal is extracted. During the last two decades, several efficient methods have reported in literature, which involve compression of ECG signal without losing and preserving the relevant clinical information for the accurate detection and classification. Multi-resolution decomposition of signal is efficient for extracting the content information [10]. In this technique, wavelet transform has been exploited for the ECG processing and extracting the information. Recently, several other methods [6-9] have been developed based on wavelet or wavelet packets promise that it is an efficient power tool for compressing and analysis of ECG signal.

In this paper, wavelet based ECG compression technique is reviewed and the effect of different quantization levels on compression is explored.

2 Discrete Wavelet Transform

Wavelets transform is a method to analyze a signal in time and frequency domain, it is an effective tool for the analysis of time-varying non stationary signal like ECG [7]. Wavelet transform gives the multiresolution decomposition of the signal. There are three basic concepts of multiresolution: subband coding, vector space and pyramid structure coding [10]. DWT decomposes a signal at several n levels in different frequency bands. Each level decomposes a signal into the approximation coefficients (low frequency band of processing signal) and the detail coefficients (high frequency band of processing signal) [10] as show in Fig. 1.

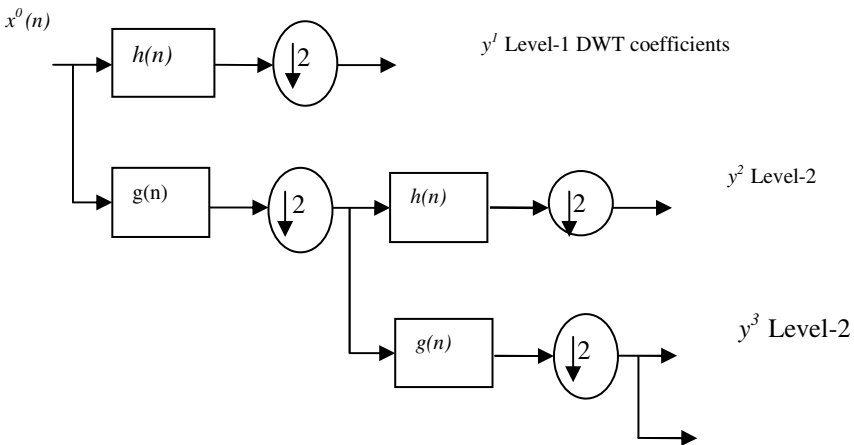


Fig. 1. Filter bank representation of DWT decomposition

At each step of DWT decomposition, there are two outputs: the scaling coefficients $x^{j+1}(n)$ and the wavelet coefficients $y^{j+1}(n)$. These coefficients are given:

$$x^{j+1}(n) = \sum_{i=1}^{2n} h(2n-i)x^j(n) \quad (1)$$

and

$$y^{j+1}(n) = \sum_{i=1}^{2n} g(2n-i)x^j(n) \quad (2)$$

where, the original signal is represented by $x^0(n)$ and j show the scaling number. Here $g(n)$ and $h(n)$ represent the low pass and high pass filter, respectively. The output of scaling function is input of next level of decomposition, known as approximation coefficients. The approximation coefficients are low-pass filter coefficients and high-pass filter coefficient are detail coefficients of any decomposed signal.

3 Quantization

Quantization is a process of representation of a set of continues value to finite discrete set of values. A signal divides into a number of interval, each interval having own codeword in the quantized value. In wavelet based compression, after thresholding the wavelet coefficients vector $x^j(n)$ is quantized [11-13]. Due to quantization process, the perfect reconstruction of the original signal is not possible at the reconstruction side. The quantization process depends on these parameters: maximum value (M_{\max}), minimum value (M_{\min}) in the signal and number of quantization level $L=2^m$ (An m -bit uniform quantizer is used). Once these parameters are found, then step size (Δ) is computed by

$$\Delta = (M_{\max} - M_{\min})/L \quad (3)$$

In the uniform quantization, step-size Δ is depend upon the number of quantization levels, its associated with the value of m . which captains the information in form of bit/symbol of quantized signal. A detailed discussion on the quantization process is given in [11-13].

4 Methodology for ECG Compression

A wavelet based methodology of ECG compression is shown in Fig. 2. This technique involves three steps for the ECG signal compression: DWT decomposition, threshold and quantization, and entropy encoding. After DWT decomposition of ECG signal, its wavelet coefficients are selected on the basis of energy packing efficiency of each subband. After decomposition of the ECG signal, a thresholding is applied to the wavelet coefficients, which makes a fixed percentage of wavelet coefficients equal to zero. There are two types of the thresholding: global and level thresholding. In level thresholding, the threshold value is calculated using Birge-Massart strategy [14, 15]. While, in global thresholding, the threshold value is set manually, this value is chosen from wavelet coefficient ($0 \dots x_{\max}^j$) where x_{\max}^j is maximum coefficient in the decomposition. Detailed discussion on thresholding is given in [12-16].

Further, uniform quantization is performed on these coefficients. The actual compression is achieved at this stage and this compression can be further enhanced with the help of entropy encoding technique (Huffman) [12, 13, 16]. In Huffman encoding, the probabilities of occurrence of the symbols in the signal are computed. These symbols indices in the quantization table, these symbols arranged according to the probabilities of occurrence in descending order and a binary tree and codeword table is created. Finally, the compressed ECG signal is obtained at the output of entropy encoder.

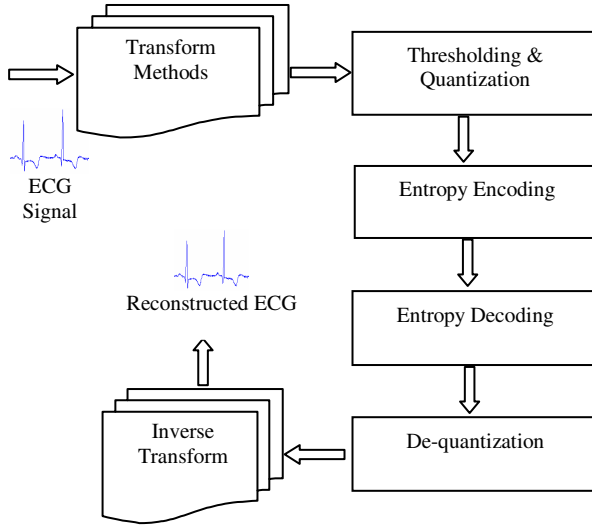


Fig. 2. Compression methodology for ECG signals

5 Methodology for ECG Compression

In this paper, ECG signal compression is achieved using the methodology discussed in Section IV and the effect of different quantization level is seen on the compression. The performance is evaluated by considering the fidelity of the reconstructed signal to the original signal. For this, many fidelity assessment parameters are considered such as Compression ratio (CR), Percent root mean square difference (PRD), Mean square error (MSE) and Signal to noise ratio (SNR) given in [11-16]: ECG records have been obtained from MIT-BIH Arrhythmia Database [17]. Here, different wavelet filters, and global thresholding are exploited for signal compression. The simulation results obtained in each case are included in Table 1. A comparative analysis of different m -bit quantizer at different thresholds is depicted in Fig. 3.

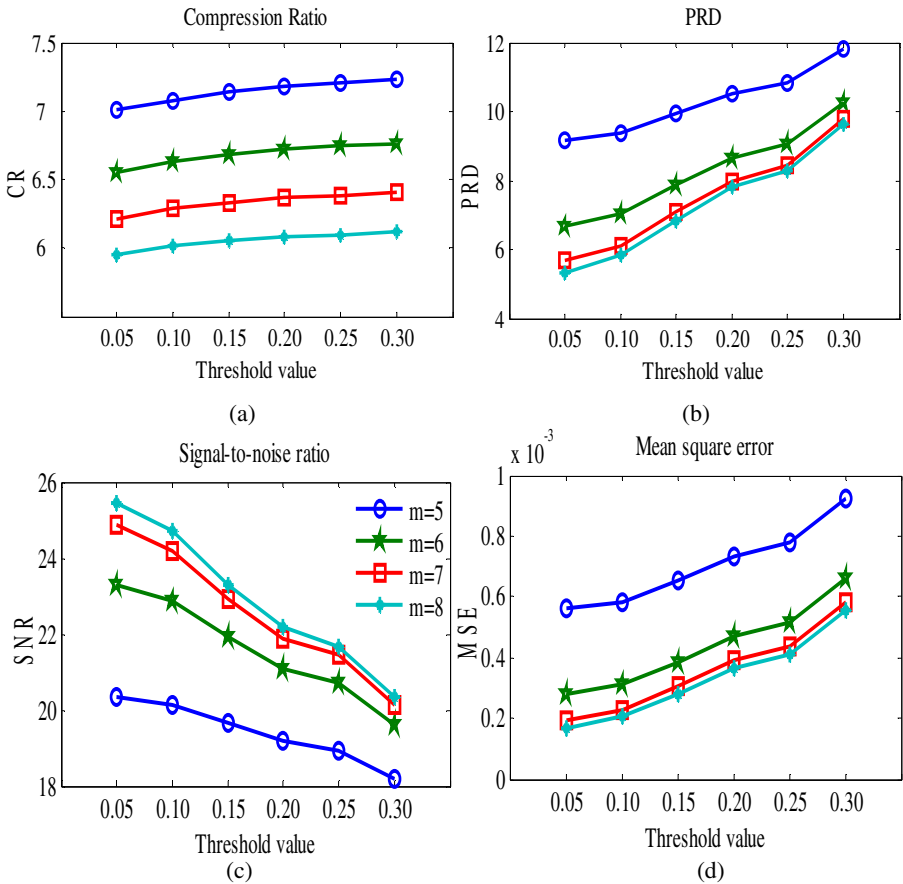


Fig. 3. Variation of performance measuring parameters of wavelet based compression with different threshold value and quantization level. (a) CR (b) PRD (c) SNR (d) MSE

It is evident from Table I that at low level quantization, compression ratio is more as compared to high level quantization. While, at higher quantization level, all fidelity measuring parameters are more improved and gives good compression as well as preserving more.

6 Conclusions

In this paper, a wavelet based methodology is presented for the ECG signal compression. A comparative study of performance of different uniform quantization levels for the ECG signal compression is explored. The simulation results obtained illustrate that good compression ratio can be achieved at low level quantization and good reconstruction of the original signal can be achieved at higher quantization levels.

Table 1. Variation of fidelity measuring parameter at different quantization levels

Wavelet Filters	Quantization Levels	<i>CR</i>	<i>PRD</i>	<i>MSE</i>	<i>SNR</i>
Haar	5	8.53	11.33	8.51×10^{-4}	18.51
db10	5	8.37	10.53	7.32×10^{-4}	19.16
coif5	5	8.08	9.33	5.84×10^{-4}	20.14
sym8	5	8.53	9.96	6.62×10^{-4}	19.59
Haar	6	7.55	9.51	5.66×10^{-4}	20.27
db10	6	7.45	8.65	4.70×10^{-4}	21.09
coif5	6	7.14	7.01	3.12×10^{-4}	22.86
sym8	6	7.60	7.79	3.84×10^{-4}	21.96
Haar	7	6.82	9.04	5.00×10^{-4}	20.81
db10	7	6.74	8.00	3.91×10^{-4}	21.88
coif5	7	6.49	6.25	2.41×10^{-4}	23.99
sym8	7	6.82	7.15	3.14×10^{-4}	22.83
Haar	8	6.26	8.94	4.83×10^{-4}	20.97
db10	8	6.16	7.80	3.66×10^{-4}	22.17
coif5	8	5.95	6.06	2.23×10^{-4}	24.31
sym8	8	6.24	6.98	2.96×10^{-4}	23.09

References

1. Martis, R.J., Chakraborty, C., Ray, A.K.: An integrated ECG feature extraction scheme using PCA and wavelet transform. In: Proceedings of IEEE India Council Conference (INDICON 2009), pp. 88–93 (2009)
2. Sastry, R.V.S., Rajgopal, K.: ECG Compression using Wavelet Transform. In: Proceedings of RC IEEE-EMBS, BMESI, vol. 14, pp. 499–500 (1995)
3. Nielsen, M., Kamavuako, E.N., Andersen, M.M., Lucas, M.-F.L., Farina, D.: Optimal wavelets for biomedical signal compression. *Med. Bio. Eng. Comput.* 44, 561–568 (2006)
4. Khorrami, H., Moavenian, M.: A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification. *Expert Systems with Applications* 37(8), 5751–5757 (2010)
5. Manikandan, M.S., Dandapat, S.: Wavelet threshold based TDL and TDR algorithms for real-time ECG signal compression. *Biomedical Signal Processing and Control* 3(1), 44–66 (2008)
6. Hunga, K.-C., Tsaia, C.-F., Ku, C.-T., Wanga, H.S.: A linear quality control design for high efficient wavelet-based ECG data compression. *Computer Methods and Programs in Biomedicine* 9(4), 109–117 (2009)
7. Saritha, C., Sukanya, V., Murthy, Y.N.: ECG Signal Analysis Using Wavelet Transforms. *Bulg. J. Physics* 35, 68–77 (2008)
8. Ahmed, S.M., Al-Shrouf, A., Abo-Zahhad, M.: ECG data compression using optimal non-orthogonal wavelet transform. *Medical Engineering & Physics* 22(1), 39–46 (2000)

9. Rajoub, B.A.: An Efficient Coding Algorithm for the Compression of ECG Signals Using the Wavelet Transform. *IEEE Transactions on Biomedical Engineering* 49(4), 355–362 (2002)
10. Mallat, S.G.: A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7) (1989)
11. Manikandan, M.S., Dandapat, S.: Wavelet threshold based ECG compression using USZZQ and Huffman coding of DSM. *Biomedical Signal Processing and Control* 1, 261–270 (2006)
12. Chen, J., Wang, F., Zhang, Y., Shi, X.: ECG compression using uniform scalar dead-zone quantization and conditional entropy coding. *Medical Engineering & Physics* 30, 523–530 (2008)
13. Ebrahimzadeh, A., Azarbad, M.: ECG Compression using Wavelet transform and Three-Level Quantization. In: *IEEE Conference proceedings IDC*, vol. 6, pp. 250–254 (2010)
14. Misiti, M., Misiti, Y., Oppenheim, G., Poggi, J.: *Matlab tool box*. The Math Works, Inc. (2000)
15. Najih, A.M.M.A., Ramli, A.R.B., Prakash, V., Syed, A.R.: Speech Compression using Discrete Wavelet Transform. In: *4th National Conference on Telecommunication Technology Proceedings*, Shah Alam, Malaysia, pp. 1–4 (2003)
16. Tohumoglu, G., Sezgin, K.E.: ECG signal compression by multi-iteration EZW coding for different wavelets and thresholds. *Computers in Biology and Medicine* 37(2), 173–182 (2007)
17. <http://www.physionet.org>

Content Based Image Retrieval by Using an Integrated Matching Technique Based on Most Similar Highest Priority Principle on the Color and Texture Features of the Image Sub-blocks

Ch. Kavitha¹, M. Babu Rao², B. Prabhakara Rao³, and A. Govardhan⁴

¹ Associate Professor, IT department, Gudlavalleru Engineering College,
Gudlavalleru, Krishna district, Andhra Pradesh, India
kavithachaduvula@yahoo.com

² Associate Professor, CSE department, Gudlavalleru Engineering College,
Gudlavalleru, Krishna district, Andhra Pradesh, India
baburaompd@yahoo.co.in

³ Director of Evaluation, JNTUK, Kakinada, A.P, India

⁴ Principal, JNTUH college of Engineering, Jagtial, A.P, India

Abstract. In this paper, we propose an efficient technique for content based image retrieval which uses the local color and texture features of the image. Firstly the image is divided into sub blocks of equal size. The color and texture features of each sub-block are computed. Color of each sub-block is extracted by quantifying the HSV color space into non-equal intervals and the color feature is represented by cumulative histogram. Texture of each sub-block is obtained by using gray level co-occurrence matrix. An integrated matching scheme based on Most Similar Highest Priority principle is used to compare the query and target image. The adjacency matrix of a bipartite graph is formed using the sub-blocks of query and target image. This matrix is used for matching the images. Euclidean distance is used in retrieving the similar images. The efficiency of the method is demonstrated with the results.

Keywords: Image retrieval, color, texture, GLCM, integrated matching.

1 Introduction

Content based image retrieval is a technique for extracting similar images from an image database using low level features of an image. The need for efficient image retrieval is increased tremendously[1]. There are various CBIR systems which used global features [2], [3], [4] and local features [4]. From these systems it is clear that local features play a significant role in determining similarity of images.

Color and texture are the most important visual features. Because of the advantage of HSV color space is its ability to separate chromatic and achromatic components. we selected the HSV color space to extract the color features. Texture feature is a kind of visual characteristics that does not rely on color or intensity and reflects the

intrinsic phenomenon of images. So we developed a technique which captures color and texture features of sub-blocks of the image.

2 Proposed Method

2.1 Partitioning the Image into Sub-blocks

Firstly the image is partitioned into 6 (2X3) equal sized sub-blocks. The size of the sub-block in an image of size 256X384 is 128X128. The images with other than 256X384 size are resized to 256X384.

2.2 Extraction of Color of an Image Sub-block

Unequal interval quantization HSV color space according the human color perception has been applied on *H, S, and V* components. So, we divided color into eight parts. Saturation and intensity is divided into three parts separately. The quantified hue(H), saturation(S) and value(V) are showed as equation 1.

$$H = \begin{cases} 0 & \text{if } h \in [316, 20] \\ 1 & \text{if } h \in [21, 40] \\ 2 & \text{if } h \in [41, 75] \\ 3 & \text{if } h \in [76, 155] \\ 4 & \text{if } h \in [156, 190] \\ 5 & \text{if } h \in [191, 270] \\ 6 & \text{if } h \in [271, 295] \\ 7 & \text{if } h \in [296, 315] \end{cases} \quad S = \begin{cases} 0 & \text{if } s \in [0, 0.2) \\ 1 & \text{if } s \in [0.2, 0.7) \\ 2 & \text{if } s \in [0.7, 1) \end{cases} \quad (1)$$

$$V = \begin{cases} 0 & \text{if } v \in [0, 0.2) \\ 1 & \text{if } v \in [0.2, 0.7) \\ 2 & \text{if } v \in [0.7, 1) \end{cases}$$

In accordance with the quantization level above, three-dimensional feature vector for different values of H,S,V with different weight to form one-dimensional feature vector named G:

$$G = 9H + 3S + V \tag{2}$$

This paper represents the one-dimensional vector G by constructing a cumulative histogram of the color characteristics of image after using non-interval HSV quantization for G.

2.3 Extraction of Texture of an Image Sub-block

GLCM creates a matrix with the directions and distances between pixels, and then extracts meaningful statistics from the matrix as texture features. GLCM texture features commonly used are shown in the following:

$$\text{Energy } E = \sum_x \sum_y P(x, y)^2 \tag{3}$$

$$\text{Contrast } I = \sum_x \sum_y (x - y)^2 P(x, y) \tag{4}$$

$$\text{Entropy } S = - \sum_x \sum_y P(x, y) \log P(x, y) \tag{5}$$

$$\text{Inverse difference } H = \sum_x \sum_y \frac{1}{1 + (x - y)^2} P(x, y) \quad (6)$$

2.4 Integrated Image Matching

We have designed an algorithm for finding the minimum cost matching based on most similar highest priority (MSHP) principle using the adjacency matrix of the bipartite graph. The minimum distance d_{ij} of this matrix is found between sub-blocks i of query and j of target. The complexity of the matching procedure is reduced from $O(n^2)$ to $O(n)$, where n is the number of sub-blocks involved. The integrated minimum cost match distance between images is defined as: $D_{qt} = \sum_i \sum_j d_{ij}$, Where $i=1,2, \dots, n$ $j=1,2, \dots, n$. And d_{ij} is the best-match distance between sub-block i of query image q and sub-block j of target image t and D_{qt} is the distance between images q and t .

3 Experimental Setup

3.1 Data Set and Feature Set

Wang's dataset comprising of 1000 Corel images with ground truth. The image set comprises 100 images in each of 10 categories. The images are of the size 256 x 384. The feature set comprises color and texture descriptors computed for each sub-block of an image as we discussed in section 2.

3.2 Computation of Similarity

Matching of the sub-blocks is done based on the most similar highest principle. We construct the Euclidean calculation model as follows:

$$D(A, B) = \omega_1 D(F_{CA}, F_{CB}) + \omega_2 D(F_{TA}, F_{TB}) \quad (7)$$

Here ω_1 is the weight of color features, ω_2 is the weight of texture features, F_{CA} and F_{CB} represents the color features for image A and B. For a method based on GLCM, F_{TA} and F_{TB} on behalf of texture features correspond to image A and B. Here, we combine color features and texture features. The value of ω through experiments shows that at the time $\omega_1 = \omega_2 = 0.5$ has better retrieval performance.

4 Experimental Results

The experiments were carried out as explained in sections 2 and 3. The results are benchmarked with some of the existing systems using the same database[15]. The query images are shown in Fig 1. The static precision of the images (a),(b),(c) and (d) in various techniques of 20 random images is represented in the form of Table 1. From the results, it is observed that our proposed method improves the precision over the other methods.

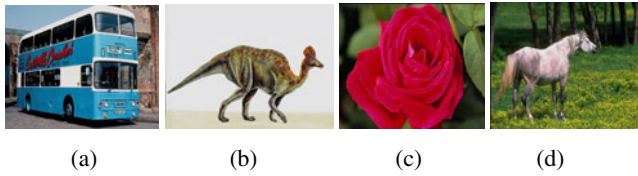


Fig. 1. Query images

Table 1. The Static Precision

P(20)	Different image retrieval techniques		
	HSVcolor+GLCM Of image	HSVcolor+GLCM of image sub-blocks	HSVcolor+GLCM of image sub-blocks with matching based on MSHP principle
A	0.45	0.5	0.6
B	0.3	0.6	0.75
C	0.8	0.85	1.0
D	0.35	0.55	0.8
Average precision	0.475	0.625	0.7875

5 Conclusions

In this paper a new image retrieval method based on color and texture features of image sub-blocks and an integrated matching scheme based on Most Similar Highest priority (MSHP) to match the images is proposed. Our experiment results demonstrate that the proposed method has better retrieval performance compared other retrieval techniques.

References

- [1] Datta, R., Joshi, D., Li, J., Wang, J.: Image Retrieval:Ideas, Influences and trends of the New Age. In: Proceedings of the 7th ACM SIGMM international workshop on multimedia information retrieval, Hilton, Singapore, November 10-11 (2005)
- [2] Niblack, W., et al.: The QBIC Project: Querying Images by Content Using Color, Texture, and Shape. In: Proc. SPIE, San Jose, CA, February 1993, vol. 1908, pp. 173–187 (1993)
- [3] Pentland, A., Picard, R., Sclaroff, S.: Photobook: Content-based Manipulation of Image Databases. In: Proc. SPIE Storage and Retrieval for Image and Video Databases II, San Jose, CA, February 1994, pp. 34–47 (1994)
- [4] Stricker, M., Orengo, M.: Similarity of Color Images. In: Proc. SPIE Storage and Retrieval for Image and Video Databases, February 1995, pp. 381–392 (1995)

Understanding the Impact of Cache Performance on Multi-core Architectures

N. Ramasubramaniam, V.V. Srinivas, and P. Pavan Kumar

Department of Computer Science and Engineering,
National Institute of Technology - Tiruchirappalli
{nrs, 206110022, 206109026}@nitt.edu

Abstract. Research, recent and old, paid much attention on reducing the cycle time and increasing the speed of execution of the processor. This led to the development of multiple core processors that distribute and share load among the many processors. The important question to answer would be, will the shared cache technology for multi-core work as efficiently as it did for uni-core processors? Our analysis in this paper, takes into consideration the impact that caches have on a uni-core environment and in a shared multi-core environment. In order to demonstrate this we use: DINERO IV for analysing the performance of uni-core environment and CACTI (Cache Access Cycle Time Indicator) for analysing the performance of multi-core environment.

Keywords: CACTI, Cache memory, performance, multi-core, shared cache.

1 Introduction

Memory hierarchy has seen a lot of improvement in recent years [8]. Cache technology serves to bridge the gap between the high speed processor and the low speed devices [5]. Recent trends show that, in order to increase the data transfer rate of the secondary memories, cache memories are placed on the secondary storage to pre-record data from secondary storage for faster access. Of-late, not much work has been done about the impact of cache memory on multi-core architectures [7]. Though the speed at which main memories or the secondary memories operate cannot be compared with the speed of processors, the access time of memories have come down and the execution time in terms of clock cycle per second has been reduced, due to the increase in the number of cores inside the processor. Moreover the cost per gigabyte of the secondary memories is also very much lower. [1] suggests the impact of cache on access time and cycle time in a uni-core environment.

1.1 Contribution

The key focus of this paper is to analyze the impact that a cache would produce when used in a uniprocessor environment and a multi-core environment. The entire paper is organized as follows: Section 2 describes the description of various tools used followed by Section 3, dealing with the experimental setup and the parameters used. Section 4 evaluates the results obtained. Section 5 concludes with the future research.

2 Tools Used

DINERO IV [2] is a cache simulator for uni-processor environment. This simulator is capable of simulating multi-level caches. It takes in a configuration file for simulating the performance of cache. The access time for DINERO is calculated based on the formula given below. Here hit time is taken as 1 clock cycle. Miss rate is taken from the simulation result and the miss penalty is taken as 8 clock cycles for L1 cache. Cycle time for L1 cache is 7.5 nanoseconds [9]. CACTI [3] is an analytical tool developed to evaluate cache access and cycle time. This tool was developed based on Wada et. als, model considers both direct and set associative cache. The computation for cache access time is done as a function of cache size, associativity and cache block size [4]. In this paper, we have used version 6.5 for all the experiments. CACTI can be accessed as a web-interfaced version [10] as well as source code version.

$$\text{Access time} = \text{Hit time} + \text{Misrate} * \text{Misspenalty}$$

3 Experimental Setup

CACTI 6.5 takes a configuration file as input. The parameters used in experiments are: cache size of 2097152,134217728 bytes with each cache consisting of 64 blocks. We have considered 45 nm chip technology in our experiments. For DINERO IV we have taken the values closest to the parameters mentioned since it does not support some of the parameters used by CACTI, such as chip technology. In this paper we have purposefully chosen these cache sizes because these are some of the cache sizes supported by CACTI-6.5. Much attention is given in this paper to CACTI-6.5 because, recent technology has evolved from uni-core to multi-core technology.

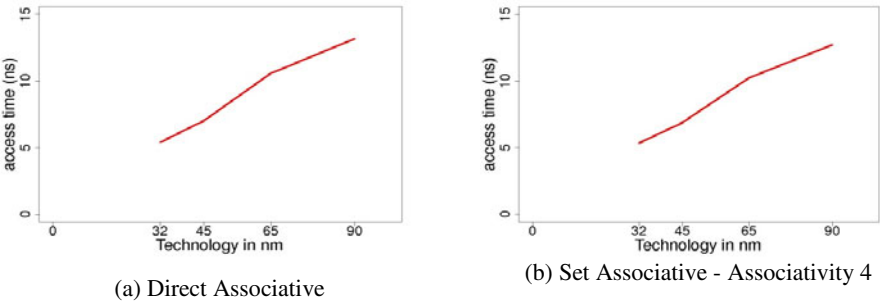
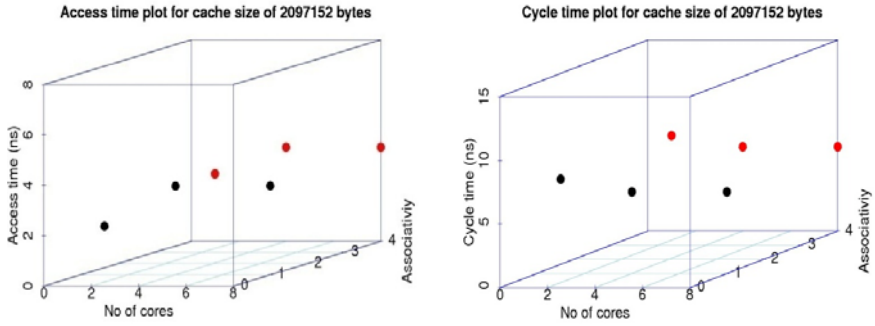


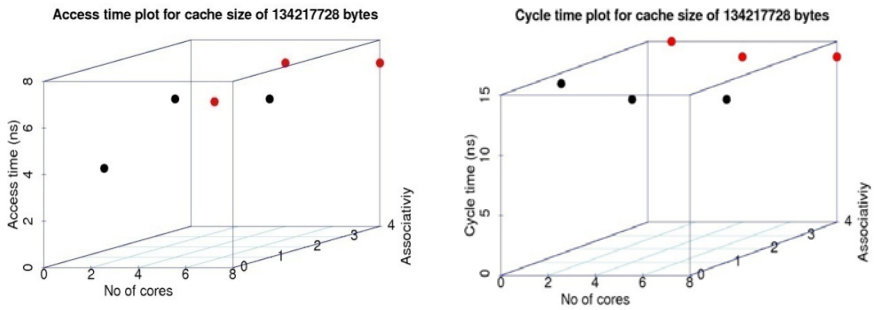
Fig. 1. Access time (ns) and Cycle time (ns) for varying manufacturing technology

4 Results Obtained and Avenues for Future Research

From Fig. 1.a and 1.c we can visualize that as the number of cores increases the cache access time increases, whereas the cycle time decreases; see Fig. 1.b and 1.d. This behaviour can be attributed to the access of shared cache by multiple processors. The



(a) Access time for cache size of 2097152 bytes in ns. (b) Cycle time for cache size of 2097152 bytes in ns.



(c) Access time for cache size of 13421728 bytes in ns. (d) Cycle time for cache size of 13421728 bytes in ns.

Fig. 2. Here x axis represents the number of cores. Y axis represents the access/cycle time in nanoseconds. Z axis represents Associativity (Here we have taken associativity of 1 and 4. The plane on front represents associativity of 1 and denoted by black dots, the plane in the rear represents associativity of 4 and denoted by red dots. The values for number of cores =1 is computed from the results obtained from DINERO IV).

Table 1. CACTI results

Technology	32	32	45	45	65	65	90	90
Number of cache banks	1	2	1	2	1	2	1	2
Access time (ns)	5.3939	5.3417	7.0117	6.879	10.59	10.27	13.1722	12.7284
Cycle time (ns)	11.01	10.95	13.72	13.575	20.14	19.79	24.1068	23.5985

only difference between Fig. 1.a and Fig. 1.c is that the range of the scale differs as the cache size increases from 2097152 and 13421728 bytes. From Fig. 2.a and 2.b we can infer that the access time for cache decreases as the technology moves from 90nm to 32 nm. This is due to the small structure, distance between the cache and the processor and high speed buses connecting cache and the processor. Further research would be to improve the models of CACTI and DINERO IV tools and to compare the results obtained from the tools with that of the real time system for the given set of

cache parameters. As mentioned in our previous paper [6], this serves as one more step in understanding the performance of multi-core architectures.

5 Conclusion

In this paper the authors have made an attempt in understanding the performance of multi-core architectures with cache memories. Thus in this paper, the time for accessing the cache memory and execution time of the processor have been understood using tools such as DINERO IV for uni-core architecture and CACTI-6.5 for multi-core architecture.

References

1. Muralimanohart, N., Balasubramonian, R., Jouppi, N.P.: Architecting Efficient Interconnections for Large Caches with CACTI-6.0. *Proceedings of IEEE Micro*, 28, 69–79 (2008)
2. Dinero IV, Trace Driven Uniprocessor Cache Simulator, <http://pages.cs.wisc.edu/~markhill/DineroIV/>
3. CACTI-6.5 (Cache Access Cycle Time Indicator), <http://www.hpl.hp.com/research/cacti/>
4. Smith, A.J.: Line (block) size choice for CPU cache memories. *Proceedings of IEEE Transactions on Computers* 100, 1063–1075 (2009)
5. Goodman, J.R.: Using cache memory to reduce processor-memory traffic. In: *Proceedings of International Symposia on Computer Architecture*, pp. 255–262 (1998)
6. Srinivas, V.V., Ramasubramaniam, N.: Understanding the performance of multicore architecture. Accepted in *International Conference in Communication, Network and Computing* (2011)
7. Banakar, R., Steinke, S., Lee, B.S., Balakrishnan, M., Marwedel, P.: Scratchpad memory: design alternatives for cache on-chip memory in embedded systems. In: *Proceedings of the Tenth International symposium on Hardware/Software Co design*, pp. 73–78 (2002)
8. Hennessy, J.L., Patterson, D.A.: *Computer Architecture: A Quantitative Approach*, 4th edn. Elsevier Inc., Amsterdam (2007)
9. Liu, Z., Zheng, K., Liu, B.: Hybrid cache architecture for high speed packet processing. *Proceedings of Computers and Digital Techniques* 1, 105–112 (2007)
10. CACTI web interface version, <http://quid.hpl.hp.com:9081/cacti/>

An Overview of Solution Approaches for Assignment Problem in Wireless Telecommunication Network

K. Rajalakshmi and M. Hima Bindu

Department of Computer Science,
Jaypee Institute of Information Technology, Noida
rajalakshmi_krishna@yahoo.com, hima.bindu@jiit.ac.in

Abstract. Currently, there is great urgency to reduce the cost of wireless telecommunication networks in order to minimize the number and cost of required facilities, and thus adds to search for optimal network designs. One such solution is optimal assignment of cells to switch in wireless telecommunication network. In this paper, we explored three basic wireless telecommunication networks, which share analogous network architecture namely Personal Communication Services (PCS), Cellular Mobile Network and Universal Mobile Telecommunication Services (UMTS). The optimal assignment problem is NP hard Complex Integer Programming problem. In this paper, we explored various issues like network architecture, formulation of objective function, constraints of objective function, algorithm used, optimization strategy, and intensification and diversification methods in algorithms.

Keywords: wireless telecommunication networks, assignment problem, NP hard, complex integer programming problem, objective function, constraints.

1 Introduction

Drastic increase in the use of mobile phone systems and the need for various adaptive demands has attracted recent research attention towards wireless telecommunication networks. In this paper, we consider three basic network architectures namely Personal Communication Services (PCS), Cellular Mobile Network, and Universal Mobile Telecommunication Services (UMTS). Broadly, the above three wireless networks share analogous network architecture. The wireless telecommunication network consists of geographically distributed hexagonal structure called cells. These cells are interconnected in a hierarchical order. Each cell has a Base Transmission Station (BTS). For communication, any registered cellular mobile device has to transmit through BTS, in turn connected through Base Station controllers (BSCs). The core network consists of Public switched Network (PSTN) and IP core network. Conventionally, cells and switches are stationary and their locations are already known. As given by Houeto et al. [2], in the real world, large proportion of budget has been allocated to the costs of network facilities that carry traffic from cell sites to switches and thus there exists pressure to reduce costs. One of the key ways to achieve cost reduction is optimal design of wireless telecommunication networks. These networks

involve the assignment of each cell to a switch, while taking into account a certain number of constraints including capacity constraints, routing redistribution and hand-offs frequency.

2 Wireless Telecommunication Networks

Basically wireless telecommunication networks like PCN, Cellular Mobile Network, and UMTS share analogous network architecture; their difference lies in the values of the four parameters, namely, frequency band, distance between cell sites, rate of data transfer and technology used by the communication network. As given by Grillo et al. [4], the frequency band of PCS is 1.7 GHz to 1.88 GHz, which is comparatively higher than the frequency band width of Cellular Mobile network with 800 MHz and UMTS frequency with 1885-2025 MHz. Chamberland et al. in [5] detail the design problem of Cellular Mobile Network. In terms of, the number of cell sites required, PCS needs cells sites within few feet apart, impacting need for more number of network devices, which as per Houeto et al. [2] is the major factor for increase in network operating cost. UMTS is a third generation telecommunication technology with a data transfer rate of 42Mbs, higher than PCS and Cellular Mobile Network. Most common technology used by UMTS is W-CDMA.

3 Objective Functions

Various parameters to formulate the preliminary objective function, for single homed signal handoff assignment problem are listed below:

- N: Number of cells in the Network
- M: Number of Switches in the Network
- C_{ik} : Cable cost for existing link between cell i and switch k
- H_{ij} : Handoff Cost between cell i and cell j per unit time
- X_{ik} : 1 if cell i and switch k are connected, otherwise zero.

As given in [1], along with assigning cells to switches in an optimal manner using link cost, handoff cost is also considered. The handoff cost consists of both complex handoff and the simple handoff cost. Let H_{ij} be the cost per unit of time for a simple handoff between cells i and j , providing cell i and j are connected to the same switch. Let H'_{ij} be the cost per unit time for a complex handoff between cells i and j , connected to different switches. Thus the objective function is given as

$$\min \left(\sum_{i=1}^N \sum_{k=1}^M C_{ik} X_{ik} + \sum_{i=1}^N \sum_{j=1}^N H_{ij} + \sum_{i=1}^N \sum_{j=1}^N H'_{ij} \right) \quad (1)$$

In [2] [6], H_{ij} and H'_{ij} are proportional to the handoff frequency between cells i and j and thus deduces the overall handoff cost as,

$$h_{ij} = H'_{ij} - H_{ij} \quad (2)$$

As per Quintero et al. [3], the total operating cost of a cellular network includes the monthly amortization cost of installed switches. Let INS_k express the monthly amortization cost for each installed switch k ($k = 1 \dots m$). Thus objective function given in equation (2) becomes,

$$\min \left(\sum_{i=1}^N \sum_{k=1}^M C_{ik} X_{ik} + \sum_{k=1}^M INS_k + \sum_{i=1}^N \sum_{j=1}^N H_{ij} + \sum_{i=1}^N \sum_{j=1}^N H'_{ij} \right) \tag{3}$$

4 Constraints

In general the objective function of minimization of total network operating cost is constrained by factors like (1) single homed constraint (2) switch constraint and (3) handoff constraint.

Single homed constraint:

The single homed cell indicates that each cell is connected to only one single switch, that is, cell i is linked to only one switch at any time.

Switch capacity constraint:

The number of connectivity requested through BTS of each cell for a particular switch, must be less than or equal to the remaining call handling capacity P_k currently available at switch k .

Handoff constraints:

To formulate the handoff cost per unit, we introduce

$$Z_{ijk} = X_{ik} X_{jk} \text{ for } i, j = 1, 2, 3 \dots N \text{ and } k = 1, 2, \dots, M \tag{4}$$

When cell i and cell j are connected to same switch k , then Z_{ijk} equals one and otherwise equals zero. Equation (4) is a nonlinear binary product. Converting this nonlinear binary constraint into linear binary constraint we have

$$Z_{ijk} \leq X_{ik} ; Z_{ijk} \leq X_{jk} ; Z_{ijk} \geq X_{ik} + X_{jk} - 1 ; Z_{ijk} \geq 0 ; \tag{5}$$

The generalized handoff constraint between two cells is given by

$$Y_{ij} = \sum_{k=1}^M Z_{ijk} \text{ for } i \neq j \text{ and } i, j = 1, 2, \dots, N \tag{6}$$

5 Algorithmes

In 1995, Merchant et al. [1] introduced heuristic methods for solving NP-hard Assignment Problem. Comparative analysis of Integer Linear Programming with Heuristics confirmed the excellent performance of heuristic algorithm. However, the heuristic algorithm exhibits high probability of getting struck at local optima. The Taboo

Search method is an improvement of the general descending algorithm, because it attempts to avoid the trap of local minimum as in [2]. The taboo list is used to avoid the solution already explored during the generation of the set of neighbor candidates. Quintero et al. [3] proposes Genetic Algorithms, which are robust search techniques based on natural selection and genetic production mechanisms perform a search by evolving a population of candidate solutions through nondeterministic operators and by incrementally improving the individual solutions through selection, crossover, and mutation. Memetic algorithms (MAs) [6] are population-based heuristic search, inspired by Dawkins' notion. A meme is defined as a unit of information that reproduces itself while people exchange ideas. A meme is usually modified by the person before passing it to the next generation. In memetic algorithm global search is made on entire population, like genetic algorithms. As described by Menon et al. [7], the simulated annealing (SA) is a simple and efficient heuristic. It controls slow convergence of local search. The effectual annealing technique avoids unnecessary exchange of assignments between cells and switches. The simulated annealing algorithm performs search for feasible solution on those neighborhood solutions neglected by local search algorithm. The acceptance probability allows the current assignment for further local search if the objective cost is better than the current best solution.

6 Conclusion

In this paper overview of assignment problem in three different wireless networks has been discussed. The overview is based on various issues like (i) network architecture (ii) objective function (iii) constraints (iv) algorithms (v) optimization strategy. Algorithms like memetic algorithm, genetic algorithm, tabu search algorithm, simulated annealing have been discussed.

References

1. Merchant, A., Sengupta, B.: Assignment of cells to switches in PCS networks. *IEEE/ACM Transaction on Networking* 3(5), 521–526 (1995)
2. Houeto, F., Pierre, S.: Assigning cells to switches in cellular mobile networks using taboo search. *IEEE Transactions on Systems, Man and Cybernetic - Part B: Cybernetic* 32(3), 351–356 (2002)
3. Quintero, Pierre, S.: On the Design of Large-Scale UMTS Mobile Networks Using Hybrid Genetic Algorithms. *IEEE Transactions on Vehicular Technology* 57(4), 2498–2508 (2008)
4. Grillo, D., Sasaki, A., Skoog, R.A.: Personal Communications—Services, Architecture, and Performance Issues. *IEEE Journal on Selected Areas in Communications* 15(8), 1385–1389 (1997)
5. Chamberland, S., Pierre, S.: On the Design Problem of Cellular Wireless Networks. *Wireless Networks* 11, 489–496 (2005)
6. Quintero, Pierre, S.: A memetic algorithm for assigning cells to switches in cellular mobile networks. *IEEE Communications Letters* 6(11), 484–486 (2002)
7. Menon, S., Gupta, R.: Assigning cells to switches in cellular networks by incorporating a pricing mechanism into simulated annealing. *IEEE Transaction on Systems, Man and Cybernetics, Part B: Cybernetics* 34(1), 558–565 (2004)

Design of Domain Specific Language for Web Services QoS Constraints Definition

Monika Sikri

Cisco Systems India Pvt Ltd,
SEZ Unit, Cessna Business Park,
Bangalore Karnataka 560103
msikri@cisco.com

Abstract. Semantic Webservices (SWS) has raised interest in mechanisms for Ontological representation of Web Services. A number of mechanisms most notably WSMO and OWL-S are being developed to represent the same. An important area in description of Web Services is the QoS characterization and discovery which is the focus of research for this paper. A Domain Specific language is being proposed for definition of observable QoS characteristics and conditions. The syntax of this proposed language is being kept closer to WSML considering it the standard modeling language.

Keywords: Web Services, QoS, SOA, WSML.

1 Introduction

Semantic Web Services are slowly moving from research papers to actual deployment. Development in this domain is accelerated with the availability of a number of open source projects like WSMX [2] to support fast adoption. An area of interest in this domain is development of mechanism for QoS model representation. Two common ontologies exist in the form of WSMO [15] and OWL-S [11] for definition of SWS. This work focuses on defining a Domain Specific Language (DSL) for representation of QoS Parameters. Emphasis in this paper is primarily on DSL design which can be both used for definition of QoS capability and also specifying QoS constraints as required by Web Service client. A domain specific language is developed as an embedded language for Groovy and hence inherits some groovy constructs. The use of standard language allows for fast development and early adoption of DSL.

2 Domain Specific Language

The proposed language is created as an embedded DSL for Groovy. Groovy [1] is a dynamic language completely written on top of Java. It provides an easy learning curve for Java developer [16]. Its dynamic nature supports an easy development of embedded DSLs [5]. The use of standard language also ensures that conditions on QoS

constraints can be easily expressed without the need of an elaborate grammar as in WSML Rule Syntax [3]. The user can use standard Java or Groovy syntax interchangeably.

As a Domain language for definition of QoS characteristics and constraints in Web Services, the syntax to specify basic QoS characteristics and constraints needs to be first specified and determined. At the same time, care is being taken to keep it closer to ontological hierarchy of WSML structure.

3 Related Work

Specifying non functional properties in Web Services has been approached in different ways .Extending WSDL to specify QoS is one of the approaches as has been proposed by authors in [4] .This would require definition of concrete QoS properties. Model-Driven specification of QoS parameters is another approach as also being specified in [13]. UML system model is used to capture QoS in Web service functionality .This model driven approach is based on UML and is specific to it.QoS based ontology model have also been proposed by [20].It uses OWL-S to define ontology whereas the study of this paper focuses on WSML. Specifying QoS in UDDI and using broker based architecture [14] to access service is another approach. This approach is limited by UDDI capabilities. QoS ontologies to semantically specify QoS information in Web Services is another approach towards semantic Web Services [9].There are existing ontologies and language which addresses different aspects of QoS. A comparative study has been conducted of various discovery mechanisms [12] and ontologies [18] to understand the pros and cons of various approaches. Frameworks have also been proposed for QoS ontology based Web Service selection [10].Authors in [7] have proposed QoS specification using OWL to cater to clients based on contribution level. QoS ontology based on WSMO [19] as part of its definition specifies WSML which is a RDF/XML mechanism for definition of Semantic Webservices. This is a standard ontological model for definition of QoS model .The work conducted [17] has focused into modeling QoS characteristics using WSMO though the disadvantage is it has further extended WSMO. The WSML framework provides a proposed mechanism to specify QoS parameters both functional and non-functional. Qos is not central to WSML and its definition using nonfunctional properties and rule syntax is quite complex. The work proposed in this paper is intended to provide a DSL for better representation of QoS parameters and is in line with WSML terminology .The DSL is also being implemented in Groovy. It is based on constraint programming [6] ideas and provide a verbose syntactical mechanism for defining QoS constraints.

4 Syntax of Proposed DSL

Listing 1.1 provides a sample of the Domain Specific Language. Various important parts of DSL syntax and elements of QoS characteristics are then defined.

```

1 QosLanguage{
2
3   QosConcept {
4       name "Availability "
5       description "Describes the availability capability of Web Service"
6       unit Integer
7       lowerBound "0"
8       upperBound "100"
9       }
10
11  QosWebService{
12      name "webserviceQos"
13      capability "StockWS"
14      qosinterface{
15          Availability {
16              description " availability description"
17              value "8"
18          }
19      }
20  }
21
22      QosAxiom{
23          name "myQosCondition"
24          concept "Availability "
25          description "Describe the mechanism of comparing the availability "
26          conditionf{ $value →
27              return ( value > 50)
28          }
29      }
30
31      QosGoal{
32          description "Goals we wanted to achieve"
33          name "goalWeb"
34          axiomRequired [ "myQosCondition" ]
35      }
36  }

```

Listing 1.1. Domain Specific Language for Webservices Qos Parameters

DSL syntax is an important aspect of this design since it would play an important role in its adoption. The salient features that this language exhibits are the separation of multiple statements using newline and use of equal sign between property name and its value to separate them. Various keywords are followed by curly brackets.

4.1 Key Elements of DSL

QosLanguage: This keyword forms the external most qualifier which is used to assert the presence of QoS constraint language. Line 1 in Listings 1.1 defines this tag.

QosConcept: QosConcept is used to define QoS parameter/characteristics. It is an important element since it defines the parameters on which QoS selection is based.

The concept is used to define both functional and non functional parameters like availability and security. Lines 3 to 9 in Listings 1.1 define the QoSConcept. Various parameters that are defined as part of QoSConcept are as follows

description	A verbal comment on the content of this tag.
Name	Unique name to identify this particular QoSConcept.
dataType	Defines the datatype used to represent this QoSConcept. This Datatype should be a well defined Groovy Datatype (or Java datatype).
lowerBound	Defines the lower expectable bound on the value of this QoS parameter. The value is defined as a String. And it is converted into dataType unit.
upperBound	defined the upper threshold value on this parameter

QosWebService: This tag is used to define expected/observed QoS values for Web Service provider and associates value for each QoS concept. This parameter is used for binding QoS language to actual webservices provider. Elaborate mechanism need to be developed for mapping the data to WSDL or UDDI based systems. Lines 11 to 20 in Listings 1.1 define the QosWebService. Its various parameters are as follows

description	A verbal comment on the content of this tag.
name	unique name to identify this particular WebService.
capability	Use to associate webservice interface with this tag
qosinterface	Embedded tag inside webservice which is used to specify values for actual QOS concepts. A nested tag, which contains one tag for each of QoSconcept. The rough syntax for Qosinterface is provided in listing 1.2.

QoSAXiom: This tag forms the center piece of QoS Language. It is used to define actual conditions on QoS parameters. Condition on QoSAXiom is defined as closure. It contains a groovy syntax for defining a condition statement on passed value of the concept say "Availability". The term axiom is chosen from WSMML ontology to provide for easy interoperability. Lines 22 to 29 in Listings 1.1 define the QoSAXiom.

The parameters for QoSAXiom are as follows

description	A verbal comment on the content of this tag.
Name	Specifies a unique name to identify this axiom.
concept	name of the Qosconcept on which conditions are defined.
condition	Closure structure to define actual condition for verification.

Listing 1.2 defines syntax for qosinterface

```

1 QoSWebService{
2   name....
3   qosinterface {
4     <<qosconceptname>>{
5       description = " "
6       value = " "
7     }
8   }
9 }
```

Condition tag contains a closure which must return a boolean value true or false. This implies, that the last statement in this Groovy block must evaluate to true or false. Since this is a groovy block, user can write a valid groovy code to define the condition. One explicit value, which is current QoS concept value of a known Web Service, is passed to this closure.

QosGoal: It is used to define requirements of Web Service consumer. It defines conditions that web service consumer requires to be satisfied. The actual set of conditions on individual QoS parameters is defined as part of QoS Axioms. Lines 31 to 35 in Listings 1.1 define QosGoal. The goals are used to specify set of axioms that need to be satisfied as part of this webservice consumer. The syntax of QosGoal is as follows

QoSGoal is also a binding point for Qos Language to the actual webservice client (or consumer). The success evaluation of goal ensures the selection of a QoS webservice endpoint.

description	A verbal comment on the content of this tag.
name	unique name to identify this goal this axiom.
axiomRequired	A groovy string array containing names of QoS Axiom (conditions) that needs to be satisfied.

5 Conclusions

The intent of this work is to provide an easy mechanism for webservice QoS parameters definition. QoS support for Web Services adds a competitive behavior to Web Service and client interaction. It adds a lot of value in essential for an enterprise environment [8] in usage of Web Services. This intends to provide at the same time some level of inter op with WSML standard. Besides it also helps service consumer to make the right choice based on differential QoS provided among similar Web Services. The resultant embedded DSL can be easily be used to define QoS conditions in Semantic Web Services environment.

Acknowledgments. I would like to thank my organization Cisco Systems Inc., and all anonymous referees for the valuable contribution and suggestions.

References

1. Groovy: An agile dynamic language for the java platform, <http://groovy.codehaus.org/>
2. Web service execution environment, <http://www.wsmx.org/>
3. de Bruijn, J., Lausen, H., Polleres, A., Fensel, D.: The wsml rule languages for the semantic web, <http://www.w3.org/2004/12/rules-ws/paper/128/>
4. Dai, C., Wang, Z.: A flexible extension of WSDL to describe non-functional attributes. In: 2010 2nd International Conference on e-Business and Information System Security (EBISS), May 2010, pp. 1–4 (2010)
5. Devijver, S.: A groovy DSL from scratch in 2 hours, <http://groovy.dzone.com/news/groovy-dsl-scratch-2-hours>
6. Garcia, J.M., Ruiz, D., Ruiz-Cortes, A., Martin-Diaz, O., Resinas, M.: An Hybrid, QoS-Aware Discovery of Semantic Web Services Using Constraint Programming. In: Krämer, B.J., Lin, K.-J., Narasimhan, P. (eds.) ICSOC 2007. LNCS, vol. 4749, pp. 69–80. Springer, Heidelberg (2007), http://dx.doi.org/10.1007/978-3-540-74974-5_6
7. Ha, Y., Park, H.S.: Qos Based on Client information for Semantic Web Service. In: Advanced Software Engineering and Its Applications, pp. 237–240 (2008)
8. Kritikos, K., Plexousakis, D.: Requirements for Qos-based Web service Description and Discovery. IEEE Transactions on Services Computing 2, 320–337 (2009)
9. Li, H., Du, X., Tian, X.: Towards Semantic Web Services Discovery with QoS Support Using Specific Ontologies. In: International Conference on Semantics, Knowledge and Grid, pp. 358–361 (2007)
10. Li-li, Q., Yan, C.: QoS Ontology based efficient web services selection. In: International Conference on Management Science and Engineering, pp. 45–50 (2009)
11. Martin, D., Burstein, M., Hobbs, E., Lassila, O., Mcdermott, D., Mcilraith, S., Narayanan, S., Parsia, B., Payne, T., Sirin, E., Srinivasan, N., Sycara, K.: OWL-S: Semantic Markup for Web Services (November 2004), <http://www.w3.org/Submission/OWL-S/>
12. Nair, M.K., Gopalakrishna, V.: Article:Look Before You Leap: A Survey of Web Service Discovery. International Journal of Computer Applications 7, 22–30 (2010); published By Foundation of Computer Science
13. Ortiz, G., Bordbar, B.: Model-Driven Quality of Service for Web Services: An Aspect-Oriented approach. In: IEEE International Conference Web Services (ICWS 2008), pp. 751–748 (2008)
14. Rajendran, T., Balasubramanie, P.: An Optimal Agent-Based Architecture for Dynamic Web Service Discovery with QoS. In: International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–7 (2010)
15. Roman, D., Keller, U., Lausen, H., de Bruijn, J., Lara, R., Stollberg, M., Polleres, A., Feier, C., Bussler, C., Fensel, D.: Web Service Modeling Ontology. Appl. Ontol. 1, 77–106 (2005), <http://portal.acm.org/citation.cfm?id=1412350.1412357>
16. Subramaniam, V.: Programming Groovy (2008)
17. Toma, I., Foxvog, D., Jaeger, M.C.: Modeling QoS Characteristics in WSMO. In: Proceedings of the 1st workshop on Middleware for Service Oriented Computing (MW4SOC 2006), pp. 42–47. ACM, New York (2006), <http://doi.acm.org/10.1145/1169091.1169098>
18. Tran, V.X., Tsuji, H.: A Survey and Analysis on Semantics in QoS for Web Services. In: International Conference on Advanced Information Networking and Applications, pp. 379–385 (2009)
19. Wang, X., Vitvar, T., Kerrigan, M., Toma, I.: A QoS-Aware Selection Model for Semantic Web Services, pp. 390–401 (2006), http://dx.doi.org/10.1007/11948148_32
20. Zhang, H., Gao, W.: A Research on QoS-based Ontology Model for Web Services Discovery. In: International Workshop on Knowledge Discovery and Data Mining, pp. 786–789 (2009)

Modified Auxiliary Channel Diffie Hellman Encrypted Key Exchange Authentication Protocol

Nitya Ramachandran¹ and P. Yogesh²

¹ Department of Computer Science & Engineering
Anna University, Chennai
nityaram86@gmail.com

² Department of Information Science & Technology
Anna University, Chennai
yogesh@annauniv.edu

Abstract. In the wireless network, security is a major concern. Bluetooth network is also highly vulnerable to attacks. This paper proposes a protocol called, Modified Auxiliary Channel Diffie Hellman Encrypted Key Exchange authentication Protocol (MACDHEKE), which authenticates the two previously unknown Bluetooth enabled devices. The user inputs a low security, low entropy pins and converts it into high security, high entropy shared key. It also authenticates the public key as being sent by the claimed sources or devices. It also analyzes how secure this protocol is against man in the middle attack as well as passive eavesdropping.

Keywords: Auxiliary channel, pin, shared key, elliptic curve, Diffie Hellman key exchange protocol.

1 Introduction

A wireless network is a network where the devices are connected via radio signals rather than wires. One of the important wireless networks is Bluetooth also called Wireless Personal Area Network (WPANs). Bluetooth network connects devices in close range say 1m - 100m approximately. It uses 2.4 GHz short range radio frequency unlicensed Industrial Scientific Medical (ISM) band.

Since Bluetooth is a wireless technology, it is highly vulnerable to attacks. One major attack is man in the middle attack, abbreviated as MITM attack. In this attack, a third device gets in between two devices and establishes a connection impersonating as a valid user. He can then intercept all messages send between the devices. This is a form of active eavesdropping where the attacker can modify the messages. Another form of attack is passive eavesdropping where the attacker only listens to the messages without modifying them. These attacks occur when two Bluetooth devices try to pair with each other.

Rest of the paper is organized as follows. The second section deals with related work. The MACDHEKE protocol, MITM attacks are given in the third and fourth section respectively.

2 Related Work

The different combinations of connectability and discoverability capabilities of a Bluetooth device can be divided into three categories, or security levels namely silent, private and public [5]. Since Bluetooth is a wireless communication system, it is highly vulnerable to attacks such as jamming or intercepting and fabrication of information being passed to the piconet devices [1]. Therefore, Bluetooth security is a major concern now days.

There is a lot of research going on in Bluetooth pairing. Even with longer alphanumeric PIN's, full protection against active eavesdropping cannot be achieved. It is found that MITM attacks on Bluetooth versions up to 2.0+EDR can be still performed [6]. Bluetooth versions 2.1+EDR and 3.0+HS (High Speed) add a new specification for the pairing procedure, namely SSP [7, 10]. The new Simple Pairing procedure in Bluetooth version 2.1 [8] tries to remedy a number of the weaknesses of previous versions. Paper [11] suggests a method for authentication and key agreement achieved through short string comparison and distance bounds.

The above version is enhanced by a key agreement procedure during pairing using an authenticated challenge-response procedure over an auxiliary channel [9]. Compared with [9] where only unidirectional auxiliary channel is required, [12] specifies verification of the symmetric key using commitment value, pin and nonce and thus require bi directional auxiliary channel.

3 Design of MACDHEKE Authentication Protocol

MACDHEKE is designed with aim of preventing man in the middle attack during pairing between two devices.

The proposed protocol is as shown in figure 1. First, public keys, PKa and PKb of device A and B are exchanged. Then device A and device B generates its own temporary elliptic curve public private key pair. That is for device A, the pair generated is (TPKa, TSKa) and for device B, pair is (TPKb, TSKb). The PIN is then sent over auxiliary channel. Device A then encrypts TPKa with pin displayed in device B and inputted in device A and send to device B. B then calculates temporary DHKey, Kba = P192 (TSKb, TPKa). That is, device B uses its temporary private key and device A's temporary public key. It then calculates commitment value, Cb = f' (PKa, PKb, Kba, TPKa), Where, f' (W, X, Y, Z) = HMAC-SHA256_Y(W||X||Z). This commitment value is again encrypted using Vernam cipher. Temporary DHKey, Kba can be used to encrypt the commitment value. But since there is possibility of the intruder to compute this DHKey, instead using Kba directly, simple secret cryptographic function is used by two devices to compute a symmetric shared key using DHKey, Kba as function argument. This encrypted commitment value is send across to device A along with its temporary public key encrypted with pin A.

Device A then calculates its temporary DHKey, Kab = P192 (TSKa, TPKb) and using this calculates the commitment of device B. This is verified against the commitment value received from B after decryption. If the device B get verified, it means, A is sure that it's to device B it is communicating. Device A calculates its commitment value Ca = f'(PKb,PKa,Kab,TPKb), encrypts it using symmetric shared key and

sends it to B. B decrypts it, calculates commitment value of device A and verifies. If verification succeeds, then B is sure that it is communicating with the device A. Now original DHKey is calculated by device A using P192 (SKa, PKb) and by device B using P192 (SKb, PKa). After the last step, both devices are assured that they have exchanged public key values and that a DH key is established between them ($DHKey_{ab} == DHKey_{ba}$). The temporary public and private keys are discarded after the final DH key is established. It is important that TPK_a is not revealed as it would allow an attacker to brute force the PIN.

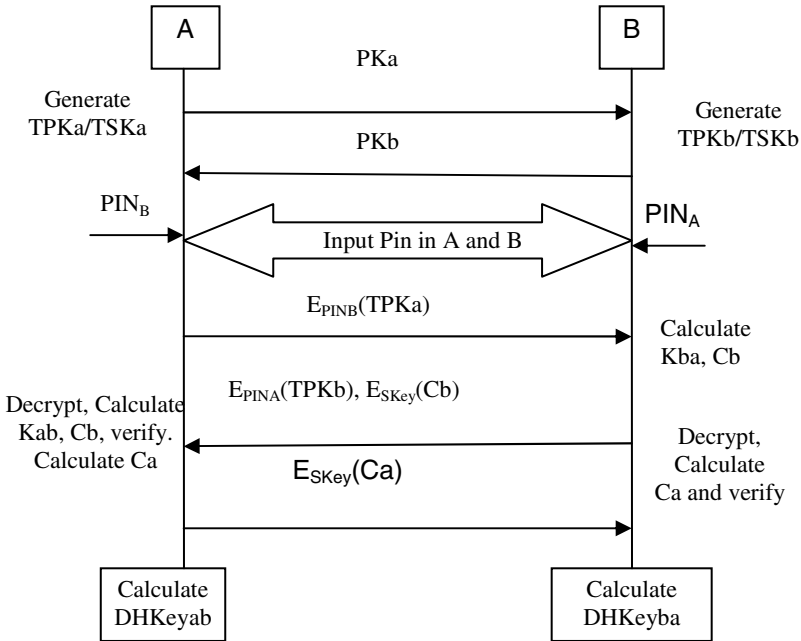


Fig. 1. MACDHEKE protocol

Analysis of the this cryptographic protocol is done using ProVerif which is a software tool for automated reasoning about the security properties found in cryptographic protocols. Secrecy of a key k can be modeled as a query for the attacker’s knowledge k . If ProVerif’s algorithm terminates and the set of attacker knowledge does not include k , then k has been proved to be secret [8].

MACDHEKE protocol can be written in pi calculus. The device A is specified as initiator and device as responder. The query is checking whether pin A, pin B and symmetric key is present in the knowledge base of the attacker or not after execution of the protocol. If the keys are part of knowledge base of attacker, it means pins and key is no longer secret. In this protocol, that attacker obtaining the values of pin A, pin B and symmetric key is false.

4 MITM Attack Scenario

Figure 2 shows the attack scenario. The attacker first injects its own public key into the network. Device A and B as well as the attacker will now generate temporary public private key pair. Attacker then receives the encrypted temporary public key, $E_{pinB}(TPKa)$ from device A which is encrypted with $pinB$ send over the auxiliary channel.

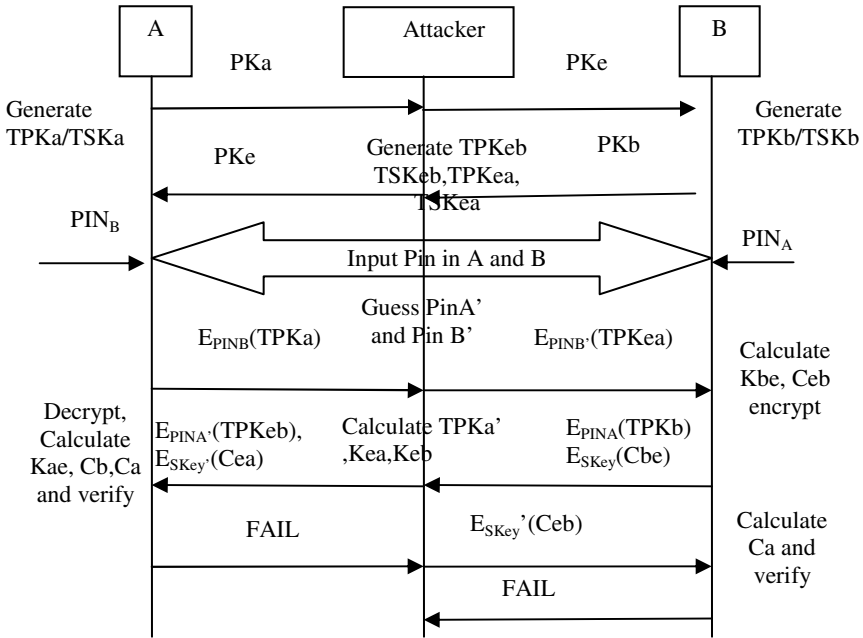


Fig. 2. Attack Scenario

Attacker will not know this pin. So it guesses a pin, $pinB'$ and encrypts its temporary public key with $pinB'$ and send it over to device B. Attacker also decrypts $E_{pinB}(TPKa)$ to obtain $TPKa'$. Device B decrypts it and calculates temporary DHKey, Kbe as $P192 (TSKb, D_{pinB}(E_{pinB'}(TPKe)))$. Then B calculates commitment value, $Ceb = f(Pke, Pkb, Kbe, D_{pinB}(E_{pinB'}(TPKe)))$. This value is encrypted as in previous case. Attacker obtains this encrypted value and temporary public key of B encrypted with pin A. The attacker again makes a guess on the pin of A as well as the key and encrypts Cea calculated and also encrypts his temporary public key with guessed pin of A. At this point, attacker is mounting attack on device A. He then sends $E_{pinA'}(TPKbe)$ and $E_{SKey'}(Cea)$ to device A.

Device A decrypts the received message to get $TPKbe'$ and Cea' . $TPKbe'$ will be equal to $TPKbe$ only if $pinA = pinA'$ and Cea will be equal to Cea' only if $Skey = Skey'$. Device A calculates DHKey, Kae and the commitment value, Cae . Here Cae will be equal to Cea only if $Kae = Kea$. It then verifies with received Cea' . This will

fail since Kea is not equal to Kae because, $TPKbe'$ is not equal to $TPKbe$. The device A aborts the connection.

Attacker then launches attack on device B by calculating the commitment value, Ceb . It then encrypts Ceb with key' and sends it to device B. Device B then calculates Cbe . This verification also fails since attacker does not know $pinB$ and key because of which Kbe and Keb will not equal and therefore Cbe and Ceb also will not be equal. Thus attacker fails in establishing a connection with either device A or B. So MITM attack is prevented since both A or B detects the attack.

5 Conclusion

Modified Auxiliary Channel Diffie Hellman Encrypted Key Exchange authentication protocol is proposed to establish a connection between two Bluetooth enabled devices. It also authenticates the two previously unknown Bluetooth enabled devices. The scheme protects against MITM attack and passive eavesdropping.

References

1. André, N., Klingsheim: J2ME Bluetooth Programming. Department of Informatics University of Bergen (June 2004)
2. Anoop, M.S.: Elliptic Curve Cryptography. An Implementation Guide
3. Hwang, M.-S.: A Secure Protocol for Bluetooth Piconets Using Elliptic Curve Cryptography. *Telecommunication Systems* 29(3), 165–180 (2006)
4. Chris, K., Cockrum: Implementation of an Elliptic Curve Cryptosystem on an 8-bit Microcontroller. Springer, Heidelberg (2009)
5. Haataja, K., Toivanen, P.: Two practical man-in-the-middle attacks on Bluetooth secure simple pairing and countermeasures. *IEEE Transactions on Wireless Communications* 9(1), 384–392 (2010)
6. Haataja, K.: Security threats and countermeasures in Bluetooth-enabled Systems. Ph.D. dissertation, University of Kuopio, Department of Computer Science (February 2009)
7. Bluetooth SIG, Bluetooth Specifications 1.0–3.0+HS, <http://www.bluetooth.com/Bluetooth/Technology/Building/Specifications>
8. Nilsson, D.K., Larson, U.E., Jonsson, E.: Auxiliary Channel Diffe-Hellman Encrypted Key-Exchange Authentication. In: ACM QShine 2008, Hong Kong, China (July 2008)
9. Nilsson, D.K., Larson, U.E., Jonsson, E.: Unidirectional Auxiliary Channel Challenge-Response Authentication. In: Proceedings of the 7th Annual IEEE Wireless Telecommunications Symposium (WTS), Pomona, CA, USA (2008)
10. Bluetooth Special Interest Group. Simple pairing white paper (2006)
11. Cagalj, M., Capkun, S., Hubaux, J.-P.: Key Agreement in Peer-to-Peer Wireless Networks. In: Proceedings of the IEEE Special Issue on Cryptography and Security, pp. 467–478 (2006)
12. Jakobsson, M.: Lecture Notes in 1400/I590: Issues in Security and Privacy, <http://www.informatics.indiana.edu/markus/i400/>

Bilateral Partitioning Based Character Recognition for Vehicle License Plate

Siddhartha Choubey, G.R. Sinha, and Abha Choubey

Shri Shankracharya College Of Engg and Technology, Junwani Bhilai
sidd25876@gmail.com, ganeshsinha2003@gmail.com,
abha.is.shukla@gmail.com

Abstract. This paper presents a new methodology for the image segmentation and character recognition from standard Indian License number plates. This method first, gets input of the segmented characters that is partitioned by bilateral partitioning method, in which we eliminate similar part from the character and match it by judging template and return identified character. This partitioning may be performed horizontally or vertically. The characters are to be partitioned horizontally or vertically depend on their subgroup, and before this subgroup the characters are grouped on the basis of the number of holes in it. The subgroup is made on the basis of some similar features like |, /, \, _ , (, - , etc. Suppose we have alphabet T and I here similar portion is I than both will go to same subgroup and we partition it horizontally. This method eliminates the problem of confusion between similar looking elements like C, G and T, I, l, J etc by exploiting the small but significant differences among them.

Keywords: Bilateral, Euler, Partitioning, Segmentation, Pixel Dentsity Distribution.

1 Introduction

At present, a vital part in transportation is played by vehicles. In recent years, the usage of vehicles has been increasing on account of population growth and human needs. Hence control of vehicles is becoming a huge problem that is hard to solve [1]. An image processing technology called License Plate Recognition (LPR) that is used to identify vehicles by their license plates is a kind of automatic vehicle identification [2]. Image acquisition, license plate detection, character segmentation and character recognition are four major phases of License Plate Recognition [3-4]. Several areas including traffic volume control, unsupervised park monitoring, traffic law enforcement and auto toll collections on highways extensively use license plate recognition applications [5-6]. In this paper we have introduced a new algorithm for character extraction and recognition based on based on fuzzy logic and analyzing some pattern recognition based works.

2 Related Works

Mi-Ae Ko et al. [7] proposed a algorithm for recognizing optimal solution and good optimization effect in fast speed, by using filled function method and BP neural network. Experimental results show that the performance of the proposed method is simple and robust. Fei Lu Mei Xie et al. [8] focussed on extra dot appear in the license plate, this extra “dot” can completely affect the “shape descriptors”, with an XOR-based kernel, is used. The XOR-based kernel can improve the performance of the system drastically. Jian-xia Wang et al. [9] emphasised on the time complexity for template matching procedure. In order to enhance recognition speed and recognition rate, an improved template matching method is presented.

3 System Overview

We proposed a novel method to recognize vehicle number plate elements which are captured by any camera. The system has two stages viz. pre-processing and post-processing. Pre-processing begins with taking input image then vehicle plate is extracted. Characters of the number plate are segmented. Post-processing stage includes feature extraction and storing then in the form of templates for matching purpose followed by character recognition process.

4 Methodology

The first step is vehicle number plate detection which is followed by character and number segmentation, feature extraction & recognition of the extracted feature and character recognition.

4.1 Preprocessing (Vehicle Number Plate Extraction)

In preprocessing we capture the image (Fig. 1) and convert it to gray image shown in Fig. 2.

4.1.1 Vehicle Number Plate Elements Segmentation

Various preprocessing steps used are: (a) Converting image to gray scale image, (b) Removing all objects smaller than 100 pixels. Fig. 3 shows the result of canny edge filter. Fig. 4(a) shows vehicle plate extraction and converting into black and white gives Fig. 4(b). Finally, Fig. 4(c) shows the result with noise removed.



Fig. 1. Captured Image



Fig. 2. Gray scale image



Fig. 3. Result of canny filter

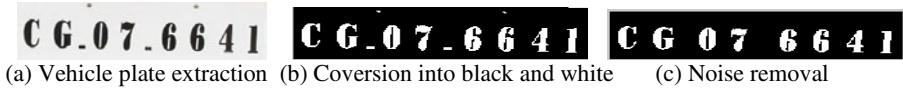


Fig. 4. Results of extraction and noise removal

4.1.2 Removing Extra Area from Left and Right by Detection of First and Last Columns for Each Character and Number

The first step in segmentation process is to cutoff the background from each character and number from the license plate. Vertical scanning is used to detect first and last columns for each character and number is conducted before horizontal scanning as explained in Algorithm.

4.1.3 Remove Extra Area from Upper and Lower of Each Component

Horizontal scanning will be done to detect the first and last rows from the result of the previous step. The results of this operation will be images that contain only the characters or numbers without any extra area. This result will help to extract image features easily. Fig. 5 shows the result of vertical segmentation and Fig. 6 depicts the result horizontal segmentation.



Fig. 5. Vertical Segmentation



Fig. 6. Horizontal Segmentation

4.2 Post Processing: Feature Extraction

The Features extraction step contains the following steps:

4.2.1 Algorithm (Post Processing) Detection of (Character and Numbers): Grouping Process

- STEP 1: Calculate the number of holes, h = number of holes in character
- STEP 2: Calculate Euler number based on number of holes: $e=c-h$, where, $c=1 //c$ is the number of connected components (always equal to one)
- STEP 3: Calculate Euler group based on Euler number: If $e = 1$, then $eg=1$; else if $e=0$ then $eg=2$; else if $e=-1$ then $eg=3$; where eg is the Euler group.

Sub Grouping Process

- STEP 4: For each Euler group match with subgroup deciding template for subgroup deciding templates find their indexes.
- STEP 5: Maximum match determines the subgroup. Return the respective subgroup.
- STEP 6: Now once subgroup is decided compare with differentiating template: For the subgroup templates find their indexes.
- STEP 7: Maximum match determines the element: For maximum match return the respective element (character and number).

It is clear that this grouping process will increase system accuracy of recognizing characters and numbers and it will also reduce the time of recognition process.

4.2.2 Character and Number Recognition

Here, we use a new methodology in which the characters are separated into subgroups on the basis of their similarity or different physical appearance. Euler group 1 is divided into the following subgroups: Subgroup C: consists of C and G. Subgroup E: Consists of E and F, Subgroup I: Consists of I, T, J & 1 and Subgroup S: Consists of 5 and S. Euler group 2 is divided into the following subgroups: Subgroup O: consists of O and Q; Subgroup P : consists of P and R. All rest elements are not categorized into any subgroup and have a separate template for each of them. Each group has a template; the template that matches with the input the most decides the group of the character. The subgroup is decided on the basis of indexes calculated by formulae:

Index of a particular template = (Number of pixels that matched with that template) / (Total number of pixels in that template). The subgroup whose template has highest index with respect to the character is chosen as the subgroup.

4.3 Recognition of Elements belonging to Same Subgroup

After sub-grouping the physical characteristics of each element that differentiate it from other elements are used to identify the element. In Subgroup E: E and F are members of this subgroup. The template shown in Fig. 7 is used to check, if the input belongs to this group then the template shown in Fig. 8 is used to detect if the input is E or F as this template is a part of E but not F.



Fig. 7. Template to judge group



Fig. 8. Template to judge whether E or F

5 Result and Conclusion

In this paper, we have proposed a new method for character recognition which very effectively eliminates the problem of confusion between characters of same type like C & G which the previous algorithms were not able to do with high precision. Initial experiments have shown success in differentiating between similar looking characters. In spite of being very effective on standard number plate suffers with the disadvantage that it is not useful on those plate which use non-standard characters to represent the vehicle's registration number. Fig. 9 shows results for some characters.






Subjected Character					
Results	M	H	0	2	G

Fig. 9. Results for some characters

References

- [1] Shan, B.: License Plate Character Segmentation and Recognition Based on RBF Neural Network. In: 2nd International Workshop on Education Technology and Computer Science, vol. 2, pp. 86–89 (2010)
- [2] Broumandnia, Fathy: Application of pattern recognition for Farsi license plate recognition. ICGST International Journal on Graphics, Vision and Image Processing 5(2), 25–31 (2005)
- [3] Cancer, H., Gecim, H.S., Alkar, A.Z.: Efficient Embedded Neural – Network Based License Plate Recognition System. International Journal of Information and Security 57(5), 2675–2683 (2008)
- [4] Feng, J., Li, Y., Chen, M.: The Research of vehicle License Plate Character Recognition Method Based on Artificial Neural Network. In: 2nd International Asia Conference on Informatics In Control, Automation And Robotics, vol. 8(1), pp. 317–320 (2010)
- [5] Tsai, I.-C., Wu, J.-C., Hsieh, J.-W., Chen, Y.-S.: Recognition of Vehicle License Plates from a Video Sequence. IAENG International Journal of Computer Science 36(1), 26–33 (2009)
- [6] Deb, K., Chae, H.-U., Jo, K.-H.: Vehicle License Plate Detection Method Based on Sliding Concentric Windows and Histogram. Journal of Computers 4(8), 771–777 (2009)
- [7] Zhang, Y., Xu, Y., Ding, G.: License plate character recognition based on fill function Method Training BP Neural Network. In: Control and Decision Conference (CCDC 2008), Chinese, pp. 3886–3891 (2008)
- [8] Lu, F., Xie, M.: An Efficient Method of License Plate Location in Complex Scene. In: Second International Conference on Computer Modeling and Simulation, ICCMS 2010, Sanya, Hainan, January 22-24, vol. 2 (2010)
- [9] Wang, J.-X., Zhou, W.-Z., Xue, J.-F., Liu, X.-X.: The research and realization of vehicle license plate character segmentation and recognition technology. In: Proceedings of the 2010 International Conference on Wavelet Analysis and Pattern Recognition, Qingdao, July 11-14 (2010)

Strategies for Parallelizing KMeans Data Clustering Algorithm

S. Mohanavalli, S.M. Jaisakthi, and C. Aravindan

SSN College of Engineering, Chennai, India
{mohanas, jaisakthism, aravindanc}@ssn.edu.in

Abstract. Data Clustering is a descriptive data mining task of finding groups of objects such that the objects in a group will be similar (or related) to one another and different from (or unrelated to) the objects in other groups [5]. The motivation behind this research paper is to explore KMeans partitioning algorithm in the currently available parallel architecture using parallel programming models. Parallel KMeans algorithms have been implemented for a shared memory model using OpenMP programming and distributed memory model using MPI programming. A hybrid version of OpenMP in MPI programming also has been experimented. The performance of the parallel algorithms were analysed to compare the speedup obtained and to study the Amdhals effect. The computational time of hybrid method was reduced by 50% compared to MPI and was also more efficient with balanced load.

Keywords: Data Mining, Clustering, Parallel KMeans, OpenMP, MPI.

1 Introduction

Data Mining is an art of extracting interesting (non-trivial, implicit, previously unknown and potentially useful) information or patterns from large data sets [5]. Clustering is an unsupervised data mining technique applied on a given data set to partition it into groups of similar data objects. The KMeans algorithm is one of the simplest & popularly practised partitioning algorithm based on similarity distance [4]. Sequential design of such a computationally intensive algorithm may not be very efficient and hence can be parallelized to improve performance. The success of such straight forward parallelization of a sequential algorithm relies on a good data organization and decomposition strategy [2]. Many researchers have worked in the area of parallel data clustering. Zhang et al. have proposed an enhanced parallel algorithm using Master/Slave model and dynamic load balancing technique [12]. Rao et al. have explored the potential of Multi-Core hardware under OpenMP API and POSIX threads for both static and dynamic threads [11]. Chin et al. have addressed the problem of higher communication overhead, by adopting hybrid MPI and OpenMP programming model [1]. Zhou et. al have tried a distributed KMeans algorithm to improve the local clustering and decrease network load [13]. Jin et al. have focused in reducing the number of iterations performed to cluster the data by carefully choosing the initial centroids by sampling [9]. Datta et al. have proposed two solutions to

address the distributed clustering problem in a P2P network using local synchronization and uniform sampling of peers [3]. Most of the above research works aim to speed up the convergence and increase the efficiency of the algorithm. In this research paper the parallel approaches to perform KMeans clustering has been experimented and observed that the hybrid method outperforms the other methods.

2 Proposed System Design

The sequential KMeans algorithm design has an iterative structure where the centroid updations are performed in the outer loop and the cluster labels are assigned to each data point in the inner loop. The latter task was performed parallelly using OpenMP fork-join model. In the MPI model, data points were distributed evenly among the processors, data labeling and updation of new centroids were done parallelly by means of message passing between the MPI processes. In the hybrid model a combination of MPI with OpenMP was tried where the data points were evenly distributed among the MPI processes and the task of data labeling for each local data within the process was done using OpenMP directives. The performance of the algorithms were compared in terms of speedup and efficiency for varying problem sizes and the hybrid method was observed as the best parallel algorithm for clustering the input data set.

3 Experimental Results and Performance Analysis

A parallel algorithm is analysed for its cost effectiveness in terms of speedup and efficiency (processor utilization). A cost optimal parallel algorithm has speedup p (number of processors) and efficiency 1 in an ideal situation [10].

Table 1. Observed Execution Time for Parallel K-Means using MPI and Hybrid

Clustering Time in seconds										
k	5		7		9		11		13	
N	MPI	Hybrid	MPI	Hybrid	MPI	Hybrid	MPI	Hybrid	MPI	Hybrid
2	17.47	9.01	17.9	9.22	18.39	9.58	18.78	9.61	19.2	9.83
3	12.05	6.49	12.29	6.59	12.57	6.63	12.89	6.76	13.17	6.93
4	9.31	5.11	9.49	5.14	9.73	5.3	9.86	5.65	10.15	5.67
5	7.61	7.73	7.92	8.02	8.21	7.17	8.46	7.36	8.34	7.74
6	6.69	6.71	6.77	6.88	7.1	7.09	7.01	7.17	7.42	7.35
7	6.42	5.89	5.84	5.98	6.08	6.09	6.16	6.22	6.26	6.4
8	7.34	5.33	5.28	5.32	5.42	5.47	5.69	5.68	5.76	5.68
9	6.66	6.26	6.48	6.41	6.53	6.49	7.59	6.71	6.86	6.86
10	7.57	6.33	6.34	6.48	6.42	6.55	8.43	6.59	6.78	6.84

The sequential and all the versions of parallel KMeans algorithm were implemented and initially tested with benchmark data sets from the UCI repository [6] to ensure the correctness of the algorithms. Their parallel performance was tested with a forest data set with 581012 data points and 54 attributes which has ideally 7 clusters [6]. The algorithms were run for various values of k and n and the observed clustering

time is tabulated in Table 1 and Table 2. The speedup and efficiency obtained by running the algorithms is tabulated in Table 3 and Table 4. The algorithms show linear speedup and the Amdhals effect [10] was also studied to find the saturation point in execution time with increase in number of processors. Both the MPI as well as Hybrid MPI/OpenMP versions, show much improvement in speedup and efficiency compared to the other versions as summarized in Table 2.

Table 2. Performance Comparison for Sequential, OpenMP, MPI, Hybrid

Performance Analysis											
K	Best Clustering Time in sec				Speedup*				Efficiency*		
	SEQ	OMP	MPI	Hybrid	OMP	MPI	Hybrid	OMP	MPI	Hybrid	
5	18.8	16.9	6.42	5.11	1.11	2.93	3.68	0.56	0.41	0.46	
7	32.7	24.47	5.28	5.14	1.34	6.19	6.15	0.67	0.77	0.77	
9	37.9	30.23	5.42	5.3	1.25	6.99	7.15	0.63	0.87	0.9	
11	76.2	39.25	5.69	5.65	1.94	13.39	13.49	0.97	1.67	1.68	
13	51.73	26.62	5.76	5.67	1.94	8.98	9.12	0.97	1.12	1.14	
Avg	43.47	27.49	5.77	5.45	1.52	7.67	7.89	0.76	0.98	0.99	

* - N is the minimum no of process required as per Amdhal's effect

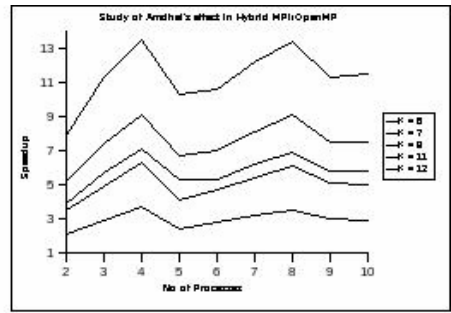
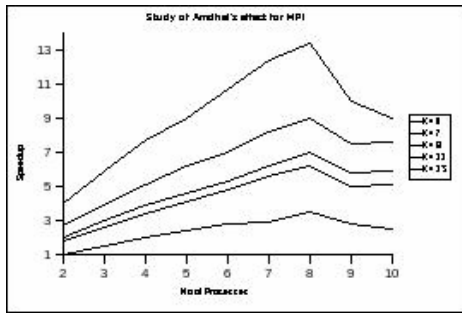


Fig. 1. Speedup for Parallel KMeans using MPI

Fig. 2. Speedup for Parallel KMeans using Hybrid MPI/OpenMP

The Amdhal's effect studied for various runs of the MPI and hybrid versions are depicted in Figure 1 & 2. The efficiency obtained using the hybrid method was almost double that of MPI algorithm, as the number of processors required to achieve best execution time using hybrid was half that required with MPI. The experiments were run in a cluster with 4 dual core systems. For varying values of k and n, the hybrid algorithm showed its best speedup with 4 nodes as it utilized both the cores to run the MPI code embedded with OpenMP threads. For higher n values the effect of load imbalance was seen as a drop in speedup and improved with load balance as shown in in Figure 2. For MPI the speedup gradually increased with increase in n and saturated at n = 8, beyond which the performance dropped as in Figure 1.

4 Conclusion and Future Work

In this research work, parallel k means clustering algorithms suitable for shared as well as distributed memory systems were implemented and studied for feasibility of parallelism and scalability of the algorithm. It was evidently shown that the hybrid version of the parallel algorithms was more efficient than the other approaches considering only the clustering time and is thus more suitable for distributed data mining. It is proposed to further improve the design to reduce the effect of I/O time by using appropriate file systems and parallel MPI-I/O routines.

References

1. Wu, C.-C., Lai, L.-F., Yang, C.-T., Chiu, P.-H.: Using hybrid MPI and OpenMP programming to optimize communications in parallel loop selfscheduling schemes for multi-core PC clusters. *Journal of Supercomputing* (2009), doi: 10.1007
2. Skillicorn, D.B.: Strategies for parallel data mining. *IEEE Concurrency* 7, 26–35 (1999)
3. Datta, S., Gianella, C.R., Kargupta, H.: Approximate distributed k-means clustering over a peer-to-peer network. *IEEE Transactions on Knowledge and Data Engineering* 21(10), 1372–1388 (2009)
4. Dhillon, I., Modha, D.: A Data Clustering algorithm on distributed memory multiprocessors. *IEEE Transactions on Knowledge and Data Engineering (KDD)* 1999, 47–56 (1999)
5. Han, J., Kamber, M.: *Data Mining: Concepts and Techniques*, 2nd edn. Morgan Kaufmann, San Francisco (2006)
6. <http://archive.ics.uci.edu/ml1> UC Irvine Machine Learning Repository
7. <http://www.openmp.org>
8. <http://www-unix.mcs.anl.gov/mpi>
9. Jin, R., Goswami, A., Agrawal, G.: Fast and Exact out of core and distributed KMeans clustering. *Knowledge and Information Systems*, 17–40 (2006)
10. Quinn, M.J.: *Parallel Programming in C with MPI and OpenMP*. Tata Mc- Graw Hill (2003)
11. Rao, S.N.T., Prasad, E.V., Venkatehwarulu, A.: Critical Performance Study of Memory Mapping on Multi-Core Processors: An Experiment with k-means Algorithm with Large Data Mining Data Sets . *International Journal of Computer Applications* (0975 8887) 1(9) (2010)
12. Zhang, X.Z., Mao, J., Ling Ou, L.: The Study of Parallel KMeans Algorithm. In: *Proceedings of the 6th WCAIAC*, pp. 5868–5871. IEEE, Los Alamitos (2006)
13. Zhou, J., Liu, Z.: Distributed Clustering Based on K-means and CPGA. In: *Proceedings of FSKD(2)*, pp. 444–447. IEEE, Los Alamitos (2008), doi:10.1109

A Comparative Study of Different Queuing Models Used in Network Routers for Congestion Avoidance

Narendran Rajagopalan and C. Mala

Department of Computer Science and Engineering,
National Institute of Technology, Tiruchirappalli,
Tamil Nadu, India-620015
narenraj1@gmail.com, mala@nitt.edu

Abstract. As the network traffic is increasing exponentially, the queuing model used in a network decides the degree of congestion that is possible in the network infrastructure. Hence using a suitable queuing model based upon the network infrastructure like buffer size and the type of data traffic flowing through the network will help in better utilization of system resources by minimizing congestion to the best. In this paper, we study the different queuing models that have evolved over time and how they suit the application.

Keywords: Queue Management, Temporal Fairness, Wireless LAN, RED, QoS, ECN.

1 Introduction

As the network has evolved, the traffic has increased exponentially and also the computing power of the intermediate nodes like routers and gateways, leading to more packet drop and making queuing models an active area of research. Queuing disciplines[1][2] are algorithms which control how packets are buffered and transferred through the network. The most basic queuing model proposed earlier was First In First Out (FIFO) or Drop Tail[4][6]. In this model, all the packets are treated equally by servicing them in the same order that they were placed into the queue. FIFO is the widely implemented queuing algorithm in the Internet. All the incoming packets are buffered until the buffer is full, then the incoming packets are dropped until there is memory to accept the incoming packets.

The two issues of FIFO queue are lockout and fullqueues. Lockout is a phenomenon in which more flows are denied of service as few flows occupy the queue to capacity. Fullqueue can occur due to timing issues, leading to large oscillations in the network utilization as more packets are dropped. It forces many flows to reduce their load due to congestion notification and falls below the network capacity. This oscillation between high and low nullifies the very purpose of using buffers for smoother flow.

Random Early Detection(RED)[3][4] queue model was proposed to overcome the lockout and fullqueue issues. The design goals of RED include congestion avoidance by controlling the average queue size, avoidance of global synchronization and bias against bursty traffic. In RED model, the average queue size is marked low and the

actual queue size is allowed to fluctuate to accommodate bursty traffic. It calculates the average queue size using a low pass filter, with an exponential weighted moving average. If the average queue size is below the minimum threshold, none of the incoming packets are marked called as no drop mode. Marking is a process of sending a notification to the process about the congestion, which is possible only if Explicit Congestion Notification(ECN) support is available, else the packet is dropped. If the average queue size is above the maximum threshold, all the incoming packets are marked, called as forced drop mode. When the average queue size is between the minimum threshold and maximum threshold, the incoming packet is marked with the probability P_a , which is calculated as a function of the average queue size called probabilistic drop mode. The probability of a packet from a particular flow getting marked is proportional to the share of the bandwidth used by the flow.

In the BLUE[5] model, instead of using the average queue size, the rate of packet loss and link utilization history is used for congestion avoidance. Practically, network traffic is non-poisson, making average queue length as an indicator to congestion dubious. In the implementation, ECN uses two bits of the type of service (ToS) field in the IP header. When BLUE decides that a packet must be dropped or marked, it examines one of the two bits to determine if the flow is ECN capable. If it is not ECN-capable, the packet is simply dropped. If the flow is ECN-capable, the other bit is set and used as a signal to the TCP receiver that congestion has occurred. The TCP receiver, upon receiving this signal, modifies the TCP header of the return acknowledgment using a currently unused bit in the TCP flags field. Upon receipt of a TCP segment with this bit set, the TCP sender invokes congestion-control mechanisms as if it had detected a packet loss.

The Stochastic Fair BLUE(SFB)[5] queuing model is based on bloom filters for protecting TCP flows against UDP flows with very little overhead and small amount of buffer space. In a WLAN, Access Point is a potential bottleneck because the most widely used TCP services such as web browsing or file retrieval has a very high data traffic on the downlink compared to the uplink traffic.

2 Analysis of Different Queuing Algorithms

The advantages of FIFO or Drop tail queuing model are, it is simple and easy to implement. It does not reorder the packets and has deterministic maximum delay. FIFO is best suited for networks with no bursty traffic hence the queue depth being low. FIFO also has the following drawbacks. It is not suitable for Quality of Service(QoS) as it impacts all flows equally and does not differentiate priority classes of traffic. When the traffic flow is bursty, it results in increased delay and jitter. During congestion, unfairly benefits UDP traffic over TCP.

RED algorithm is the most well known AQM algorithm. It is costly to implement and requires special supports like Explicit Congestion Notification(ECN). The average queue size is used as the indication for congestion measurement. In BLUE algorithm, the congestion measurement is achieved by rate of packet loss and link utilization history. This algorithm is best suited for bursty non-poisson traffic. The above mentioned AQM algorithms do not work well for multi-rate WLANs, and squeeze TCP traffic and increase UDP traffic flow.

It can be observed from the above studies that none of the algorithms can be generalized to be best for all kinds of traffic and network topologies. Hence proper analytical

study is to be done to understand the advantages and disadvantages of different queuing algorithms and choose one best suiting the required case.

3 Simulation of Different Queuing Models Using NS-2.33

To observe the behavior of different queuing models on a network model, Network Simulator 2.33[7][8] is used. The number of nodes used is 10, with 5 senders and 5

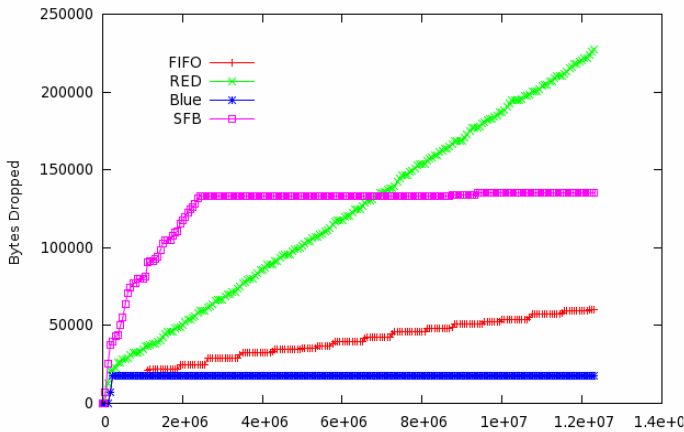


Fig. 1. Rate of Bytes dropped for Total Transmission for Different Queuing Models

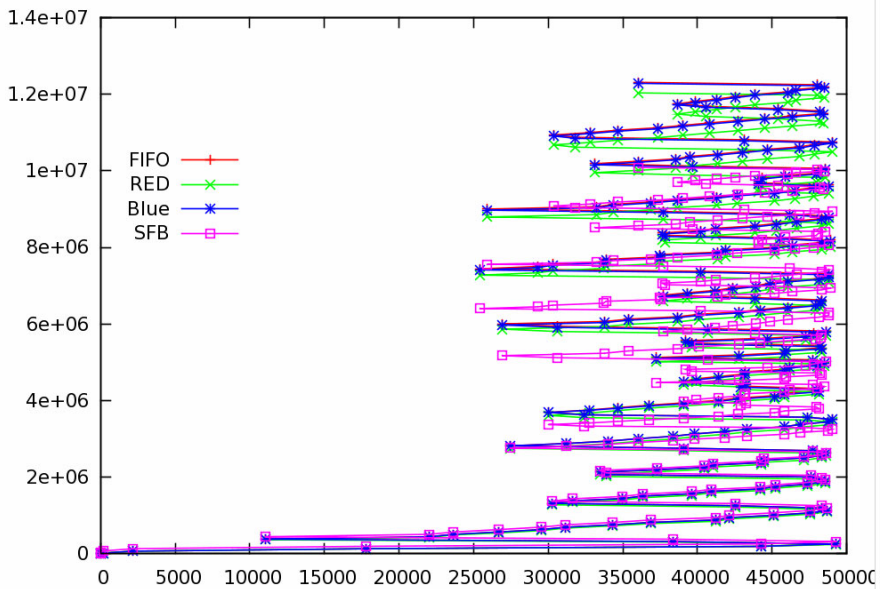


Fig. 2. Graph representing Bytes in Queue during transmission for Different Queuing Models

receivers, each with a bandwidth of 10Mbps. The senders and the receivers are separated by two routers linked together with a bandwidth of 1 mbps link. This link between the routers is buffered with a maximum queue limit of 50. The simulation time is 100 seconds. Different queuing models are applied between the routers and analyzed. It can be observed from the graphs that FIFO model performs well until the queue gets filled up. After the queue is full, the incoming packets are dropped continuously, hence the packet drop rate is exponential. The contents of the queue remain continuously full until some packets are delivered. But in the case of RED, BLUE and SFB queue models, the total bytes in the queue are kept in check with ECN ability. Hence perform better than FIFO in a congested heavy traffic environment.

References

1. Seok, Y., Park, J., Choi, Y.: Queue Management Algorithm for Multirate Wireless Local Area Networks. In: The 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication Proceedings (2003)
2. Huang, J., Wang, J., Jia, W.: Downlink Temporal Fairness in 802.11 WLAN Adopting the Virtual Queue Management. In: IEEE WCNC 2007 Proceedings (2007)
3. Barrera, I.D., Bohacek, S., Arce, G.R.: Statistical Detection of Congestion in Routers. IEEE Transactions on Signal Processing (2009)
4. Almomani, O., Ghazali, O., Hassan, S., Nor, S.A., Madi, M.: Impact of Large Block FEC with different Queue sizes of Drop Tail and RED Queue Policy on Video streaming Quality over Internet. In: Second International Conference on Network Applications Protocols and Services (NETAPPS). IEEE, Los Alamitos (2010)
5. Feng, W., Shin, K.G., Kandlur, D.: The BLUE Active Queue Management Algorithms. IEEE, Los Alamitos (2001)
6. Stanojevic, R., Shorten, R.N., Kellett, C.M.: Adaptive tuning of drop-tail buffers for reducing queuing delay. IEEE Communications Letters (2006)
7. Gao, W., Wang, J., Chen, J., Chen, S.: A Prediction-based Fair Active Queue Management Algorithm. In: Proceedings of the 2005 International Conference on Parallel Processing (2005)
8. Network Simulator, <http://www.isi.edu/nsnam/ns>

SAR Image Classification Using PCA and Texture Analysis

Mandeep Singh¹ and Gunjit Kaur²

¹ Asst. Prof., Electrical and Instrumentation Engineering Dept.
Thapar University, Patiala-147004 India

² Associate System Engineer, IBM India Pvt Ltd
Pune, Maharastra- 411057 India

Abstract. In Synthetic Aperture Radar (SAR) images, texture and intensity are two important parameters on the basis of which classification can be done. In this paper, 20 texture features are analyzed for SAR image classification into two classes like water and urban areas. Texture measures are extracted and then these textural features are further shortlisted using statistical approach, discriminative power distance and principal component analysis (PCA). In this study 30 SAR images are studied for 20 texture features. Finally, most effective 5 texture features are shortlisted for the classification of SAR images and accuracy is calculated by Specificity and Sensitivity test. The results obtained from test images give an accuracy of 95% for image classification.

Keywords: Texture analysis, Classification SAR images, PCA analysis.

1 Introduction

SAR (Synthetic Aperture Radar) image classification has been used in numerous applications like map updating, urban area classification and classification of extremely randomized clustering forests, Automatic Target Recognition (ATR) and a list to mention. Texture holds very useful information in synthetic aperture radar (SAR) image classification. There are different approaches which can be used to identify texture patterns in a given image. For efficient classification of water cover or land cover or urban area coverage, textural measures have to be chosen suitably [1]. Texture is an intrinsic property of virtually all surfaces. Classification, integrating both intensity and textural measures, will be effective in segmenting SAR images. The use of sum and difference histograms is presented for texture-based classification [2]. Clausi has compared and fused co-occurrence features, Gabor, and MRF features for sea SAR imagery classification [3]. The use of wavelet analysis for identifying specific targets is performed [4]. Dekker has investigated texture measures like histogram measures, lacunarity, wavelet energy, and fractal dimension for map updating capability [5]. In this work, the images are taken from Image database of Essex University (UK). The images are analysed to find texture measures using GLCM (Gray Level Co occurrence Matrix) and GLRLM (Gray Level Run Length Matrix). The Discriminative power and PCA (Principal component analysis) are used for dimension reduction.

2 Methodology

In this paper, the various Haralick's measures of texture [6] are computed and analysed to find the water and building areas in the SAR images. Some of the texture measures are also found using GLRLM. A gray level run is a set of consecutive, collinear picture points having the same gray level value [8]. The length of the run is the number of picture pixels in the run. The main reason for the use of the GLRLM's features is that it reflects the size of the texture elements. In the present work, total seven run length features are used to find the texture features. Chu et al [10] proposed two new features to extract gray levels information in the GLRLM and they are also used in this study; Low Gray Level Run Emphasis (LGRE), High Gray Level Run Emphasis (HGRE). The above 20 texture features are calculated for 30 SAR images taken from Image database of Essex University (UK). 30 sub images each of water and urban area are used to form two separate training sets, and for each image 20 texture parameters are extracted. Some of the SAR test images are shown in figure 1.



Fig. 1. Test image of Water (*left*) and test image of Urban area (*right*)

The images similar to these are analysed. From each image, a square ROI window (128×128) for Water and Urban land is selected and called them as Test image. After finding these training sets, box plots of these training sets are plotted taking two sets of data at a time. On the basis of box plots, it can be found that which texture feature has more value of class separability. To further reduce the dimension, discriminative power distance is used to discriminate between two sets on the basis of Fisher's discriminative distance.

$$d = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

where subscript 1 is for water and subscript 2 is for land. Therefore, μ_1 is the mean of water and μ_2 is mean of urban area coverage and σ_1^2 and σ_2^2 are the variances of water and urban area coverage respectively. After calculating the 'd', top 10 features having $d > 1.5$. To further reduce the dimension of the required number of features, we have found out Pearson's correlation coefficient of all the texture features for both water and urban area coverage. In the last, Principal Component Analysis is used to verify if the reduced feature sets are the same as obtained.

3 Results and Discussion

The training sets of water and urban area coverage are plotted in box plots to find out if a given feature is discriminative. From the box plots, we have reduced the 20 features to 14 discriminative features. Figure 2 shows the box plot of ASM and Contrast.

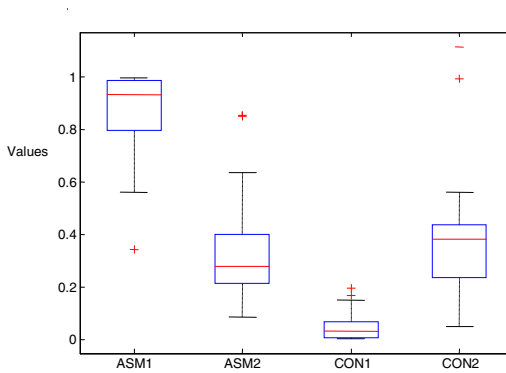


Fig. 2. Box Plots of ASM and CONTRAST for two classes

For example, in figure 2, the ASM1 (ASM of water training data) is non-overlapping with ASM2 (ASM of urban area training data). So ASM is the feature which can be shortlisted for the analysis. After analysing 14 features are selected and shown in table 1.

Table 1. Selected 14 features after using Box-Whisker plot

	WATER		Urban		d
	Mean	S.D.	Mean	S.D.	
ASM	0.871	0.160	0.340	0.190	2.138
CONTRAST	0.048	0.054	0.373	0.223	1.416
DIFF_ENTROPY	0.072	0.062	0.280	0.084	1.992
DIFF_VARIANCE	0.045	0.044	0.250	0.125	1.547
ENTROPY	0.123	0.032	0.684	0.225	2.468
IDM	0.975	0.027	0.824	0.066	2.118
INFO_MEAS_CORR2	0.465	0.228	0.917	0.075	1.883
SUM_ENTROPY	0.116	0.075	0.563	0.164	2.479
SRE (Short Run Emphasis)	0.892	0.042	0.970	0.011	1.797
LRE (Long Run Emphasis)	1.593	0.313	1.125	0.054	1.473
GLN (Gray Level Non uniformity)	1119	337	262	80	2.474
RLN(Run Length Non Uniformity)	10917	1840	14830	680	1.995
RP(Run Percentage)	0.858	0.056	0.961	0.015	1.777
HRGE	285690	278950	1490376	1001152	1.159

Then on the basis of Fisher’s discriminative distance ‘d’, we further reduced the number of features to 10. The correlation between these features is then calculated and ultimately we have reduced the feature set to just 5 features. We have analysed that the selected features like **GLN, Sum Entropy, RLN, ASM and Info_meas_corr2** are having highest discriminative distance and are least correlated. Then, we have calculated the sensitivity, specificity and accuracy to evaluate the proposed set of features for classification as tabulated in table 2.

Table 2. Specificity and Sensitivity test result

True Positive	False Positive	False Negative	True Negative	SENSITIVITY	SPECIFICITY	ACCURACY
9	0	1	10	90%	100%	95%

Thus, the proposed system has a sensitivity of 90% which means the system recognizes 90% of the images containing water. The specificity of 100% means that system is very specific in identifying water images and always identified Urban area as Urban area. Finally, the proposed system is 95% accurate. Furthermore PCA is applied to the feature set of final 5 texture features obtained after using discriminative power distance and correlation coefficient. The following plots are obtained as shown:

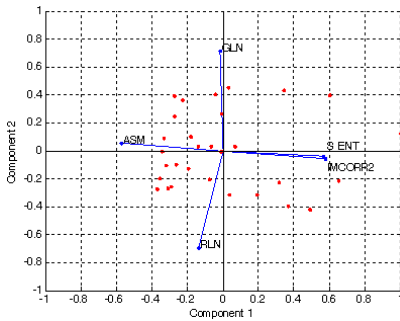


Fig. 3. PCA applied Water images

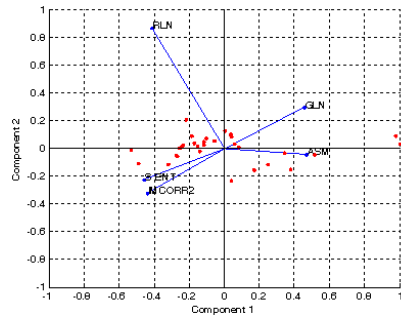


Fig. 4. PCA applied to Urban area images

From the above shown PCA plots, it is clear that all the 5 texture features are discriminating between the two training sets.

4 Conclusion

We have studied 30 images for 20 texture features. We have analysed these features and found the 5 highly discriminative features by using box plots, discriminative power distance and Pearson’s coefficient of correlation. The final five features are

- GLN (Gray Level Non-Uniformity) from GLRLM
- Sum of Entropy from GLCM
- ASM (Angular Second Moment) from GLCM

- RLN (Run Length Non-Uniformity) from GLRLM
- Information measure of correlation 2 from GLCM

Finally, we have evaluated the performance of these features on the test images and found that sensitivity is 90%, specificity is 100% and the accuracy for the classification has been 95%. So, we proposed a classification system for SAR images based on highly discriminative texture features. Furthermore, we have also applied PCA on these five features and concluded that these features are discriminative.

References

1. Chamundeeswari, V.V., Singh, D., Singh, K.: An Analysis of Texture Measures in PCA-Based Unsupervised Classification of SAR Images. *IEEE Geoscience and Remote Sensing Letters* 6, 214–218 (2009)
2. Unser, H.: Sum and difference histograms for texture classification. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8*(1), 118–125 (1986)
3. Clausi, D.A.: Comparison and fusion of co-occurrence, Gabor and MRF texture features for classification of SAR sea ice imagery. *Atmos. Oceans* 39(4), 183–194 (2001)
4. Espinal, F., Jaweth, B.D., Kubota, T.: Wavelet based fractal signature analysis for automatic target recognition. *Opt. Eng.* 37(1), 166–174 (1988)
5. Dekker, R.J.: Texture analysis and classification of ERS SAR images for map updating of urban areas in The Netherlands. *IEEE Trans. Geosci. Remote Sens.* 1(9), 1950–1958 (2003)
6. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics* (3), 610–621 (1973)
7. Rangayyan, R.M.: *Biomedical Image Analysis*. CRC Press, Washington DC (2005)
8. Loh, H.H., Leu, J.G., Luo, R.C.: The Analysis of Natural Textures Using Run Length Features. *IEEE Transactions on Industrial Electronics* 2, 323–328 (1988)
9. Galloway, M.M.: Texture Analysis Using Gray Level Run Lengths. *Computer Graphics Image Processing* 4, 172–179 (1975)
10. Chu, A., Sehgal, C.M., Greenleaf, J.F.: Use of Gray Value Distribution of Run Lengths for Texture Analysis. *Pattern Recognition Letter* 11, 415–420 (1990)

Performance of WiMAX/ IEEE 802.16 with Different Modulation and Coding

Shubhangi R. Chaudhary

Assistant Professor in E&TC dept.
Cummins College of Engineering, for women
Pune, India
shubhangirc@yahoo.com

Abstract. WiMAX (Worldwide Interoperability for Microwave Access) is the IEEE 802.16 Standards based wireless technology that provides MAN (Metropolitan Area Network) broadband and IP connectivity. WiMAX is promising technology which offers high speed voice, video and data services based on OFDM. It offers both line of sight and non line of sight wireless communication. In this paper performance of WiMAX with different modulation (BPSK, QPSK and QAM) and coding is studied on the basis of Bit Error Rate, Signal to Noise Ratio and error probability.

Keywords: WiMAX, OFDM, AMC, BER, SNR.

1 Introduction

In wireless communication, the demand of all types of services is not only voice and data but also multimedia services. A aims for the design of more intelligent communication systems, capable of providing spectrally efficient and flexible data rate access. WiMAX (Worldwide Interoperability for Microwave Access) is a new wireless OFDM-based technology that provides high throughput broadband connection over long distances based on IEEE802.16 wireless (Metropolitan Area Network) MAN air interface standard [1]. WiMAX supports a variety of modulation and coding schemes to adapt and adjust the transmission parameters based on the link quality, improving the spectrum efficiency of the system, and reaching, in this way, the capacity limits of the underlying wireless channel [4].

The WiMAX physical layer is based on OFDM. OFDM is the transmission scheme of choice to enable high-speed data, video, and multimedia communications and is used by a variety of commercial broadband systems [6]. Adaptive modulation (AMC) can effectively improve the bit error rate (BER) performance on radio channels, which had suffered from shadowing and multipath fading. Adaptive modulation enables a WiMAX system to optimize the throughput based on the propagation conditions. Using adaptive modulation scheme, WiMAX system can choose the highest order modulation provided the channel conditions are good. As the signal-to-noise ratio (SNR) is very good near the base station (BS), so higher order modulation scheme is used in this area to increase the throughput. However, in areas close to the

cell boundary, the SNR is normally poor. So, the system steps down to a lower order modulation scheme to maintain the connection quality and link stability. The supported modulations are BPSK, QPSK, 16-QAM and 64-QAM [6], [8].



Fig. 1. Scheme for the utilization of AMC

In AMC, not only the modulation order but also the forward error correction (FEC) schemes are varied by adjusting their code rate to the variations in the communication channel. An example of utilization of the cited AMC scheme is illustrated in Fig. 1. It shows that as the range increases, the system steps down to a lower modulation, but as closer to the base station, higher order modulations can be used for increased throughput.

2 System Model

The WiMAX system model is as shown in fig. 2.

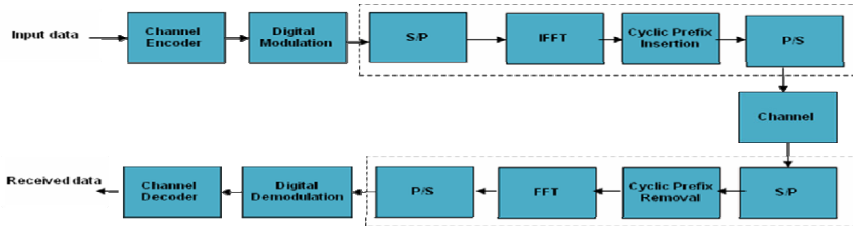


Fig. 2. WiMAX System Model

2.1 Transmitter

The data is generated from a random source, consists of a series of ones and zeros. The generated data is passed on to the next stage, either to the FEC block or directly to the symbol mapping if FEC is not used.

Channel coding

There are various combinations of modulations and code rates available in the OFDM. Channel coding includes the randomization of data, forward error correction

(FEC) encoding and interleaving. The generated random bits are coded by a concatenated Reed-Solomon (RS) and Convolutional encoder.

Modulation

There are different modulation types available for modulating the data onto the sub-carriers: BPSK, QPSK, 16QAM, and 64QAM.

IFFT: To convert mapped data, which is assigned to all allocated data subcarriers of the OFDM symbol, from frequency domain into time domain, the IFFT is used. We can compute time duration of the IFFT time signal by multiply the number of FFT bins by the sample period. Zeros are added at the end and beginning of OFDM symbol. These zero carriers are used as guard band to prevent inter channel interference (ICI).

Cyclic Prefix insertion (CP)

To avoid inter symbol interference (ISI) a cyclic prefix is inserted before each transmitted symbol. That is achieved by copying the last part of an OFDM symbol to the beginning. WiMAX supports four different duration of cyclic prefix (i.e. assuming is the ratio of guard time to OFDM symbol time; this ratio is equal to 1/32, 1/6, 1/8 and 1/4). This data is fed to the channel which represents 'Rayleigh fading channel model' and also implements multipath as shown in system model.

2.2 Receiver

The WiMAX receiver is the reverse function of WiMAX transmitter. The received signal is then passed through the serial to parallel converter. Then to next block.

Removal of CP: The first step after the arrival of data is to remove CP as shown in figure 2. We know that CP has no effect in case of using AWGN channel. It is useful when the multipath channel is used. If CP larger than the delay multipath ,the ISI is completely removed.

Fast Fourier Transform (FFT)

To convert received data from time domain to frequency domain, the FFT is used. Afterward, the zeros, which were added at the end and beginning of OFDM symbol (guard bands) at the transmitter are removed from the assigned places.

Demodulation

Demodulator converts the waveforms created at the modulation to the original transformed bits. The demodulator is used for decision rules with the goal of making a decision about which bit "zero" or "one", was sent.

Channel Decoder: The channel decoder consists of Viterbi decoder and RS decoder the sequence of bits coming from demodulator pass to channel decoder. The channel decoder tries to recover the original bits.

3 Results

The simulation result based on the adaptive modulation technique for BER calculation was observed with MATLAB7. The adaptive modulation techniques used in the WiMAX are BPSK, QPSK, 16-QAM, 64-QAM and 256-QAM respectively. Binary Phase Shift Keying (BPSK) is more power efficient and needs less bandwidth. On the other hand 64-Qadrature Amplitude Modulation (64-QAM) has higher bandwidth with very good output. During all simulations we got, BPSK has the lowest BER and 256-QAM has the highest BER than other modulation techniques.

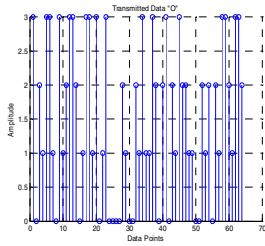


Fig. 3. Transmitted Data

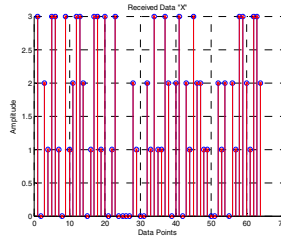


Fig. 4. Received Data

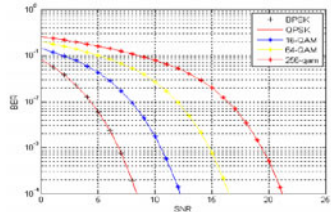


Fig. 5. OFDM with Adaptive Modulation Techniques In PURE AWGN

4 Conclusion

Thus depending on the channel condition modulation and coding can be employed.

For larger range lower modulation ie BPSK is more power efficient and need less bandwidth, lowest BER but for near Base station higher modulation ie 64QAM is used which has higher bandwidth and highest BER. While QPSK and 16 QAM are for middle range.

References

1. Eklund, C.: IEEE Standard 802.16: A Technical overview of wirelessMANTM Air Interface for Broadband Wireless Access. IEEE Communication Mag. 40(6), 98–107 (2002)
2. Ghosh, A., Wolter, D.R., Andrews, J.G., Chen, R.: Broadband Wireless Access with WiMax/802.16: Current Performance Benchmarks and Future Potential. IEEE Communication Mag. 43(2), 129–136 (2005)
3. Arafat, O., Dimyati, K.: Performance Parameter of Mobile WiMAX: A Study on the Physical Layer of Mobile WiMAX under Different Communication Channels & Modulation Technique. In: Second International Conference on Computer Engineering and Applications (ICCEA), pp. 533–537 (2010)
4. Andrews, J.G., Ghosh, A., Muhamed, R.: Fundamentals of WiMAX Understanding Broadband Wireless Networks. Prentice-Hall, Englewood Cliffs (2007)

5. Ashraful Islam, M., Zahid Hasan, M.: Performance Evaluation of WiMAX Physical Layer under Adaptive Modulation Techniques and Communication Channels. *International Journal of Computer Science and Information Security* 5(1) (2009)
6. Fazel, K., Kaiser, S.: *Multi-Carrier and Spread Spectrum System*, 2nd edn. John Wiley and Sons Ltd., New York (2003)
7. Proakis: *Digital Communications*, 4th edn.
8. Zerrouki, H.: High Throughput of WiMAX MIMO-OFDM Including Adaptive Modulation and Coding. *International Journal of Computer Science and Information Security* 7(1) (2010)

A Novel Stair-Case Replication (SCR) Based Fault Tolerance for MPI Applications

Sanjay Bansal¹, Sanjeev Sharma², and Ishita Trivedi³

¹ Research Scholar, Rajiv Gandhi Prodhogiki Vishwavidya
Bhopal, India

sanjaybansalrgpv@gmail.com

² Head, School of Information Technology

Rajiv Gandhi Prodhogiki Vishwavidya

Bhopal, India

sanjeev@rgtu.net

Abstract. When computational clusters increase in size, their mean time to failure reduces drastically. We generally use checkpoint to minimize the loss of computation. Most check pointing techniques, however, require central storage for storing checkpoints. This results in a bottleneck and severely limits the scalability of check pointing, while also proving to be too expensive for dedicated check pointing networks and storage systems. We propose a Stair-Case Replication (SCR) Based MPI check pointing facility. Our reference implementation is based on LAM/MPI; however, it is directly applicable to any MPI implementation. We use the staircase method of fault-tolerant MPI with asynchronous replication, eliminating the need for central or network storage. We evaluate centralized storage, a Sun-X4500-based solution, an EMC storage area network (SAN), and the Ibrix commercial parallel file system and show that they are not scalable, particularly after 64 CPUs. We use the staircase MPI method which allows the access point in a lower complexity level to the higher complexity level which improves the efficiency of the previous method.

Keywords: check pointing, Fault tolerance, MPI, SAN.

1 Introduction

Over the last years, e-Learning, and in particular Computer-Supported Collaborative Learning (CSCL) [1], [2] applications have been evolving accordingly with more and more demanding pedagogical and technological requirements. The aim is to enable the collaborative learning experience in open, dynamic, large-scale and heterogeneous environments [3], [4], [5]. In contemporary application servers, dynamic reconfiguration capability is well addressed to provide necessary flexibility— a component in an application can be dynamically loaded, unloaded, or upgraded at runtime on platforms like JEE (Java Enterprise Edition) [6] and CORBA (Common Object Request Broker Architecture) [7], and platforms for Fractal [8] and OpenCOM [9] component model. Despite the benefits of dynamic reconfiguration, it should not impair other properties, such as availability and reliability. Software-implemented fault tolerance (FT)

mechanisms have been provided in application servers to achieve high availability and reliability. As a result, the FT mechanism should be reconfigured either [7].

In current practices, replication-based FT is widely used in application servers to improve availability and reliability of components, such as EJBs (Enterprise Java Beans) and CORBA components [10]. Moreover, adaptive FT is considered in current practices, but most of the studies concentrate on changing FT policies in a mechanism to cope with either varied timing constraints [11], or other non-functional requirements [12]. This fine grain adaption is also called parameter adaptation [13], which changes parameters in a mechanism to modify its behaviors.

Since MPI [14] provides no mechanism to recover from such failures, a single node failure will halt the execution of the entire computation. Thus, there exists great interest in the research community for a truly fault-tolerant MPI implementation. Several groups have included checkpointing within various MPI implementations. MVAPICH2 now includes support for kernel-level checkpointing of InfiniBand MPI processes [15]. Sankaran et al. also describe a kernel-level checkpointing strategy within LAM/MPI [16], [17]. Fig1 shows fault tolerant arrangement and fault tolerance support.

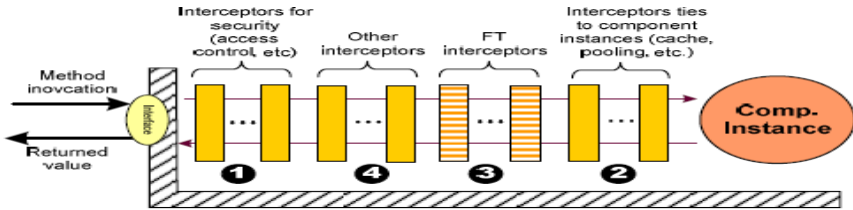


Fig. 1. Fault Tolerant Arrangement

The remaining of this paper is organized as follows. We discuss LAM/MPI in Section 2. In Section 3 we discuss about Checkpointing. The Recent Scenario in section 4. In section 5 we discuss about Algorithm. The conclusions and future directions are given in Section 6. Finally references are given.

2 LAM/MPI

MPI is suitable for parallel machines such as the IBM SP, SGI Origin, etc., but it also works well in clusters of workstations. Taking advantage of the availability of the clusters of workstations at Dalhousie, we are interested in using MPI as a single parallel virtual machine with multiple nodes. LAM is a daemon based implementation of MPI. Initially, the program lambboot spawns LAM daemons based on the list of host machines provided by the user.

LAM/MPI [16] is a research implementation of the MPI-1.2 standard [14] with portions of the MPI-2 standard. The most commonly used module, however, is the TCP module, which provides basic TCP communication between LAM processes. A modification of this module, CRTCP, provides a bookmark mechanism for checkpointing libraries to ensure that a message channel is clear.

3 Checkpointing

Checkpointing is commonly performed at one of three levels: the kernel-level, the ser-level, or the application level including language level, hardware level, and virtual machine level, also exist). In kernel-level checkpointing, the checkpointer is implemented as a kernel module, making checkpointing fairly straightforward. However, the checkpoint itself is heavily reliant on the operating system (kernel version, process IDs, etc.). Upon failure, all processes are rolled back to the most recent checkpoint/consistent state. Message logging requires distributed systems to keep track of interprocess messages in order to bring a checkpoint up-to-date.

4 Recent Scenario

One first step in the characterization of any computer system is the use of benchmarking, which allows for the analysis of the performance behavior of a system when different workloads representing the whole spectrum of possible loads are applied. Other different works have appeared for modeling a grid. Bratosin et al. in provide a formal description of grids by using Colored Petri Nets (CPN).

5 Stair-Case Replication (Proposed Method)

First compute the random replica method by the below algorithm which is given by John Paul.

Algorithm 1. Compute random replica placements.

Input-The number of replicas (Int r)

Input-The number of nodes

Output-Replica a [0...n-1]

Step1: for all I from 0 to n do

Step 2: Preload node with replica i

Step 3: for all i=0 to n do

Step 4: for j=0 to r do

z=select random node

v=Replicas[i][j] until z!=i

Step 5: for all k such tha k>=0 to r do

Step 6: if replicas[i][k]!=v

Valid replica=1

Step 7: for all k such that 0<=k<r do

Ifreplica[i][k]==vllreplicas[i][j]==repl

icas[z][k]

Valid-replica=0

Step 8: Finish

Algorithm 2 Stair-Case (Replica Result Array) Input: Input the Array

Output: enhanced replica model for fault tolerance

1. For all I such that 0< I <n do

2. Apply asynchronous replication

3. Eliminate central storage

4. Calculate the index by replica method

5. for all I such that 0 <=i < n

6. for all J such that 0 <=j <n do

7. apply replication on the lower node to the higher node

8. Finish

After that apply this algorithm to stair-case replication method on the computed replica placements to the better efficiency and the performance.

6 Conclusions and Future Work

We have shown that it is possible to effectively checkpoint MPI applications using the LAM/MPI implementation with low overhead by using staircase technique to go to the lower index to the higher index. Our stair-case replication implementation has proven to be highly effective and resilient to node failures. Because of lower to higher index concept.

References

- [1] Koschmann, T.: Paradigm shifts and instructional technology. CSCL (1996)
- [2] Dillenbourg, P.: Introduction. Elsevier Science, Amsterdam (1999)
- [3] Pankatrius, V., Vossen, G.: Towards E-Learning Grids. In: IEEE Workshop on Knowledge Grid and Grid Intelligence, Halifax, New Scotia, Canada (2003)
- [4] Caballé, S., Xhafa, F., Daradoumis, T.: A service-oriented platform for the enhancement and effectiveness of the collaborative learning process in distributed environments. In: Chung, S. (ed.) OTM 2007, Part II. LNCS, vol. 4804, pp. 1280–1287. Springer, Heidelberg (2007)
- [5] Bahrami, K., Abedi, M., Daemi, B.: AICT 2007, pp. 29–35. IEEE Computer Society, Los Alamitos (2008)
- [6] Wang, Q., Huang, G., Shen, J., Mei, H., Yang, F.: COMPSAC 2003, November 3-6, pp. 230–235 (2003)
- [7] Blair, G.S., Blair, L., Issarny, V., et al.: Proc. of Middleware, pp. 164–184 (2000)
- [8] Bruneton, E., Coupaye, T., Leclercq, M., Quema, V., Sterain, J.-B.: An open component model and its support in java. In: Crnković, I., Stafford, J.A., Schmidt, H.W., Wallnau, K. (eds.) CBSE 2004. LNCS, vol. 3054, pp. 7–22. Springer, Heidelberg (2004)
- [9] <http://www.martinfowler.com/articles/injection.html>
- [10] Narasimhan, P.: Transparent fault tolerance for CORBA (1999)
- [11] Kim, K., Lawrence, T.: Adaptive fault tolerance in complex real-time distributed computer applications. *Computer Communications* 15(4) (May 1992)
- [12] Frohofer, L., Goeschka, K.M., Osrael, J.: Middleware support for adaptive dependability. In: Cerqueira, R., Pasquale, F. (eds.) *Middleware 2007*. LNCS, vol. 4834, pp. 308–327. Springer, Heidelberg (2007)
- [13] McKinley, P., Sadjadi, S., Kasten, E., Cheng, B.: Composing adaptive software. *IEEE Computer* 37(07), 56–64 (2004)
- [14] The MPI Forum. MPI: A Message Passing Interface. In: Proc. Ann. Supercomputing Conf. (SC 1993) (ICPP 2006), pp. 471–478 (2006)
- [15] Burns, G., Daoud, R., Vaigl, J.: LAM: An Open Cluster Environment for MPI. In: Proc. Supercomputing Symp., pp. 379–386 (1994)
- [16] Sankaran, S., Squyres, J.M., Barrett, B., Lumsdaine, A., Duell, J., Hargrove, P., Roman, E. (2005)
- [17] Squyres, J.M., Lumsdaine, A.: A Component Architecture for LAM/MPI (2003)
- [18] InfiniBand Trade Assoc., InfiniBand (2007), <http://www.infinibandta.org/home>
- [19] Myricom, Myrinet (2007), <http://www.myricom.com>

Evaluating Cloud Platforms- An Application Perspective

Pankaj Deep Kaur and Inderveer Chana

Computer Science and Engineering Department, Thapar University, Patiala, India
{pankajdeep.kaur, nderveer}@thapar.edu

Abstract. Cloud computing has gained tremendous consideration recently. In the ‘Era of Tera’ when data sizes are continuously escalating and traffic patterns have become unpredictable; Cloud is a viable alternative for enterprises to serve their consumers with quicker response times. The businesses can rely on service provider to host their applications and can thus focus on their core competencies. However, a large number of Cloud providers have spurred up with their offerings. As the number of Cloud Computing players grows, the run time services provided to the cloud consumers act as key differentiators. In this paper, traditional applications are compared with Cloud applications and metrics relevant to the Cloud services have been identified. Further, the services of five major Cloud providers are compared for fundamental differences that exist in their offerings on the basis of these metrics.

Keywords: Cloud computing, Web services, Utility Computing, Virtualization, IaaS, PaaS, SaaS.

1 Introduction

Cloud Computing is an emerging approach focused to decouple the delivery of computing services from the underlying technology. This new technology intends to provide *Everything as a Service (XaaS)*. The cloud service model is categorized into three layers eminently Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) [2][12]. Cloud Consumers (CC) provision computing capabilities provided by the Cloud Service Provider (CSP) and in turn pay whatever they use and as per the duration of usage [3]. The resources in the cloud are abstracted and can be accessed using easy web interfaces. The interfaces reduce complexities by exposing minimum set of capabilities based on target use cases [4]. Cloud services shifts businesses cost from Capital Expenditure (CAPEX) to Operational Expenditure (OPEX) and are often associated with Utility Computing [6].

2 Traditional Applications versus Cloud Applications

Cloud applications share the design objectives of Distributed applications. These objectives abbreviated as IDEAS refer to the Interoperability, Distributed scale

out, **Extensibility**, **Adaptivity** and **Simplicity** aspects of distributed applications [1]. Acting as a hybrid between the desktop applications and traditional web applications, a variation in the traditional application life cycle phases is required to architect cloud applications.

2.1 Application Design

Traditional applications are designed to meet a stable demand rather than fluctuating loads prevalent in cloud environments. Applications should be designed in a way to use the underlying resources only when needed and achieve instant scalability as per demand [7].

2.2 Application Deployment

Traditional applications are deployed manually or using automated scripts. Automated scripts are faster than manual deployment but suitable for smaller workloads [5]. Image Based deployment can be used for Cloud environment in which software stack image can be instantly copied onto the target system.

2.3 Application Execution

Application execution usually requires the exposure of jobs to the scheduling system and is assigned to be executed at a later stage. In contrast, cloud applications are not exposed to the scheduling system. Application execution consists of requesting an instantiation of Virtual Machine (VM) which is assigned by the user or the middle-ware [1].

3 Evaluation Metrics for the Cloud Market Place

A large number of Cloud platforms are available in the global market. The runtime management services act as a key differentiator for various cloud platforms. In this section, some metrics are used to evaluate cloud platforms and a comparative study of the cloud offerings provided by the five major industry giants eminently Amazon Web Services (AWS) [7], Google App Engine (GAE) [8], Microsoft Azure [9], Rackspace [10] and GoGrid [11] are summarized in Table 1 and Table 2.

Table 1. Comparative Table distinguishing the Cloud PaaS Industry Giants

Metrics	Microsoft Azure	Google AppEngine
Service Launched	Feb 2010	Beta Version April 2008
Application Environment	Offers a .NET based framework that scales transparently. Server sizes vary with number of CPU cores, memory and local disk space.	Offers a component based framework. Applications run within Java or Python runtime environment. No specification of server sizes.
Virtualization Technology	Modified Hyper-V hypervisor	Undisclosed
Data Storage	Blob, Tables, Queues, SQL Azure	Google BigTable, GQL, MemCache
Computation	Web Role, Worker Role	Undisclosed

Table 1. (continued)

Communication	Azure Storage Queues	Task Queues
Prog. Lang support	.NET, PHP, Java	Java, Python
Underlying Operating system	64-bit Windows Server 2008	Linux, Windows, Mac OS
Pricing	Compute time is charged based on the amount of time an instance is processing transactions.	Compute time is charged based on the amount of time an instance is deployed.
Resources Unit		
Bandwidth Out	GB	\$0.12
Bandwidth In	GB	\$0.10
CPU Time (small instance)	Hours	\$0.10
Stored Data	GB/Month	\$0.15
Recipients emailed	Recipients	\$0.0001
	N/A	
Free Usage	Yes, Introductory offers base level of monthly usage till Oct 31, 2010.	Yes, Free default quota
Service Availability	99.95% external connectivity of role instances.	Currently does not implement SLA, Atleast 99.9% availability proposed.
Data Replication	Data replicated thrice regardless of storage option used.	Defines data location (primary, alternate)and Read policy (strong or eventual consistent)
Security:User Authentication And Authorization	Windows Live ID AppFabric Access Control Services	Google Accounts and associated URL paths

Table 2. Comparative Table distinguishing the Cloud IaaS Industry Giants

Metrics	Amazon Web Services	Rackspace	GoGrid
Service Launched	2002	June 2009	March 2008
Application Environment	Delivers empty virtual machine. Server sizes vary with the amount of compute capacity.	Allows choosing Operating system and server size based on physical memory.	Allows choosing operating system and server size based on RAM allocations.
Virtualization	Xen	Xen	Xen
Data Storage	Amazon S3, EBS, Simple DB, RDS	Uses network attached storage devices	Cloud Storage
Computation	Elastic Compute Cloud Elastic Map Reduce	Rackspace Cloud Servers	GoGrid Cloud servers Dedicated servers.
Communication	Simple Queuing Service	Undisclosed	Undisclosed
Programming Language support	PHP, Java, .NET, Python, Ruby	.NET, Perl, PHP, Python, Ruby on Rails	Java, .NET, Perl, PHP, Python, Ruby on Rails

Table 2. (continued)

Underlying Operating system	Red Hat Linux, Windows Server 2003/2008, Solaris, openSUSE, Fedora, Gentoo Linux.	64-bit Linux Distributions or Windows Server 2008 , Windows Server 2003	Linux, Microsoft Windows, CentOS and Red Hat Enterprise
Pricing Resource Unit	Prices vary with instance types and regions. Charges are calculated from the time an instance is launched until it is terminated.	Prices vary with the amount of physical memory reserved. Compute time calculated based on deployed time and not utilization.	Server usage calculated by RAM hour i.e. total amount of RAM deployed multiplied by the total number of hours it has been deployed.
Bandwidth GB Out	\$0.08(over 150 TB) ~\$0.15(first 10 TB)	\$0.22	\$0.07 (57 TB)~\$0.20 (500 GB)
Bandwidth GB In	Free until Nov 1, 2010. After that \$0.10.	\$0.08	Free
CPU Time (small Hours instance)	Free until Nov 1, 2010. After that \$0.10.	\$0.06 Linux ~\$0.08 Windows	\$0.175
Stored Data GB/Month	\$0.08 Linux/Unix ~\$0.12 Windows	\$0.15 (unlimited no. of files each up to 5 GB)	\$0.15 after storage exceeds 10 GB
Recipients Recipients emailed	\$0.055(over 5000 TB) ~\$0.15(first 50 TB)	N/A	N/A
Free Usage	Incoming Bandwidth Free until Nov 1, 2010.	-Free bandwidth between different cloud servers	- F5 Hardware Load balancing -10 GB Cloud Storage/month -Up to 16 IP addresses
Service Availability	99.95% availability, 99.9% uptime	100% availability 100% uptime	100% uptime, 10,000% service credits for SLA violations
Data Replication	Multiple Availability Zones	Three data copies across logical zones	Content delivery Network (CDN)
User Authentication and Authorization	AWS Account ID Access Control List (ACL)	User ID, API access key, session authentication token	GoGrid partner GSI (GoGrid Server Images)

4 Conclusion

As the Cloud Computing technology is gaining wider acceptance, providers are offering varied feature sets for its consumers. It can be concluded that AWS is offering the largest array of cloud services and ranks best in terms of its Global presence (datacenters in four regions) and pricing. However, it lacks behind Rackspace and GoGrid in terms of 24/7 support and freebies offered. Also, the SLA of Rackspace and GoGrid ensure 100% uptime yet it needs to be considered as a business tactic that trades off the outages with the service credits. In terms of credits, both Rackspace and GoGrid provide 100% credit for its warranties while Amazon caps its credit at 10%

per period. Furthermore, Rackspace's clear specification of one hour time to resolve failures may lure consumers, as compared to other providers, which fail to specify any such time bound warranties. As PaaS provider's, both Google and Microsoft are tough competitors. However, Azure's flexibility offered in terms of running a database and web server of consumer's choice gains advantage over Google's proprietary solutions. Nevertheless, in terms of scalability, GAE's autoscale capability outshines Azure's need based configurations. For hosting purposes, GAE is the best choice as users are provided free support until an application becomes large enough where a developer can easily pay while earning profits. The cloud market being extremely dynamic, with varied feature sets, requires application needs and customer's requirement to be meticulously considered for choosing the best provider.

References

- [1] Jha, S., Katz, D.S., Luckow, A., Merzky, A., Stamou, K.: Cloud Book Chapter. In: Understanding Scientific Applications for Cloud Environments. John Wiley & Sons, Chichester (2010)
- [2] Youseff, L., Butrico, M., Da Silva, D.: Toward a Unified Ontology of Cloud Computing. In: Grid Computing Environments Workshop (GCE 2008), Austin, Texas, USA, November 2008, pp. 1–10 (2008)
- [3] Armbrust, M., et al.: Above the Clouds: A Berkeley View of Cloud Computing. UCB/EECS-2009-28 (February 10, 2009)
- [4] Jha, S., Merzky, A., Fox, G.: Using Clouds to Provide Grids Higher Levels of Abstraction and Explicit Support for Usage Modes. *Concurrency and computation: Practice and Experience* 21(8), 1087–1108 (2009)
- [5] Server provisioning Methods holding back cloud computing initiatives, Ed Scannell, http://searchcio.techtarget.com/news/article/0,289142,sid182_gci1517504,00.html
- [6] Kaur, P.D., Chana, I.: Unfolding the Distributed Computing Paradigms. In: International Conference on Advances in Computer Engineering, ACE 2010, pp. 339–342 (2010)
- [7] Amazon Web Services, <http://aws.amazon.com/what-is-aws/>
- [8] Google App Engine-Google Code, <http://code.google.com/appengine/>
- [9] Windows Azure Platform-, <http://www.microsoft.com/windowsazure/>
- [10] Rackspace Hosting, <http://www.rackspace.com/index.php>
- [11] GoGrid, <http://www.gogrid.com/cloud-hosting/>
- [12] Pastaki Rad, M., et al.: A Survey of Cloud Platforms and Their Future. In: Computational Science and Its Applications (ICCSA 2009), pp. 788–796 (2009)

An Intelligent Agent Based Temporal Action Status Access Control Model for XML Information Management

N. Jaisankar and A. Kannan

Department of Computer Science and Engineering Anna University, Chennai-25
jaisasi_win@yahoo.com, kannan@annauniv.edu

Abstract. The enormous amount of XML data growing on the Web raises several security issues that current XML standards do not address. The most important security issue in such a distributed environment is the lack of efficient Access Control and authorization for XML data currently. In this paper, we propose a generalized Role Based Access Control (RBAC) model called An Intelligent Agent Based Temporal Action Status Access Control (IATASAC) model which uses temporal and action status constraints for efficient access control. This model addresses certain shortcomings of RBAC model by the use of semi structured data; hence it is more suitable for distributed web environment. Since the proposed model has been developed for XML data and needs smart decision, it heavily relies on XPath and intelligent agents for effective querying and answering. The Experimental results show that the proposed model performs well.

Keywords: Intelligent agent, Action status, XML, Ascribed status, Temporal constraints, Role Based Access Control.

1 Introduction

As a large quantity of information is presented in XML format on the web, there are increasing demands for XML security which is not addressed by the current XML standards. Until now, research on XML security has focused only on the security of data communication, using either digital signatures or encryption technologies. However, XML security involves not only with communication security but also with managerial security. In particular, access control and authorization issues are crucial in this huge distributed environment but have seldom been addressed. Moreover, most works in the literature did not give importance to the temporal aspects. However, the most important issue in access control and authorization for XML data is how to control user accessibility to each datum based on temporal events and roles where the traditional methods are inefficient. Hence, it is necessary to identify suitable techniques to address temporal access control issues.

In this paper to increase the access to resources, we propose a generalized RBAC model. A key feature of this model is that decisions on requests from intelligent agents to access resources are determined by considering the intelligent agent's

ascribed status, action status and temporal constraints. Additional conditions of relevance are also considered in answering the access request. An agent's ascribed status together with action status gives a measure of the agent's overall status level. The agent's status level is used as the basis for determining authorized actions and thus is used in rendering a decision on the agent's access request varies from time to time. An ascribed status may be associated with, for instance: a particular role, a classification of trustworthiness, membership of an organization, time etc. An action status is determined from a history of the deliberative actions performed by the agent. Moreover, it supports various types of temporal constraints called instant as well as interval constraints on the enabling/disabling of roles, user-role assignment and permission assignment. Using temporal constraints and role effective decisions are made even when incomplete information is present. The next section provides a survey of related works.

2 Related Work

There are many works in the literature that deal with access control and data security. Sandhu et al. [1] presented a RBAC model based on users, roles and operations. Recent interest in RBAC has been motivated by the use of roles at the application level to control access to application data. Hao He et al. [2] proposed a RBAC model for XML information management the access control scheme is represented in XML itself, with XPath to specify the linkage between the access information and actual data. Sriram Mohan et.al. [3] Proposed an infrastructure for access control on XML document by specifying access constraints in the form of virtual security and enforcing the access constraints via query rewrite. Ninghui Li et.al. [4] Proposed the analysis techniques to maintain desirable security admin privelages, more specifically the authors defined a family of security analysis problems in RBAC model. Steve Barker [5] has introduced a generalized RBAC model called Action Status Access Control (ASAC) model based on the key aspect of autonomous changing of access control policies in response to events that involve agent action and notion of status. Steve Barker [6] described implementation of ASAC model and performance measures. Fenghua Li [7] compared the action based access control model with the other existing access control models and finally concluded that this model is best for web services. Elisa Bertino et.al. [8] proposed a Temporal-RBAC (TRBAC) model, a temporal extension of the RBAC the main feature of this model is it support for periodic enabling /disabling of roles and actions expressed by role triggers. James B.D Joshi et.al. [9] Proposed a generalized TRBAC model and specified various temporal constraints to use in user-role and role-permission assignment. Comparing with existing work proposed in this paper is different and new because it uses temporal and status level constraints for efficient access control using intelligent agents.

3 Preliminaries

Before formal description of proposed model, we described the basics of status level of the user agent and the use of temporal constraints as follows.

3.1 Status Level

Here, the term authorization is defined as in [6],

Authorized (A, U, O, S, L) → Check (member ((A, level (U, L), privileges (O, S)))
 Check (true) → grant
 Check (false) → deny

where U is a user, A is an action, O is an object, S is site, L is history of events, the operator *member* is the standard membership test operator, the function *level* computes the status level of the user and the function *privileges* returns a list of pairs (actions, status level of users allowed to perform the action) for a given object at a given site.

A user agent’s status level is determined from the user agent’s ascribed and action status. An ascribed status is a status that is associated with a particular role, a categorization of agents and agent’s actions status relates to a status an agent may achieve by doing some action. This idea has been described in Table 1 for a university database consisting of staff and students considered as an example in this work.

Table 1. Status Level Description

Ascribed Status	Actions performed by agent	Status Level
PG Student	Doing project and publishing papers	Teaching Assistant(TA)
PhD Scholar	Papers published in International journals with High impact factor and received Best Paper Award	Young Scientist/ Asst. Professor
Professor	Completed Best National Projects	Principal Scientific Advisor/Chairman

In table 1, the status of TA may be used as basis for determining what actions the student with TA status can and cannot do. He may subject to handle the classes. Similarly in the proposed model the ascribed status and action status of user agent ‘u’, that requests to access a resource ‘r’, is used to determine what action ‘u’ can do and cannot do.

3.2 Temporal Constraints

The proposed model permits to specify various temporal constraints called periodicity and duration constraints as in [9]. In this model, role assumes one of the three states: enable, disable and active. The temporal constraints periodicity is used to specify the exact intervals during which a role can be enabled or disabled and duration is used to specify duration for which role enabling or assignment is valid.

4 An Intelligent Agent Based Temporal Action Status Access Control Model

The different component of the proposed model described as follow

4.1 User Agent's Session

It receives inputs from a user agent and sends request to the Temporal Access Control manager (TACM) on behalf of the user agent. It is also responsible for sending response received from TACM to the user agent.

4.2 XML Repository

XML repository is a key component of the model and it loads and stores large amount of XML files and access control files. The XML format of access control files should have the list of users with the same status level. The status level of the user is achieved by considering both ascribed status and action status which is determined from the history of deliberate actions performed by the user as in table 1. Access files also include history of actions performed by the user, operation tree, role tree and class membership attributes.

4.3 Temporal Access Control Manager

TACM can be regarded as an application of the XML database. TACM intercepts all user requests and responds for each user session by querying the XML repository and checks permission for the user according to the role of a user based on status level and given temporal constraints. When a request is received by TACM it checks temporal constraints and status level of the user first and then obtains an access set S_a by combining the operation and target. Now it queries the XML database and obtains the list of role membership M . Finally it obtains global access set of the user G_a from M . It also obtains $T = S_a \cup G_a$ after checking validity of T . The TACM executes T and returns the result to the user session.

4.4 Administrative Manager

As it interacts directly with the XML database, it can modify the access control file. The main duty of System admin is as follows. Modification of user roles, temporal constraints and history of actions of user's membership attributes and any association between users and roles, or roles and XML objects. These duties include creating new roles, granting access permission to roles based on temporal constraints, removing access permission from roles, assigning new users to roles based on temporal constraints updating deliberative actions of user, assigning status level for the user's agent, deleting users who have no membership and deleting roles, a role which does not have children or user.

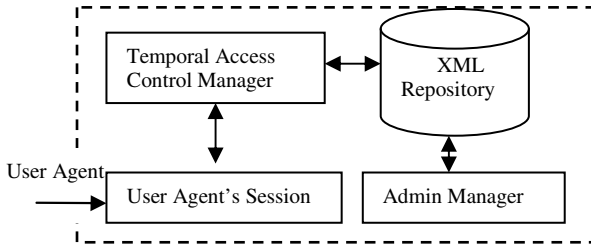


Fig. 1. System Architecture

Moreover, no direct interaction between user agent’s session and XML database allowed at all time. All interactions must be performed through the TACM. ACM as a self evolving system admin described in the above for various admin tasks. Some tasks like the total role number and role relationship and also updating the action history of users and assigning status level to users based on referring history of actions performed by the user cannot be changed without a system admin.

5 IATASAC in XML Format

The controlling information for the proposed model is stored in a configuration file in XML format. When the proposed model starts, the file is read and accordingly, user, role, and XML object associations are established.

We use the simple Extended Backus-Naur Form (EBNF) which is used for the XML specifications to design configuration file.

symbol ::= expression .

The configuration file consists of description role-tree, users, SSD, and DSD:

IATASAC -xml ::= description* role-tree user SSD* DSD where descriptions are Optional.

role-tree ::= role

The role hierarchy is represented by a role tree which has a role as root node. In the University example, University People is the root node.

role ::= role-id num-limit? role* Temporal constraints acx-f unction* admin-f n * create? (private acc-fn admin-fn* ownership-link* exception-link*) create?

The role has a unique role id, num_limit. a limitation of memberships it also has collection of job functions which may including XML access functions and administrative functions. All functions after the keyword private are not inheritable. Since ownership is unique, both ownership link and exception link are private.

acc-function ::= acc-operation* XMLPointer
 acc-operation ::= read|write|create|delete|update

An access function consists of access operations allowed and an XMLPointer pointing to the XML node object.

admin-func:tion ::= |AssignRole|DepriveRole|AddAccess|RmAccess |MvOwner
 users ::= user+, user ::= user-id user-info* p* RolePointer*

A user must have a unique user_id and may contain a short description of the user with his status level. A user must have a password. If a user has a Role Pointer pointing to a role, then the user has the membership of that role.

The XML format of Indian University role hierarchy configuration file presented in table 2.

Table 2. TASAC Model’s Configuration File in XML format

<pre><?xml version=' 1.0' ?> <!-- XML access control --> <ATASAC_xml> <users> <user id="jai"pwd="mat"></user> <user id="raj"pwd="123"></user> <user status="TAforPG"></users> <role_tree> <role id="Uiversity People" > <role id="Student" > </role> <role id="UG" > </role> <role id="PG" </role> <role id="Ph.D" > <role id="TA" > </role></role></pre>	<pre><role id="Staff"> <role id="admin"> <role id="PA"> </role> </role> <role id="Manager"> <role id="Academic"> <role id="Prof."> <role id="Dean "> <role id="VC"> </role> </role> <role id="Lect."> </role> </role> </role></role></role_tree> </ATASAC_xml ></pre>
--	---

For example, if Access(student, read, attendance) is true then the role student can read attendance, which may be owned by, say, the role professor. Since an XML file forms a tree, if all child nodes share the same access permissions, access control information can be stored only at their parent node.

6 Results and Discussion

The Fig.2 shows the number of access violations prevented by intelligent temporal access status access control model is higher than the traditional TRBAC and RBAC and Fig.3 shows that the document retrieval time while performing the proposed system is considerably less while comparing other traditional access control models.

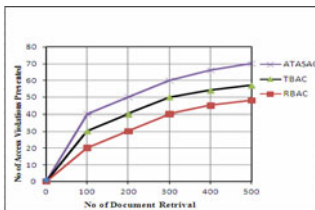


Fig. 2. Analysis of Access Violation Prevented

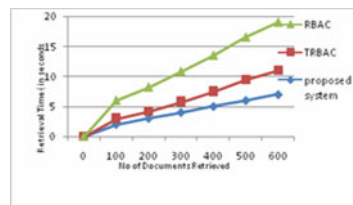


Fig. 3. Performance Analysis

7 Conclusion

In this paper, An Intelligent Agent based Temporal Action Status Access Control Model for XML information management has been proposed. The Access Control (AC) Scheme is represented in terms of XML format, with XPath to specify the linkage between the access information the ease of administration and control over the data and separation of duty. The experimental results show that the proposed model works well when compared with traditional TRBAC and RBAC. Further work in this direction is to improve the model with spatio-temporal constraints for providing improved security.

References

1. Sandhu, R.S., Coyne, E.J., Feinstein, H.L., Youman, C.E.: Role-Based Access Control Models. *IEEE Computer* 29(2), 38–47 (1996)
2. He, H., Wong, R.K.: A role-based access control model for XML repositories. In: proceeding of the first international conference on web information system engineering, vol. 1, pp. 138–145. *IEEE explorer*, Los Alamitos (2000)
3. Mohan, S., Sengupta, A., Wu, Y.: A Framework for Access Control for XML. *ACM Transactions on System and Information Security* 5, 1–38 (2006)
4. Li, N., Mahesh, V., Tripunitara: Security Analysis in Role-Based Access Control. *ACM Transactions on Information and System Security* 9(4), 139–420 (2006)
5. Barker, S.: Action Status Access Control. In: Proceedings of the 12th ACM symposium on Access control models and technologies (SAGMAT), pp. 195–204 (2007)
6. Barker, S.: Access control by Action Control. In: Proceedings of the 13th ACM symposium on Access control models and technologies (SAGMAT), pp. 143–152 (2008)
7. Li, F., Wang, W., Jianfengna, Su, H.: Action-Based Access Control for Web Services. *Journal of Information Assurance and Security* 5, 162–170 (2010)
8. Bertino, E., Bonatti, P.A., Ferrari, E.: TRBAC: A Temporal Role Based Access Control Model. *ACM Transactions on Information and System Security* 4(3), 191–223 (2001)
9. Joshi, J.B.D., Bertino, E., Latif, U.: A Generalized Temporal Role Based Access Control Model. *IEEE Transactions on Knowledge and Data Engineering* 17(1), 4–23 (2005)

An Economic Auction-Based Mechanism for Multi-service Overlay Multicast Networks

Mohammad Hossein Rezvani and Morteza Analoui

School of Computer Engineering, Iran University of Science and Technology (IUST)
16846-13114, Hengam Street, Resalat Square, Narmak, Tehran, Iran
{rezvani, analoui}@iust.ac.ir

Abstract. Recently, strategic behavior modeling has attracted much attention of the researchers focusing on designing protocols in the area of social networks. The motivation lies in the fact that the overlay peers of the social networks are selfish in nature and they typically belong to different administrative domains. In this paper, we model the strategic behavior of the selfish peers by leveraging the rich theory of mechanism design using the concept of economic auctions. By considering the bandwidth of the service offered by the origin server as the commodity, we design dynamic auctions in which downstream peers submit their value of bids for each commodity at the upstream peers.

Keywords: Social Networks, Multicasting, Resource Allocation, Strategic Behavior, Mechanism Design, Auction Games.

1 Introduction

There exists a significant body of researches towards "*social networks*" in the literature. Although online video streaming existed long before social-network-assisted sites such as YouTube, the establishment of social networks has become an equally or even more important factor toward their success. On the other hand, "*overlay multicasting*" solution has recently been accepted as a paradigm shift in order to disseminate the digital real-time contents such as teleconferencing, IPTV, and online video broadcasting in large peer-to-peer (P2P) networks.

With respect to this fact that the selfish behavior of the overlay peers is inevitable in real P2P networks, it turns out that the most important question towards designing the overlay multicast protocols is the following; "How should the selfishness of the peers is exploited, so that the aggregate throughput of the overlay network still is maximized?" There have recently been a significant body of research towards designing self-organizing overlay multicast networks by exploiting the inherent selfishness of the peers. These works can be categorized into two significant strands: "*strategic behavior modeling approaches*" such as [1, 2, 3] and "*non-strategic behavior modeling approaches*" such as [4, 5]. In the former, each peer is treated as a potential game player that seeks to maximize its utility regard to the actions that the other peers do, whereas in the latter each peer seeks to maximize its utility without taking into account the actions of the other peers. We present a revenue-maximizing monopoly auction framework by leveraging the mathematical tools from the theory of "*mechanism*

design" of microeconomics. In this framework, each offered service in a typical social network, plays the role of the commodity in the overlay network economy. The buyers in this economy are all the peers either who relay the services to their downstream peers or who are leaf nodes in the overlay multicast tree. Also, the sellers are either the origin servers or the peers who forward the media services to the other peers of the social network. The remainder of this paper is organized as follows: We discuss the related works in Section 2. Section 3 is devoted to formal description of the proposed auction mechanism for the overlay multicast network as well as its associated theories. Section 4 specifies the performance evaluation of the proposed mechanism. Finally, we conclude in Section 5.

2 Related Work

There are simultaneous works investigating social networks in popular Web 2.0 sites [6, 7]. While YouTube is also one of the targeted sites in their studies, exploring the social network for accelerating content distribution has yet to be addressed. To address the selfishness of the peers based on the non-strategic behavior modeling, much of the literature has applied pricing approaches [4, 5]. A few other works have also been proposed to allocate the bandwidth based on strategic auctions [1, 2, 3]. The authors of [3] have presented an auction-based model to improve the performance of BitTorrent. A key difference between our designed auction and that of [3] is that the bids in our auction are issued based on the price, whereas in [3], the bids are submitted based on bandwidth. The major difference between our work and the work of [1, 2] lies in the fact that their approach is in actual a bargaining approach rather than a monopoly auction in the sense that the mechanism design theory of microeconomics provides.

3 The Proposed Overlay Auction Mechanism

We consider an overlay network consisting of V end hosts denoted as $\mathcal{V} = \{1, 2, \dots, V\}$. Let us suppose that the overlay network consists of N media services, denoted as $\mathcal{N} = \{1, 2, \dots, N\}$. So, there are N origin servers among V hosts ($N < V$), each serve a distinct type of media service. Suppose the network is shared by a set of N multicast groups. Any multicast group (multicast session) consists of a media server, a set of receivers, and a set of links which the multicast group uses. Let us suppose that the overlay network consists of L physical links, denoted as $\mathcal{L} = \{1, 2, \dots, L\}$. The capacity of each link, that is the bandwidth of each physical link $l \in \mathcal{L}$ is denoted as c_l . All the nodes, except the origin servers and leaf nodes, forward the multicast stream via unicast in a peer-to-peer fashion. Fig. 1 shows an overlay network consisting of two multicast groups. In this example, we can represent the set $S = \{s_1, s_2\}$ in which s_1 (node 0) indicates one group and s_2 (node 3) indicates the other group. Here, the solid lines indicate one group and dashed lines indicate the second group. Also, the physical network consists of eight links ($L = 8$) and two

routers. Each multicast session $n \in \mathcal{N}$ consists of some unicast end-to-end flows, denoted as the set \mathcal{F}^n :

$$\mathcal{F}^n = \{f_{ij}^n \mid \exists i, j \in \mathcal{V} : \mathbf{M}_{ij}^n = 1\}. \tag{1}$$

Where \mathbf{M}^n denotes “adjacency matrix” of the multicast group n . Each flow f_{ij}^n of the multicast group n passes a subset of physical links, denoted as

$$\mathcal{L}(f_{ij}^n) \subseteq \mathcal{L}. \tag{2}$$

For each link l , we have

$$\mathcal{F}^n(l) = \{f_{ij}^n \in \mathcal{F}^n \mid l \in \mathcal{L}(f_{ij}^n)\}. \tag{3}$$

Where $\mathcal{F}^n(l)$ is the set of the flows belonging to the multicast group n and passing through the link l . Each flow $f_{ij}^n \in \mathcal{F}^n$ in the multicast group n has a rate x_{ij}^n . We show the set of all downstream nodes of each overlay node i in the multicast group n by $Chd(i, n)$. Also, the set $Buy(i)$ specifies all the multicast groups in the overlay network from which the node i receives (buys) the services. Similarly, $Sell(i)$ specifies all the multicast groups in the overlay network for which the node i provides (sells) the services.

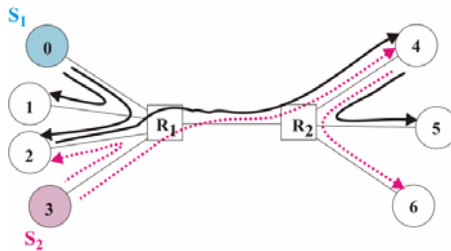


Fig. 1. Overlay network consisting of two multicast groups

There are a number of standard auction forms that the seller might use to sell the commodity: first-price, second-price, Dutch, English. Our designed mechanism, named overlay monopoly auction mechanism (OMAM), lends itself to explain the equivalence of revenue in all four auction forms. Let us suppose OMAM in n -th multicast group in which the monopolist seller s wants to sell the bandwidth of service n to one of $K_{s,n}$ bidders. OMAM is a collection of $K_{s,n}$ probability assignment functions, $pr_{i,n}(v_{i,n}, v_{-i,n})$, $i \in Chd(s, n)$, $n \in \mathcal{N}$ and $K_{s,n}$ cost functions $c_{i,n}(v_{i,n}, v_{-i,n})$. For each bidder i and every vector of values $(v_{i,n}, v_{-i,n})$, $pr_{i,n}(v_{i,n}, v_{-i,n}) \in [0, 1]$ denotes the probability that bidder i receives the bandwidth of service n (in actually the commodity in multicast group n) and $c_{i,n}(v_{i,n}, v_{-i,n}) \in \mathbb{R}$ denotes the payment that bidder i must make to seller s . Also,

note that $-i, n$ denotes all other bidders apart from bidder i in multicast group n . Consequently, the sum of the probabilities $\sum_{i \in Chd(s, n)} pr_{i, n}(v_{i, n}, v_{-i, n})$, must never

exceeds unity. On the other hand, we allow this sum to fall short of unity because we want to allow the seller to keep the commodity. OMAM works as follows. Because the seller does not know the bidders' values, it asks them to report the values to it simultaneously. It then takes those reports $r_{i, n}$, which need not be truthful, and assigns the commodity to one of the bidders according to the probabilities $pr_{i, n}(v_{i, n}, v_{-i, n})$, keeping the commodity with the residual probability, and secures the payment $c_{i, n}(v_{i, n}, v_{-i, n})$ from each bidder $i \in Chd(s, n)$. It is assumed that the entire direct selling mechanism, i.e., the probability assignment functions and the cost functions, are public information, and that seller s must carry out the terms of the mechanism given the vector of the reported values. Clearly, the seller's revenue will depend on the reports submitted by the bidders. Now, the main question is that: "will the bidders be induced to report truthfully?"

Definition 1 (Incentive-Compatible OMAM). *OMAM is incentive compatible (IC) if it is an equilibrium for the bidders to report their values truthfully.*

Understanding IC OMAM will not only be the key to understanding the connection among the four standard auctions, but also it will be central to understanding revenue-maximizing auctions as well. As it is the case in microeconomics, beginning with the equilibrium of any of the four standard auctions, we can similarly construct our IC OMAM that yields the same ex-post assignment of the commodity to the bidders and the same ex-post payments by them. Interested readers can refer to chapter 9 of [8] to find an in-depth discussion on this topic as well as the associated theories.

By incentive compatibility, each bidder must find it optimal to report its true value that all other bidders do so. We now proceed to state the necessary conditions for optimal OMAM in such a way that leads to maximization of the revenue for seller s as well. At first, let us define the concept of individual rationality. In order for OMAM to be optimal for the seller, the so called property must hold.

Definition 2 (Individual Rationality). *In microeconomics, a mechanism is said individually rational (IR), if it yields each bidder, regardless of its value, a non-negative expected payoff in the truth-telling equilibrium.*

Clearly, if the expected payoff of a bidder is negative, it will simply not participate in the selling mechanism. The following theorem completely characterizes the optimal OMAM:

Theorem 1 (Optimal OMAM). *In IC IR OMAM, suppose that the private value of each bidder i in multicast group n is drawn from the continuous positive density function $f_{i, n}$ satisfying (4) as follows*

$$v_{i, n} - \frac{1 - F_{i, n}(v_{i, n})}{f_{i, n}(v_{i, n})} \text{ is strictly increasing in } v_{i, n}. \tag{4}$$

Then the probability assignments and the cost functions defined in (5) and (6) yield the monopolist seller s the largest possible expected payoff.

$$pF_{i,n}^*(v_{i,n}, v_{-i,n}) = \begin{cases} 1, & \text{if } v_{i,n} - \frac{1 - F_{i,n}(v_{i,n})}{f_{i,n}(v_{i,n})} > \max(0, v_{j,n} - \frac{1 - F_{j,n}(v_{j,n})}{f_{j,n}(v_{j,n})}) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$$c_{i,n}^*(v_{i,n}, v_{-i,n}) = pF_{i,n}^*(v_{i,n}, v_{-i,n})v_{i,n} - \int_0^{v_{i,n}} pF_{i,n}^*(w, v_{-i,n})dw \quad (6)$$

4 Performance Evaluation

In order to handle the operations of the OMAM, we consider a dedicated server, named *Overlay Market Control Server (OMCS)*. The OMCS bears the characteristics of "rendezvous point" in the former well-known research projects such as [5, 9]. The OMCS contains the information of all markets including the free uploading and downloading capacities of the peers, the rates of the services that are allocated to each peer, the distribution of the values concerning to each service, the information of each physical link, and so on. Due to space limitation, we do not mention the details of the algorithms here. We use BRITE topology generator [10] to set up our experimental network. Each overlay peer is an end host attached to a single router. The backbone includes 512 routers and 1024 backbone edges (physical links). The bandwidths of all physical links have Heavy-tailed distributions in the interval [10 Mbps, 100 Mbps]. The propagation delay of each individual underlying link is uniformly distributed in the interval [1 ms, 2 ms]. The overlay peers are randomly connected to backbone routers through access links, whose capacities are exponentially distributed with an average of 15 Mbps. We also assume $CD_i = 2 \times CU_i$, that is downloading capacity of each access link is two times bigger than its uploading capacity. The maximum tolerable delay and the maximum tolerable loss rate of each flow are 1 Second and 5% respectively. The maximum allowed bandwidths of each service, namely B^n , is 3 Mbps. The peers join the network following a Poisson process. The inter-arrival times follow an exponential distribution with an expected length of 1 second. Upon arrival, each peer randomly selects some services; then the peer stays in the network for a certain period of time, following an exponential lifetime distribution with an expected length of 30 minutes.

Figure 2 shows the average throughput of the peers (the average social welfare) resulting from the OMAM in comparison with the case in which no priced-based mechanism is used. Also, for the sake of completeness, we have compared the OMAM with the average upper bound throughput. By "average upper bound throughput", we mean the average upload capacity of non-leaf peers in all multicast trees. Clearly, the aggregate demands of the overlay peers cannot exceed the sum of the uploading capacities of the non-leaf peers of the multicast trees. So, we can gain further insights into the OMAM by evaluating it with the average upper bound throughput as the best-case metric. It is evident from Fig. 2 that the resultant average social welfare of the OMAM is so much better than the case in which no price-based mechanism is used. The reason of decreasing the average throughput is the data constraints and the network constraints of the multicast solution. According to data constraint, a peer cannot forward the stream to its downstream peers at a rate higher than

its own receiving rate. The network constraint arises mainly due to inadequate downloading or uploading capacities of the peers. As the number of peers increases, the tree becomes deeper. Thus, the degradation of throughput in lower levels of the tree results in degradation of the average throughput.

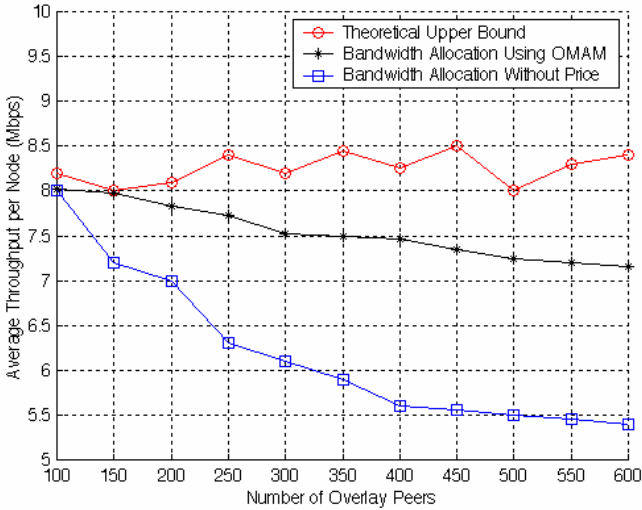


Fig. 2. Average throughput per user (average social welfare)

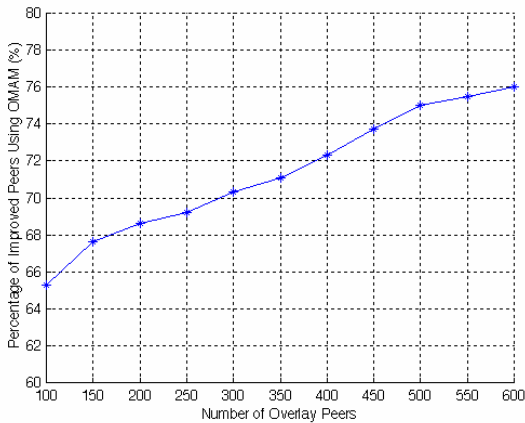


Fig. 3. Percentage of welfare improvements resulting from OMAM

Figure 3 illustrates the percentage of welfare improvements by the OMAM for different peers' population sizes. To this end, for each population of peers, we have logged the perceived utility of each peer for both cases of using the OMAM and the non-priced mechanism and then have compared these two values with each other. Next, we have normalized the number of improved peers by the total number of peers

and have represented the result in the form of percent of improvements. It is clear from the figure that using the OMAM enhances the perceived quality of the services in the multicast sessions.

5 Conclusion

In this paper, we provided a social-network-assisted framework for multi-service multi-rate overlay multicasting based on mechanism design theory of microeconomics by taking into account the inherent selfishness of the peers. The design of the strategic behavior model has been done by leveraging the concept of auction games. We proved the convergence of the auction-based strategies. Also, we showed that the mechanism is incentive compatible and individually rational and also leads to performance improvements in terms of the social welfare.

References

1. Wu, C., Li, B.: Strategies of Conflict in Coexisting Streaming Overlays. In: INFOCOM, pp. 481–489 (2007)
2. Wu, C., Li, B., Li, Z.: Dynamic Bandwidth Auctions in Multioverlay P2P Streaming with Network Coding. *IEEE Trans. Parallel Distrib. Syst.* 19(6), 806–820 (2008)
3. Levin, D., LaCurts, K., Spring, N., Bhattacharjee, B.: BitTorrent is an auction: analyzing and improving BitTorrent's incentives. In: SIGCOMM, pp. 243–254 (2008)
4. Analoui, M., Rezvani, M.H.: An Economic Case for End System Multicast. In: Berre, A.J., Gómez-Pérez, A., Tutschku, K., Fensel, D. (eds.) FIS 2010. LNCS, vol. 6369, pp. 40–48. Springer, Heidelberg (2010)
5. Analoui, M., Rezvani, M.H.: Microeconomics-based Resource Allocation in Overlay Networks by Using Non-strategic Behavior Modeling. *Elsevier J. Commun. Nonlinear Sci. Numer. Simulat.* 16(1), 493–508 (2011)
6. Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y., Moon, S.: I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System. In: Proc. of ACM IMC (2007)
7. Mislove, A., Marcon, M., Gummadi, K., Dreschel, P., Bhattacharjee, B.: Measurement and Analysis of Online Social Networks. In: Proc. of ACM IMC (2007)
8. Jehle, G.A., Reny, P.J.: *Advanced Microeconomic Theory*. Addison-Wesley, Reading (2001)
9. Pendarakis, D., Shi, S.Y., Verma, D., Waldvogel, M.: ALMI: an application layer multicast. In: 3rd USENIX Symp. on Internet Technologies and Systems (2001)
10. Medina, A., Lakhina, A., Matta, I., Byers, J.: BRIT: An Approach to Universal Topology Generation. In: Proc. IEEE Int'l Symp. Modeling, Analysis and Simulation of Computer and Telecomm. Systems, MASCOTS (2001)

Server Virtualization: To Optimizing Messaging Services by Configuring Front-End and Back-End Topology Using Exchange Server in Virtual Environments

R. Anand and T. Deenadayalan

Computer Science And Engineering,
Dhaanish Ahmed College Of Engineering,
Anna University, Chennai
Tamil Nadu

nowhereanand@yahoo.com, deenadayalan1982@gmail.com

Abstract. Microsoft Exchange Server supports the deployment of Exchange in a manner that distributes server tasks among front-end and back-end servers. This work includes messaging services such as “E-Mail”, “DiscussionForum”, “SearchEngine”, “NewsLetters”, “GuestBook” and “ClassifiedAdverts”. The E-Mail service that embeds “Exchange Server” as Back-End acts as the primary messaging service. Microsoft Exchange Server supports the deployment of Information exchange that distributes server’s tasks among Front-End and Back-End servers. Front-End server accepts requests from clients and proxies them to the appropriate Back-End server for processing, to optimize the Front-End and Back-End servers using virtualization technology. The Authorized Users are allowed to utilize the messaging services and special authorization is made when the user wishes to join the E-Mail service. After receiving a request, the front-end server uses LDAP (Lightweight Directory Access Protocol) to query the Windows 2003 server Active-Directory-service and determine which back-end server holds the requested resource. This topology reduces complexity.

Keywords: Server Virtualization, Virtualization, Exchange Server Virtualization.

1 Introduction

Microsoft Exchange 2003 Server supports the deployment of Exchange in a manner that distributes server tasks among front-end and back-end servers. A front-end server accepts requests from clients and proxies them to the appropriate back-end server for processing. The general functionality of the front-end server is to proxy requests to the correct back-end servers on behalf of the client computers; the exact functionality of the front-end server depends on the protocol and the action being performed.

After receiving a request, the front-end server uses Lightweight Directory Access Protocol (LDAP) to query the Microsoft Windows 2003 server Active Directory service and determine which back-end server holds the requested resource. The back-end

server then sends the results of the logon operation back to the front-end server, which returns the results of the operation back to the client. The front-end and back-end server topology is recommended for multiple-server organizations that use Microsoft Outlook Web Access (HTTP), POP, or IMAP and for organizations that want to provide HTTP, POP, or IMAP access to their users over the Internet.

The proposed network will allow us to centralize many, if not all. This reduction of servers, combined with Exchange 2003 integration with Active Directory, will provide a messaging environment that is significantly easier to manage. By putting Exchange 2003 on the servers, we can expect a significantly more reliable operating system due to the inherent reliability improvements incorporated within Exchange Server 2003.

The Existing system usually used IIS (Internet Information Services). This IIS stores the configuration information in the “metabase”, whereas Exchange stores configuration information in Active Directory. The metabase is a local configuration database shared by the protocols that IIS supports. In Exchange 2000 SP1 and earlier versions, DS Access used RPCs to connect to directory servers and discover the topology. In Exchange SP2 and later versions (Exchange 2003), DSAccess uses LDAP for most of the operations.

This research work has a vision to provide its user community with the best tools in a reliable, manageable, and secure environment. This provides a competitive advantage both internally and externally. In this earlier research work virtualize the server and also virtualize the Active directory tools. In this research work virtualize the server, active directory and mail server and also optimize the topology of the network. Both the topology can be optimized using Xen Hypervisor. In VMware to optimize the topology easy to one. In Xen based hypervisor to optimize the topology using open source environment. This research work is targeting the following five primary objectives such as

- *Increase Manageability of the Server and Workstation Environment.*
- *Increase the Security of the LAN Environment.*
- *Increase the Reliability of the Environment.*
- *Provide User Community with Updated Technology.*
- *Reduce the cost and System utility*

2 Detailed Description

2.1 Front-End and Back-End Topology without a Firewall

Scenario

A network willing to maintain a single namespace for their e-mail servers but cannot fit all of their users on a single server.

Setup Instructions

1. Set up a standard collection of servers running Exchange.
2. Set up a single server running Exchange configured as a front-end server.
3. Direct HTTP, POP, and IMAP users to this server, not to their back-end servers.
4. Ensure that all virtual directories and servers are configured identically on all front-end and back-end servers.

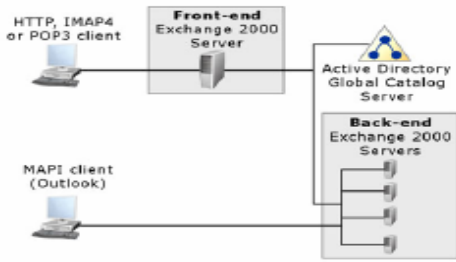


Fig. 1. Front-End and Back-End Topology

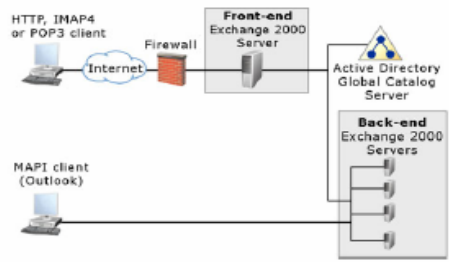


Fig. 2. Front-End and Back-End Topology behind a firewall

Issues

This is the default configuration. You do not need to perform any steps other than the standard front-end and back-end configuration steps.

If the network permits connections between the client and the back-end servers, there is nothing to prevent users from circumventing the front-end server and connecting directly to the back-end server.

If this is undesirable, you must change the network routing configuration or the back-end server configuration to prevent direct connections between a client and a back-end server.

2.2 Front-End Server behind a Firewall

Scenario

To achieve security and still provide access to Outlook Web Access, POP, or IMAP from the Internet, a corporation wants to place the Exchange system behind the corporate firewall.

Setup Instructions

1. Set up a standard Exchange front-end and back-end environment in the corporation.
2. Configure a firewall between the front-end server and the Internet. For more information about how to configure an Internet firewall for use with a front-end server running Exchange

Issues

Because the entire configuration is inside the firewall, Exchange does not require any special configuration. After a request comes through the firewall to the front-end server, the front-end server returns a response without any configuration changes.

IP address filtering is highly recommended to limit requests through the firewall to only those going to the front-end server (or servers) running Exchange and block requests through the firewall to other servers in the organization.

2.3 Configuring a Front-End Server

A front-end server is an ordinary Exchange server until it is configured as a front-end server. A front-end server must not host any users or public folders.

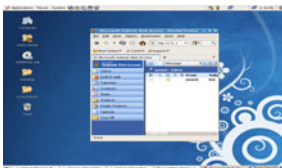
A front-end server must be a member of the same Exchange organization as the back-end servers (therefore, a member of the same Windows 2000 forest).

2.4 Configuring a Back-End Server

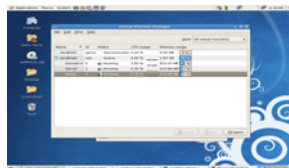
Exchange configuration is stored in Active Directory on a per-forest basis, which means that all front-end and back-end servers must be in the same forest. Back-end servers can be accessed directly if required, with no effect on the behavior of the front-end and back-end configuration. If you did not configure any extra virtual servers or directories on any front-end servers, then you do not need to configure any on the back-end server. If you created additional virtual servers or directories on any front-end servers, however, you must add matching virtual servers and directories on the back-end servers.

3 Experimental Setup

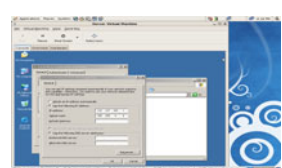
OS virtualization is achieved by inserting a layer of software between the OS and the underlying server hardware. This layer is responsible for allowing multiple OS images (and their running applications) to share the resources of a single server. Each OS believes that it has these sources of the entire machine under its control, but beneath its feet, the virtualization layer transparently ensures that resources are properly shared between different OS images and their applications. Using Xen hypervisor to configure front end server is windows 2003 server and backend server is windows 2003 server more than 2 or 3 servers can be running simultaneously. In this situation more than 2 servers takes less time to utilize the hardware. All the hardware and virtual machines information are available in hypervisor. Xen based hypervisor Domain-0 contains all the information. To avoid physical partitioning. Instead, the virtualization platform generally traps instructions issued by virtual machines and either passes the instruction through to the physical processor or emulates the instruction by issuing one or more different instructions to the physical processor and returning the expected result to the virtual processor. Depending on the virtualization platform and its configuration, it is possible for instructions from a single virtual processor to be executed sequentially across one or more physical processors in the host server. This is not in any way multiprocessing, as the instructions are not executed in parallel, but it can be performed in order to optimize the processor resource scheduler in the VMM and to help increase virtual machine performance. Using python script to configure the virtual machines. This script contains all the virtual machine configurations can use the system requirements processor: Intel Core i3M3302.13GHZ,MotherBoard:Intel5SeriesBoard3400SH,RAM:2GBDDR2,HardDisk:320GB,OS:CentOS,Hypervisor:Xen,VirtualMachine:Windows2003Server.



(a) Host Cent OS



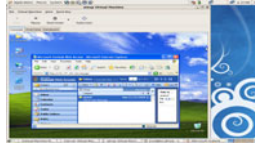
(b) Xen Virtual Manager



(c) Windows2003Server
(Guest OS)



(d) Windows Xp(Guest OS)



(e) To access Exchange Server Using Guest OS (Windows Xp)



(f) Performance Analysis of all Virtual Machines

4 Conclusion and Future Work

Server Virtualization is to optimize the front end and back end topology using exchange server. Exchange server provides secure intranet (or) internet messaging applications. The topology can be optimized using Xen hypervisor. It increases the speed and utilizes the hardware fully. Further try to optimize the topology using different virtual machine monitors then the performance is analyzed.

References

- [1] Vallée, G., Scott, S.L.: Xen-OSCAR for Cluster Virtualization. In: Min, G., Di Martino, B., Yang, L.T., Guo, M., Rünger, G. (eds.) ISPA Workshops 2006. LNCS, vol. 4331, pp. 487–498. Springer, Heidelberg (2006)
- [2] Ueno, H., Hasegawa, S., Hasegawa, T.: Virtage: Server virtualization with hardware transparency. In: Lin, H.-X., Alexander, M., Forsell, M., Knüpfer, A., Prodan, R., Sousa, L., Streit, A. (eds.) Euro-Par 2009. LNCS, vol. 6043, pp. 404–413. Springer, Heidelberg (2010)
- [3] Cafaro, M., Aloisio, G.: Grids, Clouds, and Virtualization. *Computer Communications and Networks*, 1–21 (2011)
- [4] Van Do, T., Krieger, U.R.: A performance model for maintenance tasks in an environment of virtualized servers. In: Fratta, L., Schulzrinne, H., Takahashi, Y., Spaniol, O. (eds.) NETWORKING 2009. LNCS, vol. 5550, pp. 931–942. Springer, Heidelberg (2009)
- [5] Zhang, B., Wang, X., Lai, R., Yang, L., Wang, Z., Luo, Y., Li, X.: Evaluating and optimizing I/O virtualization in kernel-based virtual machine (KVM). In: Ding, C., Shao, Z., Zheng, R. (eds.) NPC 2010. LNCS, vol. 6289, pp. 220–231. Springer, Heidelberg (2010)
- [6] Xie, W., Navathe, S., Prasad, S.K., Fisher, D., Yang, Y.: Optimizing peer virtualization and load balancing. In: Li Lee, M., Tan, K.-L., Wuwongse, V. (eds.) DASFAA 2006. LNCS, vol. 3882, pp. 357–373. Springer, Heidelberg (2006)
- [7] Hagen, W.V.: *Professional Xen® Virtualization*.
- [8] Marshall, D., Reynolds, W.A., McCrory, D.: *Advanced Server Virtualization*

Blind Source Separation for Convolutional Audio Mixing

V. Jerine Rini Rosebell, D. Sugumar, Shindu, and Sherin

Department of ECE, Karunya University, Coimbatore, Tamil Nadu
jerinebell187@gmail.com, sugumar.ssd@gmail.com

Abstract. This paper describes an efficient Blind Source Separation of speech and music, speech and music which are considered as convolutional mixtures. The convolutional mixed signals consist of source signals and same amount of delay or echo of the same source signal. Convolutional BSS of stereo mixtures is the challenging task in the audio signal processing application. BSS is a technique for estimating original source signal from their mixtures of signals. The mixed signals were decomposed by 1D multilevel discrete wavelet decomposition. Decomposition levels are changed and the signals to noise ratio (SNR) are calculated. After that ICA has been performed and the sources are separated.

Keywords: Independent Component Analysis, Blind Source Separation, Convolutional mixtures, Multilevel Wavelet decomposition.

1 Introduction

Blind Source Separation is the separation of a set of signals from a set of mixed signals without the aid of information (or with very little information) about the source signals or the mixing process [1]. BSS thus separates a set of signals into a set of original signals, such that the independency of each resulting signal is maximized, and the dependency between the signals is minimized [3]. Independent component analysis in time domain holds good for speech. It will produce better results for speech signals. Independent component analysis (ICA) is a computational method for separating a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-gaussian source signals [5]. It is a special case of blind source separation. Blind source separation refers to two class of multichannel signal processing tasks in which the goal is to extract multiple useful signals from the multiple convolutional mixtures of these signals without specific knowledge of the source properties or the mixing characteristics [6]. Convolutional BSS assumes a general multipath channel thus requiring multichannel filtering.

2 Convolutional Mixing Model

The source signals are called convolved mixtures since acoustic signals recorded simultaneously in a reverberant environment can be described as the sums of differently convolved sources. In the testing process Mixture2 is taken and it is plotted below.

$$\mathbf{x}(n) = \sum_{p=0}^P \mathbf{A}(p)\mathbf{s}(n - p) + \mathbf{n}(n) \tag{1}$$

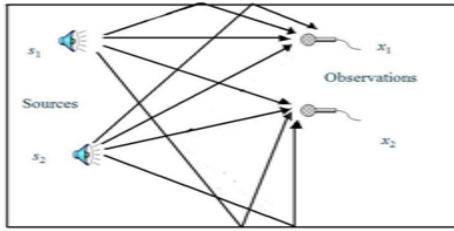


Fig. 1. Convolutive Mixing Model

3 Proposed Algorithm

Wavelet decomposition is applied on the mixed signal (speech+music). Multilevel 1D wavelet decomposition is used to decompose the mixed signal. It will produce the approximated & detailed signals. After that independent component analysis is applied on the decomposed signals. ICA effectively separates the mixed signals. Finally by applying wavelet reconstruction on the ICA Separated signals the mixed sources are separated as speech, music.

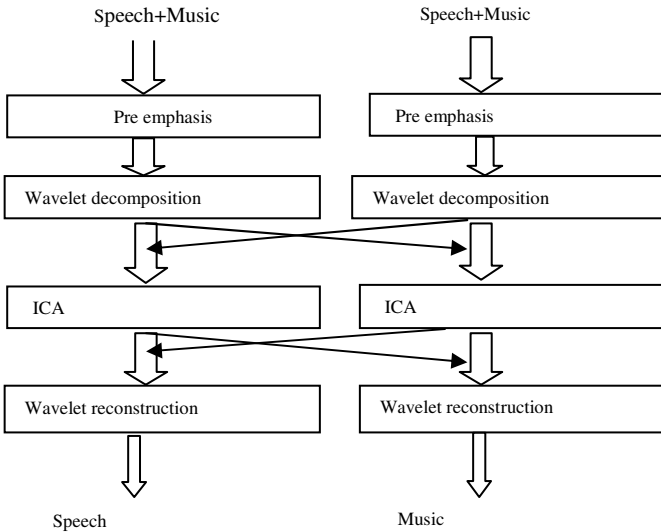


Fig. 2. Flow diagram

4 Independent Component Analysis

The independent component analysis allows two source signals to be separated from two mixed signals using statistical principles of independent and non-gaussianity [4].

The algorithm requires that there be as many sensors as input signals. For example, with three independent sources and three mixtures being recorded, the problem could be modelled as:

$$x_1(t) = as_1(t) + bs_2(t) \quad (2)$$

$$x_2(t) = cs_1(t) + ds_2(t) \quad (3)$$

5 Results and Discussions

Table 1. Snr Measurements For Mixture1

Wavelet Decomposition Levels	SNR	SNR1
4	21.7932	7.1472
5	22.5731	7.0732
6	22.700	7.1951
7	26.9025	11.2425

Table 2. Snr Measurements For Mixture 2

Wavelet Decomposition Levels	SNR	SNR1
4	21.1932	3.5927
5	21.2475	3.5923
6	21.3349	4.3686
7	19.2031	7.9077

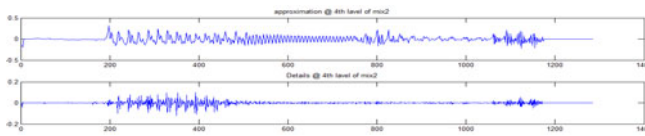


Fig. 3. ICA (approximation & detailed signals)

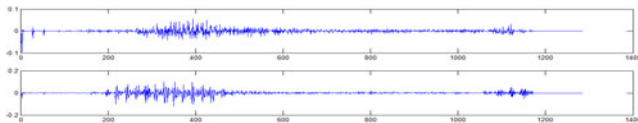


Fig. 4. Reconstructed signal (Speech, Music)

6 Conclusion

The mixture1 and mixture2 are separated using independent component analysis using wavelets decomposition method. Input signals are considered as convolutional mixing signal. The signal strength mainly depends on the levels of decomposition, As

the number of levels increases SNR also increases. The experimental results show that mixtures2 produce better results than the mixture1. For mixture1 (speech) the SNR value increases up to 6 levels after that it start decreasing. The best SNR obtained for mixture1 in the sixth level is 21.3349dB. The maximum SNR obtained in the seventh level is 26.9025dB.

References

1. Ozerov, A., Févotte, C.: Multichannel Nonnegative Matrix Factorization in Convolutional Mixtures for Audio Source Separation. *IEEE Transactions on Audio, Speech, and Language Processing* 18(3) (March 2010)
2. Cobos, M., Lopez, J.J.: Stereo audio source separation based on time–frequency masking and multilevel thresholding. *Elsevier -Digital Signal Processing* 18, 960–976 (2008)
3. Jafari, M.G., Vincent, E., Abdallah, S.A., Plumbley, M.D., Davies, M.E.: An adaptive stereo basis method for convolutional blind audio source separation. *Elsevier Neurocomputing* 71, 2087–2097 (2008)
4. Nesbitt, A., Plumbley, M.D., Davies, M.E.: Audio Source Separation with A Signal-Adaptive Local Cosine Transform. *Elsevier, Signal Processing* 87, 1848–1858 (2007)
5. Douglas, S.C., Gupta, M., Sawada Sr., H., Makino, S.: Spatio–Temporal FastICA Algorithms for the Blind Separation of Convolutional Mixtures. *IEEE Transactions on Audio, Speech, and Language Processing* 15(5) (July 2007)
6. Mei, T., Xi, J., Yin, F., Mertins, A., Chicharo, J.F.: Blind Source Separation Based on Time-Domain Optimization of a Frequency-Domain Independence Criterion. *IEEE Transactions on Audio, Speech, and Language Processing* 14(6) (November 2006)
7. Vincent, E., Gribonval, R., Févotte, C.: Performance Measurement in Blind Audio Source Separation. *IEEE Transactions on Audio, Speech, and Language Processing* 14(4) (July 2006)
8. Ding, S., Cichocki, A., Huang, J., Wei, D.: Blind Source Separation of Acoustic Signals in Realistic Environments Based on ICA in the Time-Frequency Domain. *Journal of Pervasive Computing and Communications* 1(2) (June 2005)
9. Das, N., Routray, A., Dash, P.K.: ICA Methods for Blind Source Separation of Instantaneous Mixtures: A Case Study. *Neural Information Processing – Letters and Reviews* 11(11) (2007)
10. Addison, W., Roberts, S.: *Blind Source Separation with Non-Stationary Mixing Using Wavelets*. The University of Liverpool (2006)
11. Araki, S., Makino, S., Mukai, R., Nishikawa, T., Saruwatari, H.: Fundamental Limitation of Frequency Domain Blind Source Separation for Convolved Mixture of Speech. *NTT Communication Science Laboratories and Nara Institute of Science and Technology, Japan* (2006)

ICA Based Informed Source Separation for Digitally Watermarked Audio Signals

R. Sharanya, D. Sugumar, T.L. Sujithra, Susan Mary Bose, and Divya Mary Koshy

Department of ECE, Karunya University, Coimbatore, Tamil Nadu, India
sariyu.ece@gmail.com, sugumar.ssd@gmail.com,
sujithraleon@gmail.com, susanmarybose@gmail.com,
dvy.koshy@gmail.com

Abstract. This paper presents an efficient, digital audio watermarking method based on one dimensional multilevel discrete wavelet transform(MDWT) and discrete cosine transform (DCT) for the application of copyright protection. The proposed scheme also employs informed source separation for the watermark extracted audio signal based on independent component analysis (ICA).In addition, the performance of the proposed algorithm in terms of Signal to Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR) and Normalized Root Mean Square Error (NRMSE) are also evaluated. The proposed scheme achieved good robustness against most of the attacks such as requantization, filtering, addition and multiplication of noise. The experimental result shows that the SNR value is about 36.12dB for the proposed algorithm, whereas by using only MDWT technique method gives only 30.2dB.

Keywords: Independent Component Analysis, Informed Source Separation, Audio Watermarking, Copyright protection.

1 Introduction

Digital watermarking techniques are considered to be effective solution to the problems of copyrights. Currently, watermarking techniques based on transform domain are more popular than those based on time domain since they provide higher audio quality and more robust. Also in terms of the performance of watermarks against attacks, the transform domain methods are commonly considered better than that of the time and frequency domain methods [1].Additional known audio watermarking classes include spread spectrum [3] and compressed audio techniques, in which the spread spectrum requirements of hiding a signal against an unintended listener and ensuring information privacy, are very similar to those in watermark applications However, computational complexity and synchronization overhead may be unacceptably high.

Blind Source Separation (BSS) method was proposed in [6] to separate a large number of speech sources. When the number of available audio channels (mixtures) equals or exceeds the number of individual sources Independent component analysis (ICA) is used. In this paper, an audio watermarking algorithm that satisfies the requirements of effective audio watermarking inaudibility and watermark robustness has

been proposed. The requirements were met by the proposed algorithm by exploiting the attractive properties of two powerful mathematical transforms; one dimensional Multilevel Discrete Wavelet Transform (MDWT) and Discrete Cosine Transform (DCT). In the proposed algorithm, watermark bits are embedded directly on coefficients of DCT, MDWT is taken and it is reconstructed. The watermark embedding and extraction procedure, source separation is outlined in section two, and experimental results, in section three. Conclusion is given in section four.

2 Proposed Algorithm

The proposed algorithm employs a cascade of two transforms; the discrete wavelet transform and the discrete cosine transform. The algorithm is described in this section by outlining the major steps, the watermark embedding procedure, and watermark extraction procedure [7].

2.1 Watermark Embedding/Extraction and Source Separation Procedure

The procedure is illustrated in the block diagram shown in Fig 1, and described in details in the steps which follow below,

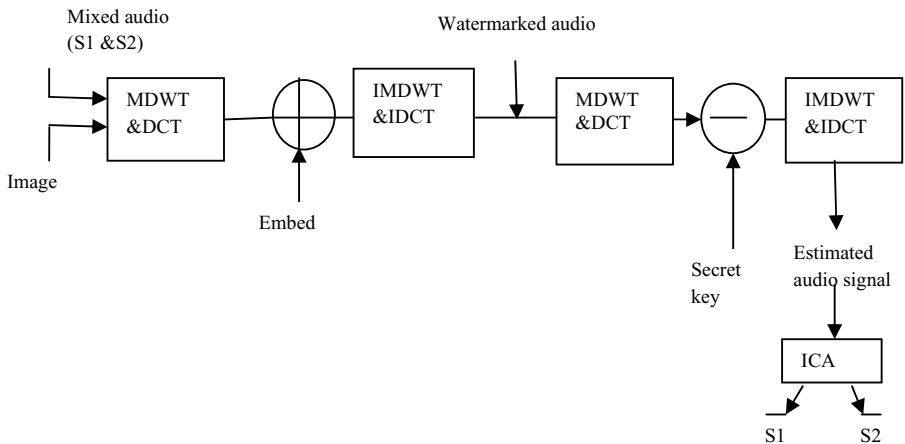


Fig. 1. Watermark Embedding/Extraction, Source Separation Procedure

Two different audio signals (music + speech) of sampling frequency 44100, bit resolution of 16 and PCM uncompressed signals are mixed linearly and image is taken. Perform a MDWT and DCT transformation. Perform inverse operation of the embedded signal by taking IDCT and IMDWT, the obtained signal is the watermarked signal. The watermarked audio signal obtained from the embedding process is taken and MDWT and DCT for the signal are taken. The binary image is then extracted by using a secret key. Assemble the extracted bits from the inverse operation of MDWT and IDCT is taken and the estimated audio signal is obtained. estimated audio signal is obtained by applying ICA and sources are separated [9].

3 Experimental Results and Discussion

In this section, the results obtained using the MDWT-DCT algorithm has been presented. Pop music and speech audio clips were used to evaluate performance of the proposed algorithm. The watermark used in our experiments is the binary image shown in Fig 2. Then performance measures such as SNR, PSNR, and NRMSE are tabulated.



Fig. 2. Detected Watermark Image

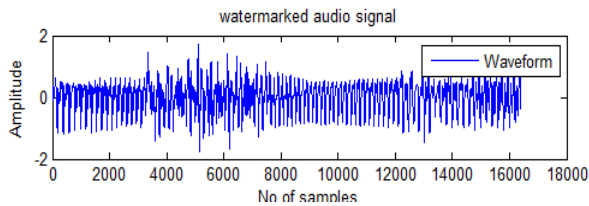


Fig. 3. Watermarked Audio Signal

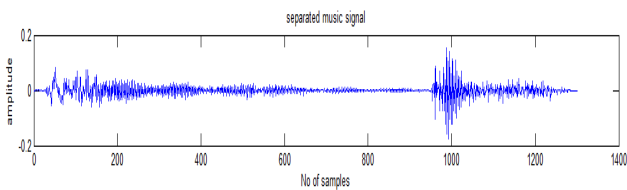


Fig. 4. Separated Music Signal after applying ICA

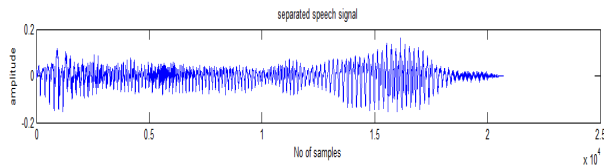


Fig. 5. Separated Speech Signal after applying ICA

Table 1. Performance Measures after Watermark Insertion & watermark Extraction for MDWT-DCT

Performance measures	After Watermark insertion	After Watermark extraction
SNR	36.1	21.6
PSNR	78.2	42.1
NRMSE	12.7	3.4

4 Conclusion

In this paper, proposed an imperceptible (inaudible) and robust audio watermarking technique based on cascading two powerful mathematical transforms; MDWT and the DCT. By virtue of cascading the two transforms, inaudibility and different levels of robustness is achieved. This paper presents that preprocessing is important where SNR measure without pre-processing process gives 0.0014dB which is very low. The simulation results obtained verify the effectiveness of audio watermarking as a reliable solution to the copyright protection. Contrary to the BSS framework, in the Informed Source Separation, source signals are available before the mix is processed. The proposed ICA based detector can be used for wavelet based watermarking for all types of multimedia data, e.g., audio, video, images, etc.

References

1. Cox, I.J., Kilian, J., Shamoon, T.: Secure Spread Spectrum Watermarking for Image, Audio and Video. *IEEE Trans. on Image Processing* 6, 1673–1687 (1997)
2. Khademi, N., Akhaee, M.A., Ahadi: Audio watermarking based on quantization index modulation in frequency domain. *Signal Processing and Communication*, 1127 (2007)
3. Kirovski, D., Malwar, H.: Robust spread spectrum watermarking. In: *Proceedings of Acoustics, speech and signal processing (ICASSP 2001)*. , vol. 3, pp. 1345–1348 (2001)
4. Kaengin, S., Airphaiboon, S., Patthoumvanh, S.: New technique for embedding watermark image into an audio signal. *Communication and Information Technology (ISCIT)* (2009)
5. Yilmaz, O., Rickard, S.: Blind separation of speech mixtures time-frequency masking. *IEEE Trans. Signal Processing* 52, 1830–1847 (2004)
6. Parvaix, M., Girin, L., Brossier, J.M.: A Watermarking based method for informed source separation of audio signals with a single sensor. *IEEE Trans.on Audio, Speech and Language Processing* 18, 1464–1475 (2010)
7. Haj, A.A., Mohammad, A.: Digital audio watermarking based on discrete wavelet transform and singular value decomposition. *European Journal of Scientific Research* 39, 6–21 (2010) ISSN 1450-216X
8. Das, N., Routray, A., Dash, P.K.: ICA methods for blind source separation of Instantaneous Mixtures. *Letters and Reviews* 11 (November 2007)
9. Douglas, S.C., Gupta, M., Sawada, H.: SpatioTemporal fast ICA algorithms for the blind separation of convolutive mixtures. *IEEE Tran. on Audio, Speech and Language Processing* 15, 1511–1520 (2007)

Evaluation of Retrieval System Using Textural Features Based on Wavelet Transform

Lidiya Xavier and I. Thusnavis Bella Mary

School of Electrical Sciences, Karunya University, Coimbatore-641114, India
lidiyaxavier@gmail.com, bellamary@karunya.edu

Abstract. A content based image retrieval system allows the user to present a query image in order to retrieve images stored in the database according to their similarity to the query image. Content based image retrieval method is used as diagnosis aid in medical fields. In this paper content based image retrieval is used for medical applications. The main objective of this paper is to evaluate the retrieval system using Textural features. The texture features are extracted by using wavelet transform. The method is evaluated on Diabetic Retinopathy Database (DRD). Here the precision rate obtained is about 60% for DRD images.

Keywords: Image Retrieval, Content based image retrieval, Pyramid-structure wavelet transform, Texture and Medical image.

1 Introduction

Image Retrieval aims to provide an effective and efficient tool for managing large image databases. Image retrieval (IR) is one of the most exciting and fastest growing research areas in the field of medical imaging [2]. A CBIR method typically converts an image into a feature vector representation and matches with the images in the database to find out the most similar images. From reference paper Mathieu Lamard [1], Aliaa.A.A Yousif [2] and Cazuguel, G. [3] different databases have been used and a comparative study is carried out. From Mathieu Lamard [1] and Cazuguel, G. [3], it is concluded that performance of DRD image is less compared to other database images. In order to improve efficiency of DRD database images pyramid-structured wavelet transform is used. In this paper evaluation of retrieval system based on Textural features is carried out. The major advantage of this approach is that little human intervention is required. The databases used here is Diabetic Retinopathy Database [1] [3].

2 Block Diagram

Figure 1 shows the basic block diagram used in this work. Textural features are extracted for both query image and images in the database. The distance (ie., similarities) between the feature vectors of the query image and database are then computed

and ranked. The database images that have highest similarity to the query image are retrieved. Then the performance analysis is carried out using precision and recall.

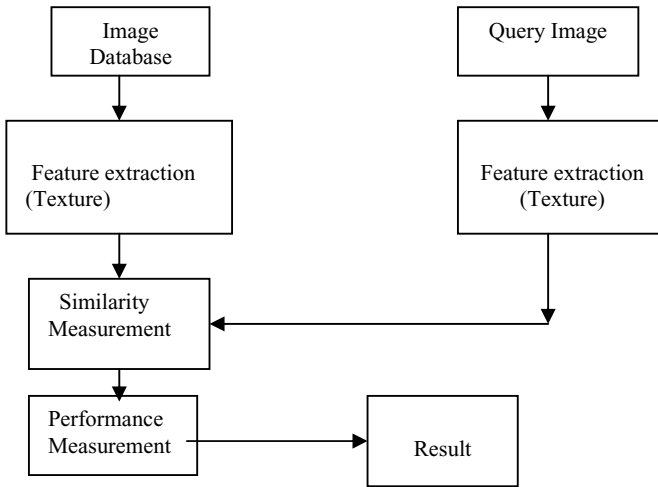


Fig. 1. Basic block diagram

3 Texture Extraction

The property of all surfaces that describes visual patterns, each having properties of homogeneity is termed as texture. Three types of texture features are used in this paper. They are energy, contrast and entropy. Energy is a measure of textural uniformity. Energy is low when all elements are equal and is useful for highlighting geometry and continuity. Here energy is calculated using pyramid-structure wavelet transform, and it will explain in the following sections. Contrast is a measure of the contrast or amount of local variation present in an image or surface. A texture of high contrast has large difference in intensity among neighboring pixels, while a texture of low contrast has small difference. Here contrast is calculated for every image in each decomposition level. Entropy is a measure of disorder or complexity. It is large for surfaces that are texturally not uniform.

3.1 Pyramid-Structured Wavelet Transform

The pyramid-structure wavelet transform indicate that it recursively decomposes sub signals in the low frequency channels [2]. Using the pyramid-structure wavelet transform, the texture image is decomposed into four sub images, as low-low, low-high, high-low and high-high sub-bands. The energy level of each sub-band is calculated. This is first level decomposition. Using the low-low sub-band for further decomposition is done. Decomposition is done up to third level in this paper. The reason for this type of decomposition is the assumption that the energy of an image is concentrated in the low-low band.

3.2 Energy Level and Euclidean Distance

Calculate the energy of all decomposed images at the same scale, using:

$$E = \frac{1}{MN} \sum_{i=1}^m \sum_{j=1}^n |X(i, j)| \tag{1}$$

where M and N are the dimensions of the image, and X is the intensity of the pixel located at row i and column j in the image map. Repeat from step 1 for the low-low sub-band image, until it becomes third level. Using the above algorithm, the energy levels of the sub-bands is calculated [2]. These energy level values are stored to be used in the Euclidean distance algorithm. And the equation is given below. The Euclidean distance D between two vectors X and Y is

$$D = \sqrt{\left(\sum_{i,j} (X - Y)^2\right)} \tag{2}$$

Using the above algorithm, the query image is searched for in the image database. The Euclidean distance is calculated between the query image and every image in the database. This process is repeated until all the images in the database have been compared with the query image.

4 Graphical Analysis

Figure 2(a) shows the graphical representation of precision versus number of retrieved images. From the graph it is understood that to retrieve all the relevant images in the database almost the maximum number is has to be retrieved for all the class. Almost 50% of the relevant image is retrieved for precision at P(3). The precision rate is about 60% when first three images are retrieved from the database. Graphical representation of recall and number of retrieved images is shown in the figure 2(b). The recall rate is high when first three images are retrieved in case of background and profiliative retinopathy compared to non profiliative. Figure 2(c) shows the graphical representation of precision versus recall. Precision value is inversely proportional to recall value. The maintains of high precision value at various levels indicates that majority of relevant images are retrieved at the early stage.

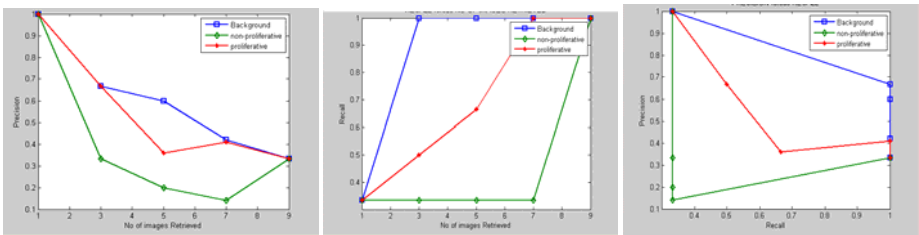


Fig. 2. (a) Precision Vs Number of Retrieved Images,(b) Recall Vs Number of Retrieved Images, (c) Precision Vs Recall

5 Conclusion

In this paper pyramid-structure wavelet transform based method is considered for content based diabetic retinopathy image retrieval. From Fig. 2 it can be concluded that the best precision rate of the 60% and recall rate of 60% is achieved using this method. In paper [1] DRD image is evaluated using adaptive nonseparable lifting scheme which results in poor retrieval rate. Hence by using this technique the retrieval rate for DRD image is improved.

References

1. Quellec, G., Lamard, M., Cazuguel, G., Cochener, B.: Adaptive Non- Separable Wavelet Transform via Lifting and its Application to Content-Based ImageRetrieval. *IEEE Transactions on Image Processing* 19(1), 25–35 (2010)
2. Yousif, A.A.A., Darwish, A.A., Mohammad, R.A.: Content based medical image retrieval based on pyramid structure wavelet. *IJCSNS International Journal of Computer Science and Network Security* 3 (March 2010)
3. Lamard, M., Cazuguel, G., Quellec, G., Bekri, L., Roux, C., Cochener, B.: Content based image retrieval based on wavelet transform coefficients distribution. In: Proc. of the 29th annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (August 2007)
4. Antani, S., Lee, D.J., Long, L.R., Thoma, G.R.: Evaluation of shape similarity measurement methods for spine X-ray images. *J. Vis. Commun. Image R* 15(3), 285–302 (2004)

Behavioural Level Watermarking Techniques for IP Identification Based on Testing in SOC Design

Newton david Raj, Josprakash, AntopremKumar, Daniel, and Joshua Thomas

School of Electrical Science, Karunya University, Coimbatore-641114, India
{newton.davidraj1, jos.prakash.av, anto.donns, danielpon10, joshuaalackal}@gmail.com

Abstract. This paper proposes a watermarking scheme for intellectual property (IP) identification based on testing in soc design. The core concept is embedding the watermarking generating circuit (WGC) and test circuit (TC) in to the soft IP core at the behavioural design level. Therefore this scheme can successfully survive synthesis, placement and routing and can identify the IP core at various design levels and IP core does not change at any levels. This method adopts current main system-on-chip (SOC). The identity of the IP is proven during the general test process without implementing any extra extraction flow. After the chip has been manufactured and packaged, it is still easy to detect the identification of the IP provider without the need of microphotograph. This approaches entail low hardware overhead, tracking costs, and processing-time costs. The proposed method solves the IP-identification problem.

Keywords: intellectual-property (IP), system-on-a-chip (SOC), very large scale integration (VLSI) design, watermarking generating circuit (WGC), test circuit (TC).

1 Introduction

Advances in semiconductor processing technology have led to the rapid increases in integrated-circuit (IC) design complexity [1],[2].The shift toward very deep sub micrometer processing technology has encouraged IC designers to design an entire system implemented on a single chip. This new paradigm, called the system-on-a-chip (SOC), has changed design methodologies. In order to reduce time to market and to increase productivity, the reuse of previously designed modules is becoming a common practice.

Design rules lead to the development of IP-identification techniques. Each IP should have identification that represents the design information, including designer identity, version, ownership rights, and provider. The identification can also provide designer information, IP tracking, ownership proof, and IP management. The ability to prove the identity of virtual components is increasing in importance. After the IP has been integrated into a whole chip and packaged, designers can still check the identity of the IP. This paper deals the design for SOC using watermark techniques for IP identification.

2 IP-Based Design Flow with Watermarking

In this section, develop an IP-identification approach using the watermarking technique. This method is developed depending on the current IP-based design flow [1]. The explanation will describe the design process of headed watermarking-sequence method.

2.1 Watermark Design

First of all, the watermark, which can intuitively represent one's identity, it is generated as a binary sequence and inserted in to each IP core. Propose a coding technique for the design of the digital watermark. The watermark is a symbol that stands for the organization's title, a laboratory's mark, or a personal name, and it is comprised of a sequence of bits. For example, the symbol "CEDECEKUC" to represent College of Engineering, Department of Electronics and Communication Engineering, Karunya University Coimbatore. It describes the symbol according to the coded table that constructed beforehand. 00010 to represent letter C, 00100 to represent E, 00011 to represent D, 00100 to represent E, 00010 to represent C, 00100 to represent E, 01010 to represent K, 10100 to represent U, 00010 to represent C. With this method of encoding, just 45 bits are needed to describe "CEDECEKUC".

2.2 WGC Design

The WGC is composed of several parallel-input–serial output (PISO) registers and inverter gates. When the test-mode signal is active (test mode = 1), the WGC will be turned on. The parallel watermark data are generated by the inverters. If the watermark value is one, the circuit directly generates the value. If the watermark value is zero, there is an inverter that translates the test-mode signal into zero. The watermark data are generated via the test-mode signal and inverters. The PISO translates the parallel watermark data into a sequence.

3 Combining TC with WGC

After the WGC has been designed, combine the TC with the WGC. How the TC is combined with the WGC is very important. two methods for combining the TC with the WGC, and we analyse the characteristics of each method of combining test circuit with wgc circuit.

3.1 Headed Watermark-Sequence Method

When the chip is in the test mode, the chip sends out first the watermark sequence. After sending out the entire watermark sequence, the chip sends the output test patterns. The watermark sequence is like the header of a bit stream. This method enables the watermark to simply be extracted. Drawback is that the watermark is easy to remove

4 Experimental Results

The IP cores were designed using Verilog HDL and were verified. The watermark function is not changed after logic synthesis because embed the watermark into the TC at the behavioural design level. After placement and routing, still detect the identity, according to the watermark sequence, without error. According to the results, the proposed method can identify the soft IP core at the behavioural, gate, and physical design levels.



Fig. 3. (a) WGCResponse (when clock=1, reset=0 test mode=1 and ready signal =1). (b) Watermark sequence and test circuit response

5 Conclusion

The watermark is a general-purpose design methodology that does not need to be designed case by case according to various IPs. The watermark function is not changed after logic synthesis, placement, and routing because the watermark is embedded into the TC at the behavioural design level. The approaches have the ability to detect the presence of the watermark and to identify the soft IP core at various design levels.

References

- [1] Fan, Y.-C.: Testing-Based Watermarking Techniques for Intellectual Property Identification in SOC Design. *IEEE Trans. Instrumentation and measurement* 57(3) (March 2008)
- [2] Martin, G., Chang, H.: *Winning the SoC Revolution: Experiences in Real Design*. Kluwer, Norwell (2003)
- [3] Cox, I.J., Miller, M.L., Bloom, J.A.: *Digital Watermarking*. Morgan Kaufmann, San Mateo (2002)
- [4] Narayan, N., Newbould, R.D., Carothers, J.D., Rodriguez, J.J., Holman, W.T.: IP protection for VLSI designs via watermarking of routes. In: *Proc. IEEE Int. Conf. ASIC/SOC*, September 2001, pp. 406–410 (2001)

Rough Set Approach for Distributed Decision Tree and Attribute Reduction in the Disseminated Environment

E. Chandra¹ and P. Ajitha²

¹ Research Supervisor & Director, Department Of Computer Science, D J Academy for Managerial Excellence, Coimbatore, Tamilnadu, India

² Research Scholar & Assistant Professor, Department Of Computer Science, D J Academy for Managerial Excellence, Coimbatore, Tamilnadu, India

ajitha@y7mail.com

Abstract. Attribute reduction is a necessitated step for the disseminated environment in regard to classification and prediction of the data. Traditional approaches were not efficient for optimal attribute reduces. Current techniques are quiet time consuming and less accuracy. Rough set approach is a mathematical technique to handle attribute reduces through data dependencies and structural methods. This paper discusses a novel algorithm for optimal attribute deduct and also increases the accuracy in predicted results and the distributed decision tree classification techniques was made use of to implement the same in the disseminated environment. Proposed algorithm for Construction of distributed decision trees with rough sets increases the accuracy and also reduces the attributes on the time of the massive data sets handling.

Keywords: Rough Sets, Distributed Decision Trees, Distributed data mining.

1 Introduction

Distributed Data Mining mines massive data sets and also heterogeneous type of data. An appropriate mathematical model is necessary to deal with the data especially of heterogeneous types where there was large number of attributes. Individual analysis of each site is different and inadequate for some special treatments [1] and also when there was uncertain, vague and imprecision data handling of them was very difficult. Rough sets, one of the mathematical model is dealt with these category of data.

Rough set was proposed by Pawalak [4], one of the most popular techniques applied in machine learning, data mining[3], pattern recognition etc. It provided the strategies to discover the data dependences and used structural methods. This paper proposes a rough sets distributed decision tree algorithm which achieves simultaneous prototype selection and feature selection, and a scheme to use the rough sets for decision tree for efficient classification in a disseminated environment.

2 Proposed Computation of Attribute Reduction

To compute a_{ij} the attribute reductions, construct the attribute subset enumeration tree by merely using the non-core attributes. Already Distributed Decision trees[9]

constructions with roughs sets was proposed in the literature survey so it was not discussed here again.

2.1 Computation Method

Sequential Attribute Reduction Algorithm SARA, similar to the algorithm put forward by Zhang [6], and the distributed attribute reduction algorithm DARA based on peer-to-peer technique both for the client and server was discussed here. The maximum length of queue Q or T should be $\frac{\perp(C-Core(C))/2}{C-U-Core(C)}$ [9]. when the

computation is amortized on computers in a peer-to-peer network to apportion Q and T, the reduction of larger data set with more attributes may be carried out efficiently as the T computation was calculated. Newly proposed Distributed Attribute Reduction Algorithm DARA is composed of the server DARA-S and the client DARA-C. Because of page constraint DARA-S is specified here with the small elimination of integration DARA-S is taken as such

Input: U , C, and D of consistent decision system
Output: all attribute reduction set R
 (1) compute Core(C), send M('C', Core(C)) to all N client computers;
 (2) for all a ∈ C -Core(C), send M('Q', {a}) to client computer C[i] in round robin way to the distributed evt and $N_Q = |A|, N_T = 0, R = \emptyset$; (3) send M ('S ') to all clients to start checking and computing the attribute reductions;(4) when receive M('R', q), put $q \rightarrow R$;(5) when receive M('T', q), set $N_T = N_T + 1$ and $q \rightarrow R$ with N_Q and N_T
 (6) when receive M ('N') , set $N_Q = N_Q - 1$; (6.1) if $N_Q = 0$ and $N_T = 0$, stop; if $N_Q = 0$, send M('P') to all clients;(7) when receive M ('Q', t) , set $N_Q = N_Q + 1$ and $N_T = N_T + 1$ and $q \rightarrow R$ with N_Q and N_T and (8) when receive M ('F') , set $N_T = N_T - 1$; then if $N_T = 0$ and $N_Q = 0$, stop; check if $N_T = 0$ goto (3). (9) send these N_Q and N_T to all sites

Algorithm 1. DARA-S: Distributed Attribute Reduction on Server

Proposed algorithm DARA-S reduces the attributes when dealing in the distributed architecture. When the computational complexity of the algorithm is defined $O(n+m)$ and $M(\log_2(n^2))O(n)$. Here M and n represents the attributes that is sent over across the sites. Computational memory that is taken is very less in the proposed algorithms of DARA-S and DARA-C.

3 Classification Based on Rough Sets

When the classification technique of Distributed Decision Tree is considered a mathematical model is necessary to produce an optimize results. For this greater purpose an algorithm was proposed so that heterogeneous data sets can be classified with greater accuracy and in minimised time in compared with the other existing methods[10].

3.1 Proposed Algorithm for Rough Sets in Distributed Decision Tree

The following is the proposed algorithm for classification based on rough sets. CT-tree[8][10] proposes algorithm for attribute reduction here both for the attribute reduction, time and accuracy was calculated using rough set theory (rCTI) –Classification Tree based on rough sets

Inputs: T_S -Test patterns, T_R -Training patterns, s -User defined threshold.
Outputs: C_{time} —Classification time, C_A —Classification accuracy.
 Steps: (1) Generate CT-tree using T_R and initialize start-time =time().
 (2) For each branch b_i of the DDT-tree, find lb_a, ub_a, cb_a (3) For each $s_j \in TS$ (a) Find n_j , set of positions of non-zero values corresponding to s_j .(b) For each branch d_k in rCT-tree, if d_k corresponds to u_{bk} then $n_j = n_j - C_{bk}$ else $n_j = n_j - U_{bk}$ (c) Find the nearest neighbour branch, ep in rPC-tree depending on maximum number of features which are common to both ep and n_j . (d) Attach the label l associated with ep to n_j (e) If $(l == label\ of\ s_j)$ then Correct= correct + 1.(4) end-time = time(). (5) $CA = (correct / |T_S| \times 100 | T_S|)$ is the number of test patterns. (6)Output $C_{time} = end-time - start-time$ and Output CA.(7) check CA and C_{time} as scheduled (7) DARA –S and DARA-C recursive call

Algorithm 2. Classification based on rough sets

lb_a =set of features from root to leaf along b_a, ub_a = set of features from root to leaf along b_a where ‘Count’ field values of the nodes is greater than or equal to s . assign $Cb_a = lb_a - ub_a$. (i) ub_a —if b_a is shared by the pattern of same class; (ii) cb_a —otherwise. Using all these attributes are selected based on some threshold and time factor is checked by taking into account the value of nodes on that specified criteria’s. The proposed algorithm deals with these aspects paved way to the following results.

Table 1. Comparison of CT-DDT and rCTI

Algorithm	Classification accuracy (%)	Classification time (in s)	Storage space(in Bytes)
CT	92.5	1713	1,302,454
rCTI	93.22	1207	1,291,347

From this table1 that rCTI gives the best classification accuracy and computational space. On comparing with the other methods the proposed algorithm specifies the 87% of accuracy in compared with the others as specified in the table and as in simulated results.

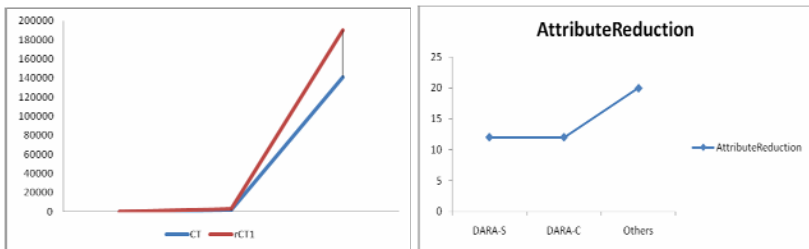


Fig. 1. Proposed Simulation results of rCTI and attribute reduction by using proposed algorithm

4 Conclusions and Future Work

When the data sets are large and need to be transmitted over the different sites chances for less accuracy and more time because of large selection of attributes. As Rough sets and indiscernibility functions was made use of the classification accuracy is maintained and also attribute reductions which lead to the computational storage is reduced as it is also maintained by reducts on decision trees. As decision trees are quite efficient for classification, this paper deals with the decision trees in distributed environment. rCT1 tree scans the databases only once. Segmentation of data can be discussed further and also association rules can be integrated for the pruning of trees, when the outliers or over fitting occurs in the distributed decision trees.

References

1. Chitcharoen, D., Pattaraintakorn, P.: Novel matrix forms of rough set flow graphs with applications to data integration. *Computers and Mathematics with Applications* 60, 2880–2897 (2010)
2. Guo, Q.L., Zhang, M.: Implement web learning environment based on data mining. *Knowledge-Based Systems* 22(6), 439–442 (2009)
3. Cios, K., Pedrycz, W., Swiniarski, R.: *Data Mining Methods for Knowledge Discovery*. Kluwer, Norwell
4. Pawlak, Z., Skowron, A.: *Rudiments of rough sets*. *Information Sciences*
5. Zhang, E.M., Song, G.Z., Ma, W., Zhang, W.: Attribute Reduction Algorithm Research Based on Rough Core and Back Elimination. In: *Proceedings of the 9th International Conference for Young Computer Scientists (ICYCS 2008)*, pp. 1624–1628. Central South University, China (2008)
6. Stanczyk, U.: On Construction of Optimised Rough Set-based Classifier. *International Journal of Mathematical Models and Methods in Applied Sciences* 2(4) (2008)
7. Sikder, I.U., Munakata, T.: Application of rough set and decision tree for characterization of premonitory factors of low seismic activity. *Expert Systems with Applications* 36, 102–110 (2009)
8. Chen, Y., et al.: A rough set approach to feature selection based on power set tree. *Knowl. Based Syst.* (2010)
9. Ma, G., Lu, Y., Wen, P., Song, E.: A novel attribute reduction algorithm based on peer-to-peer technique and rough set theory. In: *IEEE/ICME International Conference on Complex Medical Engineering* (2010)
10. Mi, J.S., Wu, W.Z., Zhang, W.X.: Approaches to knowledge reduction based on variable precision rough set model. *Information Sciences* 159(3-4), 255–272 (2004)

MIMO and Smart Antenna Technologies for 3G and 4G

Vanitha Rani Rentapalli¹ and Zafer Jawed Khan²

¹ MISTE, Associate Member of IEEE (No. 90410990) Assistant Professor,
ECE Department, Vivekananda Institute of Technology and Science, Karimnagar,
Andhra Pradesh

vanitharani@aol.in

² Vivekananda Institute of Technology and Science, Karimnagar, Andhra Pradesh
khanzjl@rediffmail.com

Abstract. Evolution of wireless access technologies is about to reach its fourth generation (4G). The adaptation of smart antenna techniques in future wireless systems is expected to have a significant impact on the efficient use of the spectrum and transparent operation across the multi technology wireless networks. With the rapid growth of wireless data traffic operators are anxious to quickly expand the capacity of their wireless networks. To address these 3GPP standards have incorporated powerful techniques for using so-called smart antennas. This paper focuses on the practical aspects of deploying smart antenna systems in the existing Radio Access Systems (RAS). Smart antenna techniques, such as multiple-input multiple-output (MIMO) systems, can extend the capabilities of 3G and 4G systems to provide customers with increased data throughput for mobile high-speed data applications.

Keywords: Adaptive array beam forming, MIMO, smart antenna, 3G, 4G.

1 Introduction

Mobile radio communications are evolving from pure telephony systems to multimedia platforms offering a variety of services ranging from simple file transfers, audio and video streaming to interactive applications and positioning tasks. These services have different constraints concerning data rate, delay, and reliability. Hence future mobile radio systems have to provide a large flexibility, and scalability to match these heterogeneous requirements. The basic concept of MIMO (Multi-Input Multi-Output) is that the transmitted signals from all transmit antennas are combined at each receiving antenna element in such a way to improve the Bit Error Rate (BER) performance or the data rate of the transmission. Smart antennas, here are referred to adaptive antennas with electrical tilt, beam width and azimuth control, which can follow relatively slow-varying traffic patterns that can form beams aimed at particular users or steer nulls to reduce interference.

Smart antenna technology is being considered for mobile platforms such as automobiles, cellular telephones and laptops. Operational experience with advanced network systems that have utilized smart antennas for several years provides practical experience applicable to the upgrade to HSPA (High Speed Packet Access) and LTE (Long Term Evolution) wireless systems. Most 3G systems are arranged to operate in 2 GHz frequency band. The 4G represents the next development stage of cellular

evolution beyond 3G, and offers an ideal basis and bandwidth to provide more efficient cellular multicast services. To increase capacity, operators need to add cell sites. Doubling the number of cell sites approximately doubles the network capacity and the throughput per user and greatly improves the peak user and the aggregate throughput. [1-4]

In this paper an outline of various smart antenna schemes for improving the capacity and coverage of the emerging generations of wireless networks is given in addition to downlink transmission modes in relation to practical antenna configurations. Smart antennas with MIMO systems which improve Quality of Service (QoS), capacity & link reliability, and spatial multiplexing are described in section 2. Sections 3 & 4 give the concept of minimizing interfering signals and enhancing signal quality by using adaptive arrays and reconfigurable techniques. Performance evaluation is discussed in section 5, simulation results have been shown in 6, and finally section 7 provides conclusion for this paper.

2 Antennas for MIMO

MIMO can offer more capacity by adding more antennas and more sectors. To meet the data rate and QoS requirements of future broadband cellular systems, their spectral efficiency and link reliability should be considerably improved, which cannot be realized by using traditional single-antenna communication techniques. To achieve these goals, MIMO systems deploy multiple antennas at both ends of the wireless link, exploit the extra spatial dimension, besides the time, frequency, and code dimensions, which allows to significantly increase the spectral efficiency and improve the link reliability.

Array processing techniques are expected to offer the capability of providing significantly reduced costs per transmitted bit in multi-service, multi technology networks. Two main categories of antenna array processing techniques are, beam forming for high element correlation and spatial multiplexing for low element correlation environments. Beam forming techniques can be used to reduce the total transmitted power while preserving the data rate. This in turn reduces the overall system interference . With these advanced technologies, the capabilities of 3G and 4G systems can be extended to provide customers with increased data throughput for mobile high speed data applications.

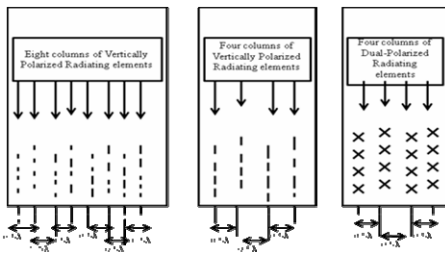


Fig. 1. Multi column planar Array Architecture

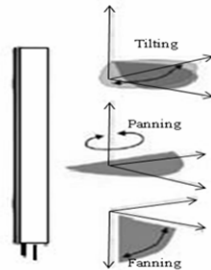


Fig. 2. Reconfigurable beam antenna

2.1 MIMO Spatial Multiplexing

Spatial multiplexing allows transmitting different streams of data simultaneously on the same resource block(s) by taking maximum use of the spatial dimension of the radio channel. These data streams can belong to one single user (single user MIMO / SU-MIMO) or to different users (multi user MIMO / MU-MIMO). Each receiving antenna may receive the data streams from all transmit antennas, the number of data streams that can be transmitted in parallel over the MIMO channel is given by $\min\{N_t, N_r\}$ is limited by the rank of the matrix H . The capacity of MIMO system scales linearly with receiving antennas at low SNR (Signal to Noise Ratio) and at high SNR. Hence spatial multiplexing is used in LTE to achieve the peak data rates and it also provides improvement in cell capacity and throughput. [5]

2.2 Transmit Diversity in LTE

Transmit diversity provides a source of diversity for averaging out the channel variation either for operation at higher (User Equipment) UE speeds (300Kmph) or for delay sensitive services at both low (15 Kmph) and medium (15-120Kmph) UE speeds. The increase in the number of parallel channels translates into an increase in the achievable data rate within the same bandwidth. This increases the signal to noise ratio at the receiver side. Instead of increasing data rate or capacity, MIMO can be used to exploit diversity and increase the robustness of data transmission.

3 Adaptive Array Beam Forming

Reliable and high performance transmission continues to be a major goal of wireless communication systems, which is significantly enhanced by arrays employing beam forming and diversity techniques. An adaptive beam forming multi-column array antenna can be considered as an advanced multiple antenna technique that will provide an improvement in the overall communication link between the base station and mobile. The basic architecture of an adaptive beam-forming antenna consists of multiple columns of radiating elements that are driven by separate transceiver networks. [7] The adaptive systems are really intelligent in the true sense and can actually be referred to as smart antennas. The smartness in these systems comes from the intelligent digital processor that is incorporated in the system which can adjust or adapt its own beam pattern in order to emphasize signals of interest and to minimize interfering signals.

The beam forming array for the downlink case is to increase the signal strength in the desired direction while reducing interference to the undesired directions. Whereas for the uplink case, this array is to improve the receiver sensitivity in the direction of the desired signal while reducing interference from the undesired directions if possible. For the beam forming antenna shown in figure 1 with 0.5 wavelength column separation, the signals between adjacent columns would be highly correlated in low angle spread environments. The two-column antenna with column separation on the order of 1.2 wavelengths would be suitable for a 4X4 MIMO scheme. In this the signals between any two sets of cross polarized ports would not be highly correlated even in low angle spread environments. [6]

Beam forming quality depends on the accuracy of the amplitude and phase values of each MIMO transceiver. Due to undesirable variations between each transmit and receive path, some degree of errors will be formed in multi-column beam forming antenna systems, resulting in significant beam forming degradation. This degraded pattern gives undesirable side lobe levels shown in figure 5, squinting of the main beam, degradation in gains as well as losing the ability to accurately position nulls. By using some type of calibration networks, beam forming antennas can minimize these errors. Typical beam forming systems deployed today require the amplitude variations be limited to ± 0.5 dB, while phase variations are limited to more than ± 5 degrees.

4 Reconfigurable Antennas

A modern telecommunication system which uses re-configurable antennas has the ability to radiate more than one pattern at different frequencies. In a continuously changing environment, re-configurable and adaptive techniques are used for adjusting the structure and parameters of the transceivers to allow them to demonstrate the best performance in a variety of the particular situations. MIMO receivers are capable of reconfiguring themselves by switching automatically between a beam forming and a spatial multiplexing. Reconfigurable beam antennas extend the range of remote beam changes from a single dimension for elevation beam steering (Remote Electrical Tilt, RET) to multiple dimensions. These antennas include the possibility to change the bore sight or azimuth direction (panning), as well as the beam width of the antenna (fanning) remotely. Reconfigurable beam antennas with tilting, panning and fanning as shown in figure 2, can help balance the load between different cells, leading to a combination of coverage, interference, and capacity improvements. Reconfigurable beam antennas can thus significantly increase the basic network coverage.

5 Performance Evaluation

Spectral efficiency measures the ability of a wireless system to deliver a given amount of billable services in a given amount of radio spectrum. The performance evaluation of LTE (E-UTRA) and HSPA (UTRA) are being discussed. Simulation parameters and assumptions follow guidelines provided in [8]. Case 1 represents interference-limited small urban macro cell environment having carrier frequency 2GHz, the inter-site distance 500 m, the bandwidth 10 MHz and UE speed 3Km/h. Case 2 represents an inter-site distance of 1732 m. The cell radii for cases 1 and 2 are 288.7 and 1000m respectively. Uplink performance for case 2 is slightly worse than for case 1 due to the larger cell radius and hence coverage is limited for users towards the cell edge. Uplink spectral efficiency in the LTE system provides more than two times improvement related to HSPA baseline system. In the down link, LTE uses 2×2 SU-MIMO which provides more than three times improvement in spectral efficiency over 1×2 baseline HSPA system. LTE downlink, 4×2 , 4×4 MIMOs provides 10% and 58% spectral efficiency gain related to 2×2 MIMO. It is also observed from the results that the MIMO technologies introduced in LTE improves the cell-edge throughput and spectrum efficiency.

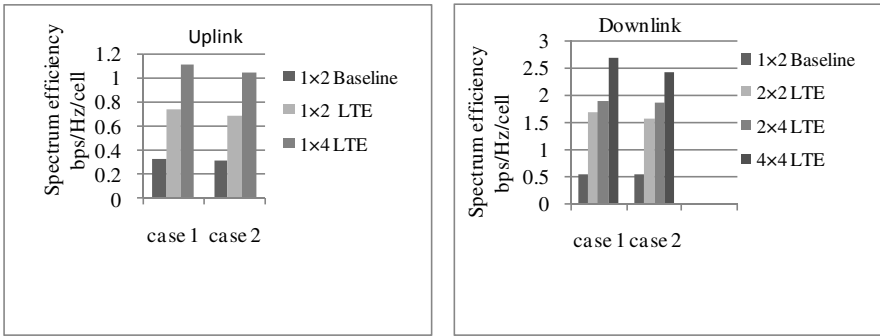


Fig. 3. Uplink and Downlink spectral efficiency

6 Results and Discussion

Simulation for eight element antenna array is performed on Matlab 7.0, using input user signal at 0 degrees and three interferers at -60, -30 and 60 degrees. Calculating an array factor for the array from -180 to 180 degrees, the response of every input user signal and interfering signal is shown in figure 4. The number of elements in the beam

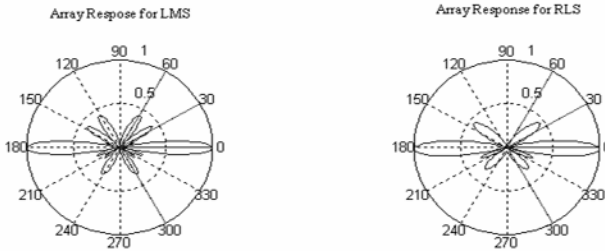


Fig. 4. Beam Pattern, User angle , and interference at angles 60,30,60

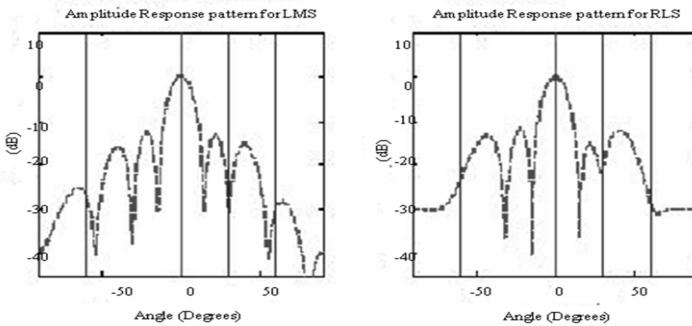


Fig. 5. Amplitude Responses, User angle at 0, Interferences at -60, 30, 60

forming array affect the complexity of the beam forming patterns. The eight column antenna array demonstrates the ability to generate sharper high resolution beam with more nulls due to selection of larger array of antenna elements. The response of major lobe formation is obtained by using Least Mean Square and Recursive Mean Square algorithms. This response has maximum signal strength in user direction as shown in figure 5. An eight column array would result in a larger overall antenna, require eight transreceivers and resulting in a significant increase in cost and complexity.

7 Conclusion

In this paper an overview of the benefits of most recent advances in smart antenna transceiver architectures is given. The array processing techniques have been taken into account to provide the basis for scalable, high data rate, high capacity system solutions and spatial multiplexing gain on the receiving as well as the transmit sides. A variety of smart antennas with remote controlled bore sight to beam-steering arrays and beam forming MIMO have been reviewed. In this context the major trends in the area of smart antennas, such as reconfigurability to varying channel propagation, antennas for MIMO with diversity techniques have been discussed. The multi antennas are used in a beam-forming transmission scheme to improve overall reception quality, increase system capacity and extended coverage.

References

1. Tse, D., Viswanath, P.: Fundamentals of wireless communication. Cambridge University Press, Cambridge (2005)
2. Khan, F.: LTE for 4G mobile broadband. Cambridge University Press, Cambridge (2009)
3. Peng, M., Wang, W.: Technologies and Standards for TD-SDMA Evolutions to IMT- Advanced. IEEE Communications Magazine 47(12), 50–57 (2009)
4. Greenspan, A., Klerer, M., Tomcik, J., Canchi, R., Wilson, J.: IEEE 802.20: Mobile Broadband Wireless Access for the 21st century. IEEE Communications Magazine 46(7), 58–59 (2008)
5. 3GPP, TS 36.211, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation (Release 8)
6. 3GPP, TS 36.212, Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding (Release 8)
7. Parkvall, S., Astely, D.: The Evolution of LTE towards IMT-Advanced Journal of Communications 4(3) (April 2009)
8. 3GPP RAN WG1 TR 25.814 V7.1.0, Physical Layer Aspects for Evolved UTRA (Release 7)

A GA-Artificial Neural Network Hybrid System for Financial Time Series Forecasting

Binoy B. Nair, S. Gnana Sai, A.N. Naveen, A. Lakshmi,
G.S. Venkatesh, and V.P. Mohandas

Amrita Vishwa Vidyapeetham, P.O Ettimadai, Coimbatore, Tamilnadu, India, PIN-641105
b_binoy@cb.amrita.edu, sai.gnana@gmail.com,
naveen4053@gmail.com, lareekath@yahoo.in, gsvenky89@gmail.com,
vp_mohandas@amrita.edu

Abstract. Accurate prediction of financial time series, such as those generated by stock markets, is a highly challenging task due to the highly nonlinear nature of such series. A novel method of predicting the next day's closing value of a stock market is proposed and empirically validated in the present study. The system uses an adaptive artificial neural network based system to predict the next day's closing value of a stock market index. The proposed system adapts itself to the changing market dynamics with the help of genetic algorithm which tunes the parameters of the neural network at the end of each trading session so that best possible accuracy is obtained. The effectiveness of the proposed system is established by testing on five international stock indices using ten different performance measures.

Keywords: Genetic algorithm, artificial neural networks, financial, time series.

1 Introduction

Prediction of financial time series is an extremely challenging problem mainly due to the fact that these series are inherently nonlinear in nature. Especially, in case of stock markets, the task becomes still more difficult as the dynamics that govern the market behavior are very hard to determine and quantify. The behavior of stock markets depend on many factors such as political (eg. general elections, government policies), economic (eg. Economic growth rate, unemployment data), natural factors (eg. natural disasters, rainfall), among many others. However, there have been many studies, for example, in [1] which indicate that stock markets do not follow a random walk, as was suggested by [2] and that, it is in fact, possible to make profits in the stock market. Two commonly used methods of analysis and prediction of stock markets are: (a) fundamental analysis and (b) technical analysis. Fundamental analysis tried to predict the behavior of a stock market based on the analysis of data such as macro-economic indicators, national/international events etc. Technical analysis, on the other hand, considers only the time-series generated by the stock-price/ market movement to arrive at conclusions on the likely future trends.

Soft computing based techniques for predicting stock markets are also now gaining prominence [2]. Comparison of the effectiveness of time delay, recurrent and probabilistic neural networks for prediction of stock trends based on historical data of the daily closing price is done in [3]. Combinations of technical indicators and soft computing techniques have been used in [4] and [5] for predicting of stock exchanges. Artificial neural networks in combination with other soft computing techniques have also been used [6], [7]. In [8], an evaluation method for evolutionary computation based stock trading systems is proposed. [9] presents the application of a novel multiple-kernel learning algorithm for improved stock market forecasting using multiple-kernel support vector regression, while in [10], a multiple objective particle swarm optimization was used to simultaneously optimize two objective functions, namely, the percentage profit and the Sharpe ratio. In [11] and [12], adaptive neuro-fuzzy systems were used to predict the stock markets. However, it has been observed that though soft computing techniques improve the prediction accuracy, selection of appropriate parameters for the prediction algorithm is of critical importance. In the present study, a GA-artificial neural network hybrid system for predicting one-day-ahead closing values in stock markets is proposed. The proposed system can select the optimum algorithm parameters by itself to ensure a high degree of prediction accuracy. The effectiveness of the proposed system is validated on five stock markets, namely, the BSE-Sensex, NSE-Nifty, FTSE-100, DJIA and Nikkei-225.

The rest of the paper is organized as follows: Section 2 presents the design of the proposed hybrid system and section 3 presents the results and the conclusions.

2 Design of the Hybrid System

The stock market prediction system proposed in the present study has three major parts. The first part involves the determination of the optimal lag from the given time series and once the lag is determined, the determination of the optimal embedding dimension. The vectors created from the time series data using the embedding dimension are then fed into the Elman neural network which generates the prediction for the next day's closing. Some of the major issues that confront the design of an artificial neural network (ANN) based system are: determining the number of neurons in the input layer, determining the number of neurons in the hidden layer and the selection of ANN training parameters such as momentum. In the proposed system, the number of input neurons is determined using the embedding dimension of the time series and the number of hidden neurons, the learning parameters of the ANN are optimized using genetic algorithm. The design of the proposed system is presented in the form of a flowchart in Fig.1.

2.1 Determination of Embedding Dimension

The first step in the prediction process, as per the proposed hybrid system, involves the identification of the embedding dimension. For this purpose, the optimum delay value needs to be computed, which is accomplished in the present study using mutual information. The value of delay for which the mutual information first reaches the minimum value is taken as the delay [13]. Once the delay has been calculated, vectors are created from the stock market data with each dimension being represented by the

value at the obtained delay. For example, a three dimensional vector created in this way will be made up of values at the instants: t , $t - (\text{delay})$ and $t - (2 * \text{delay})$. Now the optimum embedding dimension is calculated using the concept of false nearest neighbors [14].

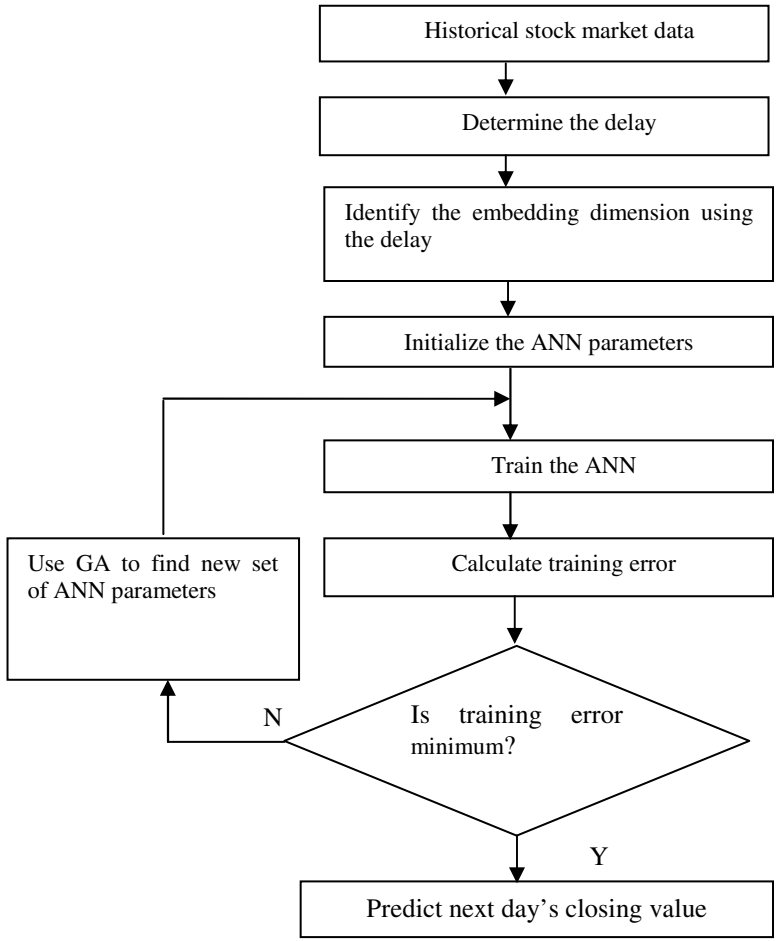


Fig. 1. Proposed hybrid system

2.2 Elman Neural Network Based Prediction

The Elman neural network (ENN) is one kind of global feedforward locally recurrent network model proposed by [15]. An ENN can be considered to be an extension of multilayer perceptron (MLP) with an additional input layer (state layer) that receives a feedback copy of the activations from the hidden layer at the previous time step. These context units in the state layer of the Elman network make it sensitive to the history of input data. In the present study, the ENN is trained using the gradient

descent with momentum backpropagation learning algorithm[16]. In the present study, Elman neural network takes in the input, with the number of input neurons being the same as the embedding dimension. The output of the system is the predicted value for next day's closing. The number of neurons in the hidden layer is optimised using genetic algorithm.

2.3 Genetic Algorithms (GA)

GA has been very widely used for solving unconstrained optimization problems [17]. These are parallel search algorithms which can be used for optimizing nonlinear functions, for example, see [4],[7]. Usually, the function which needs to be optimized, called the objective function, takes the shape of a minimization problem. In the present study, GA is used to minimize the prediction error (which is another way of saying-maximizing accuracy) by optimizing the number of neurons in the hidden layer of the ENN and the momentum parameter. Since the dynamics of stock markets change over time, the number of neurons in the hidden layer is optimized at the end of each trading session.

3 Experimental Results

3.1 Datasets

The proposed system was validated on the following five major stock markets, namely, the Bombay stock exchange sensitive index-India (BSE – Sensex) , The national stock exchange- India(NSE-Nifty), FTSE 100- London , Nikkei 225- Japan and the Dow Jones Industrial average-United States of America (DJIA). The period under consideration was the stock market closing data from January 2005 to September 2010 for all the markets under consideration. First 90% of the data was used for training and the system was validated on the remaining 10%.

3.2 Optimized System Parameters

The hybrid system was trained using the proposed method. The optimum system parameters obtained are presented in Table 1. The embedding dimension gives the number of input neurons needed for the network. The optimal number of hidden neurons and the momentum are obtained using GA.

Table 1. Hybrid system parameters obtained using the proposed system

Training Parameters	BSE-Sensex	NSE-Nifty	FTSE 100	DJIA	Nikkei 225
Hidden neurons	12	18	13	6	7
Embedding dimension	10	38	32	22	19
Delay	26	15	22	13	13
Momentum	0.93	0.98	0.82	0.92	0.90

3.3 Performance Evaluation Measures[18]

The most common measures used to evaluate the performance of forecasting methods are presented in this section. Let Y_t be the observation at time t and F_t denote the forecast of Y_t . Then, the forecast error is $e_t = Y_t - F_t$ and the percentage error is $p_t = 100e_t / Y_t$. Another method is to divide each error, by the error obtained with another standard method of forecasting (in this paper, the standard method or the base method considered is the random walk). Let $r_t = e_t / e_t^*$ denote the relative error, where e_t^* is the forecast error obtained from the base method. Usually, the base method is the “naïve” method or the random walk model. The notation $\text{mean}(x_t)$ is used to denote the sample mean of $\{x_t\}$ over the period of interest (or over the series of interest). Similarly, $\text{median}(x_t)$ is used for the sample median and $\text{gmean}(x_t)$ for the geometric mean. The most commonly used methods are given in Table 2 [18], where the subscript b refers to measures obtained from the base (random walk) method.

The results obtained from the Makridakis competitions (also known as M Competitions, M2-Competitions and M3- competitions) held to test forecasting accuracy of various forecasting methods form the basis of accuracy measures listed in Table 2. In the first M-competition [19], measures used included the MAPE, Mean squared error (MSE), MdAPE, and percentage better (PB). However, the MSE is not appropriate for comparisons between series as it is scale dependent. The MAPE also has problems when the series has values close to (or equal to) zero [20]. MAPE, MdAPE, PB, GMRAE, and MdRAE have also been used as performance measures [21]. The M3-competition [22] used three different measures of accuracy: MdRAE, sMAPE, and sMdAPE. The “symmetric” measures [23] were proposed in response to the observation that the MAPE and MdAPE have the disadvantage that they put a heavier penalty on positive errors than on negative errors. Hence, in the present study, the performance of the proposed system is evaluated using ten different performance measures to ensure the validity of results.

Table 2. Forecast performance measures

Measure	Description	Expression
MAPE	Mean absolute percentage error	$\text{mean}(p_t)$
MdAPE	Median absolute percentage error	$\text{median}(p_t)$
sMAPE	Symmetric mean absolute percentage error	$\text{mean}(2 Y_t - F_t / (Y_t + F_t))$
sMdAPE	Symmetric median absolute percentage error	$\text{median}(2 Y_t - F_t / (Y_t + F_t))$
MRAE	Mean relative absolute error	$\text{mean}(r_t)$
MdRAE	Median relative absolute error	$\text{median}(r_t)$
GMRAE	Geometric mean relative absolute error	$\text{gmean}(r_t)$
RelMAE	Relative mean absolute error	$\text{MAE} / \text{MAE}_b$
RelRMSE	Relative root mean squared error	$\text{RMSE} / \text{RMSE}_b$
LMR	Log mean squared error ratio	$\log(\text{RelRMSE})$

The results obtained using the proposed hybrid system are presented in Table 3.

Table 3. Results for the proposed hybrid system

Performance measure	BSE-Sensex	NSE-Nifty	FTSE 100	DJIA	Nikkei 225
MAPE (%)	8.22	6.17	5.26	3.98	3.86
MdAPE(%)	5.72	4.28	3.92	3.83	3.19
sMAPE (%)	0.08	0.06	0.05	0.04	0.04
sMdAPE (%)	0.08	0.06	0.05	0.04	0.04
MRAE	26.31	24.98	27.35	23.35	10.19
GMRAE	11.25	7.24	7.57	8.14	3.82
Md RAE	10.46	6.28	10.69	7.57	3.80
Rel. MAE	10.22	7.21	7.67	5.16	3.39
Rel. RMSE	9.55	7.32	7.47	4.28	3.30
LMR	4.51	3.98	4.05	2.91	2.39

The results obtained using the hybrid system are compared to those obtained using a stand-alone ENN with four hidden neurons and trained using a gradient descent with momentum backpropagation algorithm with constant momentum of 0.9. An adaptive learning rate is used, as in the case of the hybrid system, the initial value being 0.1. These parameters are randomly selected. The results obtained are given in Table 4.

Table 4. Results for stand-alone Elman neural network

Performance measure	BSE-Sensex	NSE-Nifty	FTSE 100	DJIA	Nikkei 225
MAPE (%)	12.32	16.93	14.24	10.11	6.90
Md APE(%)	11.77	17.02	15.05	10.79	6.49
sMAPE(%)	0.14	0.19	0.14	0.11	0.07
sMdAPE(%)	0.14	0.19	0.14	0.11	0.07
MRAE	49.63	65.48	65.02	76.67	19.53
GMRAE	21.07	25.41	24.42	20.53	8.36
Md RAE	20.88	28.41	20.30	25.03	8.56
Rel. MAE	16.67	20.36	20.71	14.40	6.06
Rel. RMSE	14.63	16.93	18.71	10.92	5.20
LMR	5.37	5.66	5.86	4.78	3.30

From the tables 3 and 4 it is clear that the proposed hybrid system shows a significant improvement in performance when compare to a standalone ENN. Hence, it can be said that the proposed hybrid system is well capable of predicting stock markets with a high degree of accuracy.

References

1. Atsalakis, G.S., Valavanis, K.P.: Surveying Stock Market Forecasting Techniques – Part II: Soft Computing Methods. *Expert Systems with Applications* 36, 5932–5941 (2009)
2. Fama, E.F.: Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance* 25, 383–417 (1970)

3. Saad, E.W., Prokhorov, D.V., Wunsch, D.C.: Comparative Study of Stock Trend Prediction using Time Delay, Recurrent and Probabilistic Neural Networks. *IEEE Transactions on Neural Networks* 9(6), 1456–1470 (1998)
4. Nair, B.B., Mohandas, V.P., Sakthivel, N.R.: A Genetic Algorithm Optimized Decision Tree-SVM based Stock Market Trend Prediction System. *International Journal on Computer Science and Engineering* 2(9), 2981–2988 (2010)
5. Nair, B.B., Mohandas, V.P., Sakthivel, N.R.: A Decision Tree- Rough Set Hybrid System for Stock Market Trend Prediction. *International Journal of Computer Applications* 6(9), 1–6 (2010)
6. de Faria, E.L., Albuquerque, M.P., Gonzalez, J.L., Cavalcante, J.T.P., Albuquerque, M.P.: Predicting the Brazilian Stock Market Through Neural Networks and Adaptive, exponential smoothing methods. *Expert Systems with Applications* 36(10), 12506–12509 (2009)
7. Kuo, R.J., Chen, C.H., Hwang, Y.C.: An Intelligent Stock Trading Decision Support System through Integration of Genetic Algorithm based Fuzzy Neural Network and Artificial Neural Network. *Fuzzy Sets and Systems* 118, 21–45 (2001)
8. Yeh, I.-C., Lien, C.-H., Tsai, Y.-C.: Evaluation Approach to Stock Trading System using Evolutionary Computation. *Expert Systems with Applications* 38(1), 794–803 (2011)
9. Yeh, C.-Y., Huang, C.-W., Lee, S.-J.: A Multiple-Kernel Support Vector Regression Approach for Stock Market Price Forecasting. *Expert Systems with Applications* 38(3), 2177–2186 (2011)
10. Briza, A.C., Naval Jr., P.C.: Stock Trading System based on the Multi-Objective Particle Swarm Optimization of Technical Indicators on End-of-Day Market Data. *Applied Soft Computing* 11(1), 1191–1201 (2011)
11. Gholamreza, J., Tehrani, R., Hosseinpour, D., Gholipour, R., Shadkam, S.A.S.: Application of Fuzzy-Neural Networks in Multi-Ahead Forecast of Stock Price. *African Journal of Business Management* 4(6), 903–914 (2010)
12. Atsalakis, G.S., Valavanis, K.P.: Forecasting Stock Market Short-Term Trends using a Neuro-Fuzzy based Methodology. *Expert Systems with Applications* 36, 10696–10707 (2009)
13. Fraser, A.M., Swinney, H.L.: Independent Coordinates for Strange Attractors from Mutual Information. *Phys. Rev. A* 33, 1134–1140 (1986)
14. Kennel, M., Brown, R., Abarbanel, H.: Determining Embedding Dimension for Phase-Space Reconstruction using a Geometrical Construction. *Phys. Rev. A* 45, 3403–3411 (1992)
15. Elman, J.L.: Finding Structure in Time. *Cognitive Science* 14, 179–211 (1990)
16. Han, J., Kamber, M.: *Data Mining: Concepts and Techniques*. Morgan Kaufmann, San Mateo (2006)
17. Goldberg, D.E.: *Genetic Algorithms in Search, Optimization & Machine Learning*. Addison-Wesley, Reading (1989)
18. De Gooijer, J.G., Hyndman, R.J.: 25 Years of Time Series Forecasting. *International Journal of Forecasting* 22, 443–473 (2006)
19. Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R.: The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. *Journal of Forecasting* 1, 111–153 (1982)
20. Makridakis, S., Wheelwright, S.C., Hyndman, R.J.: *Forecasting: Methods and Applications*. John Wiley and Sons, New York (1998)

21. Fildes, R., Hibon, M., Makridakis, S., Meade, N.: Generalising about Univariate Forecasting Methods: Further Empirical Evidence. *International Journal of Forecasting* 14, 339–358 (1998)
22. Makridakis, S., Hibon, M.: The M3-competition: Results, Conclusions and Implications. *International Journal of Forecasting* 16, 451–476 (2000)
23. Makridakis, S.: Accuracy measures: Theoretical and Practical Concerns. *International Journal of Forecasting* 9, 527–529 (1993)

A Preemptive View Change for Fault Tolerant Agreement Using Single Message Propagation

Poonam Saini and Awadhesh Kumar Singh

Department of Computer Engineering, National Institute of Technology,
Kurukshetra, 136119, India
nit.sainipoonam@gmail.com, aksinreck@rediffmail.com

Abstract. The paper presents a proactive approach for failure detection. In our previous work [5], we assumed a trustworthy Transaction Manager, TM, amenable to the job of view creation, detection of faulty primary as well as backup replicas and to evacuate them from the transaction processing system. In the end, TM initiates a view in case it detects faulty primary or faulty replica. The TM provides an efficient failure-resiliency in the protocol; however, it also introduces the possibility of single-point failure. To eliminate the reliance on single TM, we propose a protocol that distributes the responsibilities of a transaction manager among $3f+1$ (f are faulty) replicas and results in a distributed Transaction Manager (DTM). The article attempts to limit the failure detection time to an optimum value, i.e., single message propagation time between any two nodes.

Keywords: Proactive View Change, Byzantine Agreement, Distributed Transaction Manager.

1 Introduction

The transaction processing poses many challenges in the area of distributed computing. In comparison with other applications, fault tolerance is a more serious concern in transaction systems where the behavior of the interacting nodes may be arbitrary. Moreover, with reference to agreement, the transaction handling protocol should maintain atomicity, i.e., either the operation is to be committed or aborted. Traditional Byzantine fault tolerant protocols e.g., BFTDC [1, 2] and Zyzzyva [3], deal with failures in a *reactive* manner, i.e., they rely on the specification of the faults to initiate view change. In a particular view, one of the replicas is chosen as primary and other replicas as backups. In the middle of agreement, if time out occurs for current view, due to delay in message propagation or the primary is found faulty, the view change occurs. The *proactive/preemptive* approach, on contrary, is designed to minimize the transaction discontinuity and latency while ensuring stability as well as availability of replicas through failure notifications in advance. Towards this goal, we build a system model to analyze the failure resiliency of our protocol under both reactive and proactive approaches. An optimized agreement is run among the replicas to reach the final decision about the transaction.

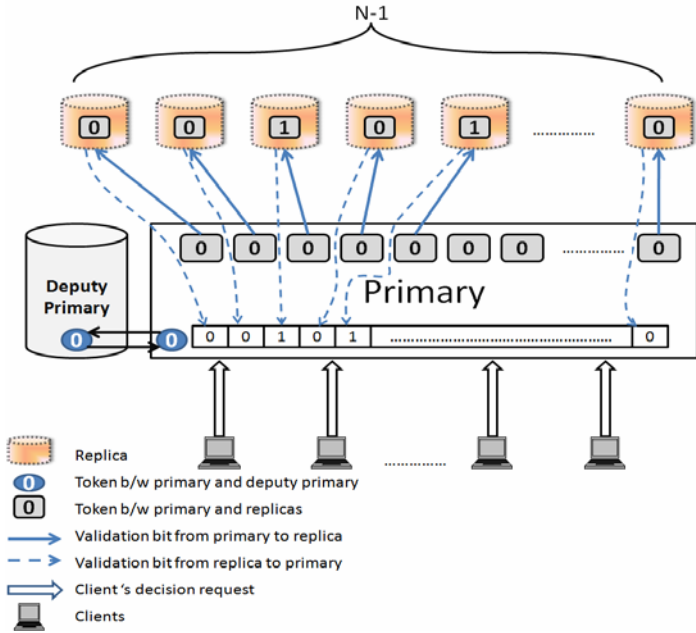


Fig. 1. Schematic view of DTM

Motivation. In most of the contemporary works, the protocols, as described in [4], replace a replica from the system when it is diagnosed as faulty. Although, the protocols produce desired result, they incur latency in order to initiate the view change which results in short-lived (i.e., transient) halts. We have devised a technique which is able to detect, in advance, the tentative fault in the system. The protocol fulfils all the requirements that are agreement, validity, and termination. Our previous protocol [5] relies on a trusted and fault-free Transaction Manager, TM. Although, it served as an efficient oracle for the proactive view change, the dependency is very high on TM. The proposed DTM (Fig 1) model intends to eliminate it and distributes the jobs of TM among the $3f+1$ replica's where f is the number of faulty one.

2 System Model

We assume a 2-tier architecture where the clients are not replicated as the Byzantine faults are considered only at the coordinator site. There are $3f+1$ coordinator replicas, among which at most f can be faulty during a transaction. . The protocol is started for a transaction whenever a commit request is received from the client. Each coordinator replica is assigned a unique id i , where i varies from 0 to $3f$. The id is required firstly, to identify the primary (P), and deputy primary (DP), from the replicas and secondly, to detect the fault in P , DP and replicas, if any. The deputy primary is chosen along with primary to avoid immediate transaction discontinuity, in case, the primary behaves arbitrarily. DP takes over the charge as soon as primary is diagnosed faulty.

3 Data Structures and Message Types

The protocol uses the following data structures and messages types:

Data Structure

- i. At all nodes
 - a. n_id : Node ID.
 - b. n_wt : Node weight randomly assigned, $0 < n_wt < 1$.
 - c. n_tb : A toggle bit $\{0, 1\}$ i.e., a priority associated with each node to nominate itself as primary and deputy primary.
 - d. p : A replica which acts as primary/coordinator for the transaction.
 - e. dp : A replica nominated as deputy primary to set back as primary in the event of primary being declared faulty.
 - f. r : Backup replicas in the system.
- ii. At primary, p node
 - a. $st_{p,r}$: A status token, initially, *false* (0), to verify the current state of replicas.
 - b. $st_{p,dp}$: A status token, initially set to *true* (1), to validate the existence of deputy primary, dp .
 - c. $rec_status_p[r]$: An array of tuples declaring the status of all replicas.
- iii. At deputy primary, dp node
 - a. $st_{dp,p}$: A status token to validate the existence of primary, p .
 - b. tp : A tentative primary status field, initially *passive*, to be set to *active* as soon as the value of st_p turns *true* (1) and primary is declared faulty.

Message types

- i. At all nodes
 - a. *View Message (VM)*: It is a view broadcast message to inform all replicas and participants about the primary, p , deputy primary, dp and backup replicas, r .
 - b. *Agreement Message (AM)*: It is the agreement message between the replicas to decide on a common value for a particular query.
 - c. *Decision Message (DM)*: It is the decision message to broadcast the final outcome of the transaction to the intended clients.

4 The Approach

4.1 The View Formation

In a transaction system, replicas move through a sequence of configurations which is termed as view (Fig 2). All the replicas enter into the transaction processing system with weighted information, n_wt . A toggle bit n_tb with value 0 or 1 is attached with all the replicas. In the next step, a primary, deputy primary and backup replicas are chosen as follows:

- a. For primary p : $\{n_wt(p) \geq 0.5 \text{ and } n_tb(p) = 1\}$.
- b. For deputy primary dp : $\{n_wt(dp) \geq 0.5, n_tb(dp) = 1, \text{ and } n_id(p) \neq n_id(dp)\}$.
- c. For replicas r : $\{n_wt(r) < 0.5 \text{ and } n_tb(r) = 0\}$.

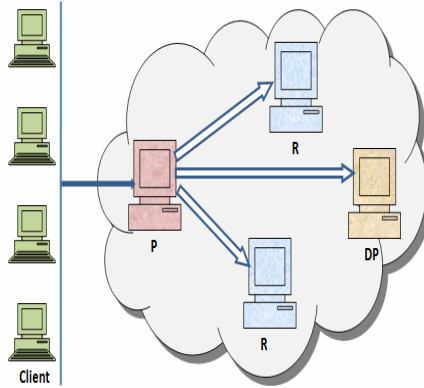


Fig. 2. The Semantic View Architecture

A view formation ends with the selection of all the above said entities. The current view message, which contains the current view number v , primary p , deputy primary dp and transaction id i , is then broadcast to each participating replica. Finally, if $3f$ replicas respond with an acknowledgement of current view, the primary begins the agreement round. In a particular view, the selection of a deputy primary ensures the pause-free processing among the replicas. However, in case if both, p and dp are fail-silent, a view change is carried out in order to provide liveness to the transaction system.

4.2 Proactive Fault Detection

Both, primary and deputy primary check each other for their correctness through a single-bit status token passing mechanism. The formal description of working is given as follows:

- a. A status token, initialized as 1 by primary, is rotated among both the entities.
- b. The value of status token ($st_{p,dp}$) is stored in the array of primary and passed to the deputy primary.
- c. Now, the deputy primary will complement the received value of the status token and stores the complemented value in status token ($st_{dp,p}$) and returns it back to primary.
- d. Afterwards, the same steps are repeated at primary. In this way, the status token circulates between primary and deputy primary.
- e. The process continues and the protocol for proactive fault detection runs simultaneously with the agreement protocol.
- f. If the value of status token stored at primary (deputy primary) and the value of status token received from deputy primary (primary) are same, then it signifies that deputy primary (primary) is declared faulty.

In the similar manner, the backup replicas are also continuously audited by the primary as follows:

- a. A different status token generated by primary with value 0 is circulated among all the backup replicas in the system.

- b. Primary declares a backup replica Byzantine faulty, whenever the backup replica returns a value of status token other than 0 e.g., it returns token value 1.
- c. In case, if no value is returned from the backup replica, the primary would declare it fail-silent or crashed.

This mechanism results in the preemptive detection of faulty primary, deputy primary, or replica, if exist, in the transaction processing system. Following the proactive detection mechanism of tentative failures into the system, an optimized Byzantine agreement protocol is executed simultaneously.

5 Correctness Proof

Lemma 1. The protocol maintains liveness. Formally,

$$(st_{p,dp} = 1 \wedge st_{dp,p} = 1) \vee (st_{p,dp} = -1 \wedge st_{dp,p} = -1) \leadsto tp.active$$

Argument: Assume that at some instant t , during transaction processing, the value of the status token ($st_{dp,p}$) in the array of deputy primary, dp complies with the last returned value of the status token by the primary p , i.e.,

$$[dp(st_{dp,p}) \wedge p(st_{p,dp}) = 1 \vee -1]$$

Then p is declared faulty by dp . Formally,

$$(st_{p,dp} = 1 \wedge st_{dp,p} = 1) \vee (st_{p,dp} = -1 \wedge st_{dp,p} = -1) \Rightarrow \mathfrak{d}(p)$$

Note: $\mathfrak{d}(p)$ denotes the faulty primary.

Now, according to the operational semantics of the protocol, the deputy primary takes over as primary, p and broadcast a view change message to all active replicas. The tentative primary field tp is set to active and status of dp is updated as primary p .

Lemma 2. The faulty replicas are exposed by non-faulty primary, eventually. Formally,

$$p(rec_status_p [r_i]) = (\perp \vee 1) \Rightarrow \mathfrak{d}(r_i)$$

Note: $\mathfrak{d}(r_i)$ denotes the faulty replicas, where the faulty replicas involves both Byzantine and fail-silent.

Argument: Assume that a replica r receives a status token with value 0 from primary p . Now, there can be two possibilities:

Case 1: The modified value of token is returned, i.e.,

$$r \text{ sets } (st_{p,r}) = \{1\} \text{ and broadcast to } p.$$

Case 2: No value is returned, i.e.,

$$r \text{ sets } (st_{p,r}) = \{ \} \text{ (a null value)}$$

Now, after the first round of transaction, the primary verifies for the status token array received from the replicas. The modified value in the array or an empty field in the array would confirm the presence of Byzantine faulty and fail-silent replicas in the system. Formally,

$$rec_status_p \in \{(\perp \vee 1)\} \Rightarrow \mathfrak{d}(r_i)$$

Lemma 3. All the non-faulty replicas will eventually reach the same view in which the replica to be chosen as primary, initially, is not faulty. Formally,

$$\forall r \in \{p, dp, r\}, p \neq \Delta(p)$$

Argument: Assume the contrary. Let x and t to be primary for view $v+1$ and $v+2$ respectively. Now, if x is primary for view $v+1$, it will send new view message if it has received $2f+1$ view change messages for view $v+1$. Let, this set is R_1 . Similarly, t will send new view message if it has received view change message for view $v+2$ from $2f+1$ replicas. Let, this set is R_2 . Now, between R_1 and R_2 there are at least $f+1$ common replica's because there can be at most f faults in the system at a time.

Note: In the worst case, out of $f+1$ replicas, at most, f replicas could be faulty.

Thus, at least one correct replica is still present. Hence, the correct replica would not send new view change message for both the replicas to be selected as primary in their respective views. Hence the lemma holds.

6 Conclusion

The paper proposed an optimized and novel proactive mechanism towards the proactive detection of faulty replicas. The distribution of responsibilities of the Transaction Manager (TM) among the replicas has removed the dependency on it. The protocol prompts to evacuate both, the Byzantine as well as fail-silent i.e., a crashed replica. The static analysis of the protocol verifies the significance of protocol under both the faults. For future work, the extension of the boundaries for arbitrary behavior of the nodes is under revision.

References

1. Castro, M., Liskov, B.: Practical Byzantine Fault Tolerance and Proactive Recovery. *ACM Transactions on Computing Systems* 20, 398–461 (2002), doi:10.1145/571637.571640
2. Zhao, W.: A Byzantine Fault Tolerant Distributed Commit Protocol. In: *IEEE International Symposium on Dependable, Autonomic and Secure Computing*, pp. 37–44 (September 2007)
3. Kotla, R., Alvisi, L., Dahlin, M., Clement, A., Wong, E.: Zyzzyva: Speculative Byzantine Fault Tolerance. *ACM Proceedings of twenty-first Symposium on Operating Systems and Principles* 41(6), 45–48 (2007)
4. Fisman, D., Kupferman, O., Lustig, Y.: On verifying fault tolerance of distributed protocols. In: Ramakrishnan, C.R., Rehof, J. (eds.) *TACAS 2008*. LNCS, vol. 4963, pp. 315–331. Springer, Heidelberg (2008)
5. Saini, P., Singh, A.K.: An Efficient Byzantine Fault Tolerant Agreement. In: *ICM2ST 2010: Proceedings of the International Conference on Methods and Models in Science and Technology*, American Institute of Physics (AIP), December 2010, vol. 1324, pp. 162–165 (2010), doi:10.1063/1.3526183, ISBN: 978-0-7354-0879-1

A Model for Detection, Classification and Identification of Spam Mails Using Decision Tree Algorithm

Hemant Pandey¹, Bhasker Pant², and Kumud Pant^{3,*}

¹ Senior Infrastructure Engineer, EDS, Pune, Maharashtra, India
helpmenu@gmail.com

² Assistant Professor, Department of IT, Graphic Era University, Dehradun, India
pantbhaskar2@gmail.com

³ Senior Research fellow, Department of Bioinformatics, MANIT, Bhopal, India
pant.kumud@gmail.com

Abstract. Spam mails are unsolicited bulk mails which are meant to fulfill some malicious purpose of the sender. They may cause economical, emotional and time losses to the recipients. Hence there is a need to understand their characteristics and distinguish them from normal in box mails. Decision tree classifier has been trained with the major characteristics of spam mails and results obtained with more than 86.7437% accuracy. This classifier can be a valuable strategy for software developers who are trying to combat this ever growing problem.

Keywords: spam mails, header characteristics, message body, classifier/predictor.

1 Introduction

E-mail spam, known as unsolicited bulk Email (UBE), junk mail, or unsolicited commercial email (UCE), is the practice of sending unwanted e-mail messages, frequently with commercial content, in large quantities to an indiscriminate set of recipients. Everyday 1000's and lack's of spam mails are send by unknown senders through out the world. At the first instance spam mails may seem harmless unwanted mails but some of the innocent users may fall prey to the ill intentions encrypted within it. The infiltration of ones mail box with these mails is not only annoying but it is also dangerous. It has the potential to cause great financial losses to the administration which deals with it and to the common men who may fall in this spam trap that has been laid for them. They are usually loaded with viruses, spy programs and software's. Realizing the importance of spam mails and problems caused by it, here we present a comparative account of random forest and decision tree algorithm for classification, filtering and prediction of spam mails. Previously the header characteristics of these mails have been used for their classification and identification [4]. The same have been used here with a comparative account of decision tree and random forest

* Corresponding author.

method. Hence this is a novel step to make general people aware of these silent monsters and also for software developers to device new hybrid models and approaches using decision tree classifier and random forest algorithm for increasing efficiency of their spam trapping ability.

2 Materials and Methods

2.1 Data Set

We collected around 61 spam mails and 37 normal mails from our gmail accounts. They were checked for various attributes and a data sheet was made for its tabulation.

2.2 The Algorithm

For analyzing our data we used two algorithms from weak.

1) Decision tree and its boosting algorithm J48 from Weka and 2) Random forest algorithm [1-3]. Here using both the above classifiers a model is developed to classify spam and normal in-box mails by using their header characteristics. After data filtering, each classifier is trained and cross-validated for 10-times with a 10-fold random sampling. The ten resulting values for each performance parameter are averaged to obtain the final figures, and Receiver Operating Characteristic (ROC) curves and TP rates vs. FP rates are plotted and analyzed [4-7]. In the past decision tree algorithm has been used for classification of plant and animal micro RNA using their various attributes and Random Forest for classification of MMP's hence are quiet powerful in classification of both supervised and unsupervised data [8, 9].

3 Results

The comparative results of both classifiers are depicted in table 1. It was seen that characteristic A was the most important criteria in distinguishing spam and normal mails. On removing which the accuracy was reduced.

The decision tree is shown in figure1.

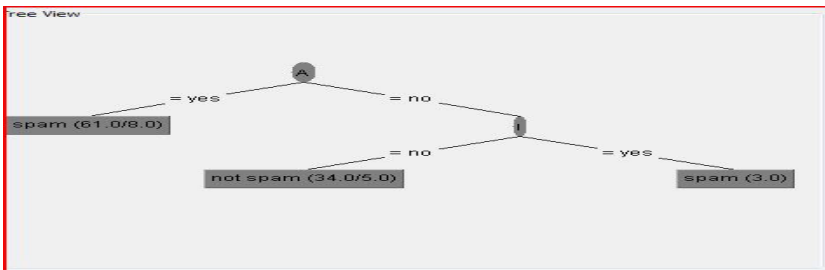


Fig. 1. J 48 Deision tree with A (recipients address not in to: or cc:field) as the principal field

Table 1. Comparative results of both the classifiers

With Decision tree				With Random Forest			
All characteristics included		Characteristic A not included		All characteristics included		Characteristic A not included	
Correctly classified instances	85 (86.7347%)	Correctly classified instances	79 (80.6122%)	Correctly classified instances	82 (83.6735%)	Correctly classified instances	78 (79.5918%)
Incorrectly classified instances	13 (13.2653%)	Incorrectly classified instances	19 (19.3878%)	Incorrectly Classified Instances	16 (16.3265%)	Incorrectly Classified Instances	20 (20.4082%)
Kappa statistics	0.7132	Kappa statistics	0.5897	Kappa statistic	0.6489	Kappa statistic	0.5658
Mean absolute error	0.2313	Mean absolute error	0.2956	Mean absolute error	0.2133	Mean absolute error	0.2618
Root mean squared error	0.3438	Root mean squared error	0.4092	Root mean squared error	0.3598	Root mean squared error	0.4254
Relative absolute error	49.1191%	Relative absolute error	62.7718%	Relative absolute error	45.3011%	Relative absolute error	55.6005%
Root relative squared error	70.8651%	Root relative squared error	84.3524%	Root relative squared error	74.1698%	Root relative squared error	87.6859%
Total number of instances	98	Total number of instances	98	Total Number of Instances	98	Total Number of Instances	98

4 Conclusion

Besides header files, other criteria can be very important in distinguishing spam from normal mails like subject line, sender domain, message content etc. The author also wishes to extend the above classifier by incorporating the above characteristic so as to further improve the classifier.

References

1. Witten, I.H., Frank, E.: Data Mining – Practical machine learning tools and techniques with Java implementations. Morgan Kaufmann, San Francisco (2005)
2. Langley, P., Sage, S.: Elements of machine learning. Morgan Kaufmann, San Francisco (1994)
3. Weka Data Mining Java Software, <http://www.cs.waikato.ac.nz/~ml/weka/>
4. Mike Spykerman – CEO Red Earth Software, Typical spam characteristics How to effectively block spam and junk mail
5. Langley, P., Sage, S.: Elements of Machine Learning. Morgan Kaufmann, San Francisco (1994)
6. Han, J., Kamber, M.: Data Mining: Concepts and Techniques. Morgan Kaufmann, San Francisco (2001)
7. Heckerman, D., Geiger, D., Chickering, D.M.: Learning Bayesian network: The combination of knowledge and statistical data. Machine Learning 20(3), 197–243 (1995)
8. Pant, B., Pant, K., Pardasani, K.R.: ‘Decision tree classifier for classification of plant and animal micro RNA’s. Communications in Computer and Information Science, Part 9 51, 443–451 (2009), doi:10.1007/978-3-642-04962-0_51.
9. Pant, B., Pant, K., Pardasani, K.R.: Machine Learning Model for Domain Based Classification of MMP’s. The Internet Journal of Genomics and Proteomics 5(2) (2010)

Author Index

- Acharya, Arup Abhinna 100
Agarkar, Ankita 222
Agarwal, Ajay 255, 368
Ajitha, P. 489
Ambeth Kumar, V.D. 214, 358
Analoui, Morteza 461
Anand, R. 468
AntopremKumar 485
Aravindan, C. 312, 323, 427
Arya, Meenakshi 123
- Babu Rao, M. 399
Bagchi, Parama 287
Balaji, Anerudh 291
Balamurugan, Karthigha 297
Bande, Shivangi 82
Bandhopadhyay, T.K. 344, 348
Banerjee, Soumik 39
Bansal, Sanjay 445
Bella Mary, I. Thusnavis 481
Bhalani, Jaymin 184
Bhuvana, J. 312
Bose, Susan Mary 477
Budyal, V.R. 375
Burse, Kavita 67
- Chakroborty, Debashish 39
Chana, Inderveer 449
Chandra, E. 489
Chaudhary, Shubhangi R. 440
Chitra, K. 196
Choubey, Abha 422
Choubey, Siddhartha 422
Chougule, Archana 190
- Danda, Aneeshwar R. 273
Daniel 485
Dash, Gananath 170
david Raj, Newton 485
Deenadayalan, T. 468
Deshmukh, Manjusha 55
Devashrayee, N.M. 33, 95
Dey, Kashinath 305
Dhayanandh, S. 178
Dilli, Ravilla 240
- Dinesh, P.M. 263
D'sa, Kiran 273
Dubey, Ashutosh K. 144
Dubey, Nilesh 82
Dubey, Vandana 82
- Fatima, Mehajabeen 348
- Gahankari, Sonal 55
Ganda, Geetika 379
Govardhan, A. 19, 399
Grace Mary Kanaga, E. 317
Gupta, B.B. 117
Gupta, Roopam 344, 348
- Haider, Mohammad 150
Halkarnikar, Pratap 190
Hanumantharaju, M.C. 162
Hari, CH.V.M.K. 227, 281
Hima Bindu, M. 407
Hiremath, S.G. 375
- Jain, A. 117
Jain, Jyoti 344
Jaisakthi, S.M. 427
Jaisankar, N. 454
Jayakumar, M. 297
Jinwala, Devesh C. 388
Joshi, Apoorv 222
Joshi, R.C. 117
Josphineleela, R. 352
Josprakash 485
Juyal, S. 117
- Kannan, A. 454
Kaur, Gunjit 435
Kaur, Pankaj Deep 449
Kaushal, B.S.S. 281
Kavitha, Ch. 399
Khan, Zafer Jawed 493
Kirar, Vishnu Pratap Singh 67
Kole, Arnab 305
Korde, Mridula 74
Koshy, Divya Mary 477
Kosta, Yogesh 184

- Kosta, Yogeshwar 111
 Krishnam Raju, K.V. 267
 Krishnan, Shobha 248
 Krishna Prasad, A.V. 301
 Kumar, A. 392
 Kumar, Abhishek 336
 Kumar, Neeraj 259, 384
 Kumar, Niti 245
 Kumar, P. Pavan 403
 Kumar, Sumit 156
 Kumar, Vimal 132
 Kushwaha, Ganesh Raj 144
- Lakshmi, A. 499
 Lavanya, N. 178
- Maheshwari, Manish 384
 Maity, Saikat 277
 Majumder, Saikat 207
 Makhijani, Jagdish 6
 Mala, C. 138, 431
 Manohara Pai, M.M. 1, 340
 Manoria, Manish 67
 Manvi, S.S. 375
 Meher, Jayakishan K. 170
 Meher, Pramod Kumar 170
 Mehta, Mayuri A. 388
 Mishra, Nibedita 170
 Misra, M. 117
 Mittal, Shaily 379
 Mohanavalli, S. 427
 Mohandas, V.P. 499
 Mohapatra, Durga P. 100
 Mohapatra, Pranab Kishor 170
 Mondal, Hemanta 245
 Motwani, Mahesh 6
 Mukherjee, Amartya 291
 Mukherjee, Debasis 245
 Murali Nath, R.S. 240
 Murugesan, D. 61
- Naik, Amisha 95
 Nair, Binoy B. 499
 Nandi, Subrata 291
 Naveen, A.N. 499
 Nayak, Rakesh 27
 Nikhil 384
 Niranjana, Manoj Kumar 6
 Nuparam 45
- Oza, Shruti 33
- Padmaja Rani, B. 301
 Padmavathi, G. 196
 Pais, Alwyn R. 336
 Pal, Shantanu 287
 Pandey, Hemant 513
 Pant, Bhasker 513
 Pant, Kumud 513
 Patel, Sanket 111, 184
 Patel, Shobhit 111, 184
 Patnaik, Debashree 100
 Patnaik, L.M. 273
 Patra, Sushovan 291
 Paul, Biju 332
 Prabhakar, R. 117
 Prabhakara Rao, B. 399
 Prachi 379
 Pradhan, Jayaram 27
 Prasadh, K. 332
 Prasad Reddy, P.V.G.D. 227
 Puranik, Minal M. 248
- Raamesh, Lilly 327
 Rahamatkar, Surendra 255
 Rahmatkar, S. 368
 Raj, Adhira 297
 Raja, K.B. 13, 273
 Rajagopalan, Narendran 431
 Rajalakshmi, K. 407
 Rajalakshmi, R. 323
 Rajesh Kumar, G. 233
 Rajkumar, S. 178
 Rajpurohit, Vijay S. 1, 340
 Rajput, Anil 6
 Ramachandran, Nitya 417
 Ramachandran, S. 162
 Ramakrishnan, M. 214, 358
 Rama Krishna, E. 233
 Ramakrishna, S. 301
 Ramakrishnan, M. 352
 Ramamohanreddy, A. 19
 Rama Rao, T. 61
 Ramasubbareddy, B. 19
 Ramasubramaniam, N. 403
 Rambabu, N. 233
 Ramesh, S. 61
 Ramesha, K. 13
 Rameshbabu, D.R. 162
 Ramya Sri, A.P. 178

- Ranjeet 392
 Rathkanthiwar, Anagha 74
 Raval, Mukesh Kumar 170
 Ravishankar, M. 162
 Reddy, B.V.R. 245
 Rentapalli, Vanitha Rani 493
 Reshma, P. 203
 Rezvani, Mohammad Hossein 461
 Rini Rosebell, V. Jerine 473

 Sabeenian, R.S. 263
 Sachan, A.K. 6
 Sadalkar, Kunal M. 336
 Saha, Sujay 305
 Saha, Sujoy 291
 Sai, S. Gnana 499
 Saini, Poonam 507
 Sairam, Ashok Singh 156
 Sastry, C.V. 27
 Sathisha, N. 273
 Sen, Praveen 255
 Sengar, Sandeep Singh 132
 Sethi, Tegjyot Singh 281
 Shanker, Udai 45
 Sharanya, R. 477
 Sharma, Abhishek 281
 Sharma, Sanjeev 445
 Sharma, Vishwas 336
 Shekar Reddy, P. Chandra 240
 Sherin 473
 Shindu 473
 Shravani, D. 301
 Shrivastava, Nishant 144
 Shunmuganathan, K.L. 87
 Siddavatam, Rajesh 123
 Sikri, Monika 411
 Sil, Jaya 277
 Singh, A.K. 117
 Singh, Akash 384
 Singh, Akhilendra Pratap 132

 Singh, Awadhesh Kumar 507
 Singh, Kalyan 245
 Singh, Mandeep 435
 Singh, Nahar 156
 Sinha, G.R. 422
 Soni, Himanshu 111
 Sreedharan, Swetha 273
 Srinivas, V.V. 106, 403
 Sugumar, D. 473, 477
 Sujithra, T.L. 477
 Suresh, T. 87
 Suresh Babu, K. 273

 Thanuja, M.K. 138
 Thilagam, P. Shanthi 336
 Thomas, Joshua 485
 Tiwari, Abhinav 259
 Trivedi, Ishita 445

 Ujwal, R. 273
 Uma, G.V. 327
 Upendra Kumar, M. 301

 Valarmathi, M.L. 317
 Valli Kumari, V. 267
 Varadhan, V.V. 106
 Venkatesh, G.S. 499
 Venkat Reddy, A. 233
 Venugopal, K.R. 273
 Verma, Kuldeep 222
 Verma, Shrish 207
 Vijayan, Vinodh P. 332
 Vijay Kumar, T.V. 150

 Wadhawan, Nisha 259
 Wairiya, Manoj 132

 Xavier, Lidiya 481

 Yadav, Arun Kumar 255, 368
 Yogesh, P. 417