ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTING 180

Zhenyu Du (Ed.)

Intelligence Computation and Evolutionary Computation

Springer

Advances in Intelligent Systems and Computing 180

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

For further volumes:
http://www.springer.com/series/11156

Zhenyu Du (Ed.)

Intelligence Computation and Evolutionary Computation

Results of 2012 International Conference of Intelligence Computation and Evolutionary Computation ICEC 2012 Held July 7, 2012 in Wuhan, China

Springer

Editor
Information Technology & Industrial
Engineering Research Center
Wuhan
China

ISSN 2194-5357 e-ISSN 2194-5365
ISBN 978-3-642-31655-5 e-ISBN 978-3-642-31656-2
DOI 10.1007/978-3-642-31656-2
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012941506

© Springer-Verlag Berlin Heidelberg 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, 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.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

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

Preface

2012 International Conference of Intelligence Computation and Evolutionary Computation (ICEC 2012) is held on July 7, 2012 in Wuhan, China. This conference is sponsored by Information Technology & Industrial Engineering Research Center. It is the most international academic conference per year in Intelligence Computation and Evolutionary Computation area. The main aim is to provide an academic platform to exchange experience and production for researchers, industrial practitioners and teachers in the fields of intelligence computation and evolutionary computation.

In this conference, we had received 400 submissions from 6 countries and regions. Each of the papers was reviewed carefully by two famous specialists in this field. After received the revised papers on the basis of the specialists' advice, finally we collected 150 papers into the proceedings. In addition to the regular publications of the papers, we have also invited several specialists and scholars in this field to make keynote speech at the conference.

We would like to take this opportunity to express our thanks to the individuals and organizations for their efforts to serve the conference. We'd like to extend our thanks to members of academic committee for their effort to the conference; especially, we'd like to thank to members of organizing committee for their hard working; finally, we would like to express our appreciations to the participants of this conference.

E.R. Zheng
Chairman of ICEC 2012

Zhenyu Du
Chairman of Publication Committee

Contents

Intelligence Computation

Advantages of Cross-Border Acquisition under International Financial Crisis 1
Zuoling Nie, Yuhai Su, Chong-jin Wang

A New Extension of LBP for Texture Classification 7
Kunlun Li, Pan Wang

Adaptive Backstepping Control for Nonlinear Systems Using Support Vector Regression 13
Liu Yinan, Zhang Shengxiu, Cao Lijia, Zhang Chao

The Application of Data Mining Technology in the Breakdown Maintenance of Warship Equipment 25
Hua Rong, Mingrong Chen, Zhen Wang

Intrusion Detection Based on Immune Principles and Fuzzy Association Rules 31
Zhang Lei, Meng Lingrui, Hou Chunjie

Temperature Model for FOG Zero-Bias Using Gaussian Process Regression 37
He Zhikun, Liu Guangbin, Zhao Xijing, Yang Jian

Pulse Wave Averaging Computation and Clinical Verification 47
Li Yang, Yunan Hu

Conditional Fault Diagnosis of Bubble Sort Graphs under the PMC Model 53
Shuming Zhou, Jian Wang, Xirong Xu, Jun-Ming Xu

Process Simulation of Micro Device with Virtual Reality Technology 61
Zheng Liu, Hua Chen

Intelligent Control of Beer Fermentation Temperature 67
Tianli Li

Discussion on Enterprise Emergency Management Decision Support System 73
Wenchang Liu

Shallow Parsing of Chinese Based on HMM Model 79
Zheng Weifa, Xie Wenliang

Application of Seismic Sedimentology Method to Beach-Bar Depositional System 87
Guo Jianqing, Lin Chengyan

Visual Sign System Design of Modern Office Building 93
Shen Lei

Research on Image Intensity Based on Matlab 101
Wang Yueping, Yi Zhi'an

A Novel Digital Audio Watermarking Algorithm Based on Wavelet Transforming 109
Liu Yanjun

The Demonstrating and Monitoring Software of Robot Control System 117
Zou Huan, Zhang Xueping, Yang Yanxin, Wang Xin

Discussion on Training Mode of Entrepreneurial Talent in Electronic Business 125
Jin Tao

The Design ECU of Vehicle with the Use of Controller Integrated PICI18F258 SCM 133
Shujun Hua

Encrypt Hardware Device Based on Cryptographic Algorithm 139
Chen Cong

The Data Mining Technology in the Application of Graduates' Employment 147
Hu Yan

Application Research of Virtual Reality Technology in Electronic Technique Teaching 153
Hu Liang, Bi Xiaoming

The Platform Design of Practice Teaching in the Electromechanical Major That Based on the Virtual Reality Technique 161
Hu Liang, Yang Wen, Huang Chengzhong

The Exploration Model of New-Type Software and the Application in the System Exploration of Teacher Management 167
Cao Yu

Education Software Design of Electronic Technique That Based on Virtual Reality Technology 173
Hu Liang, Bi Xiaoming

Study on Innovation Education Technology Based on New Computer Technology 179
Cao Yu

Research on Multimedia Database 185
Hu Liang, Huang Manyuan, Huang Chengzhong

The Experimental Teaching Platform of Mechanotronics Major That Based on Java EE Technology 191
Hu Liang, Huang Chengzhong

The Analysis of Grid Database Technology in the Computer Application System 197
Juan Chen

Study of Smart Warehouse Management System Based on the IOT 203
Wen Ding

Design and Implementation of Visual C + + Based Digital Image Processing System 209
Hu Liang

A New Multi-chaos Based Image Encryption Algorithm 215
Nan Lin, Xiaofeng Guo, Ping Xu, Yuqin Wang

Performance of an Optical Packet Switch with Limited Wavelength Converter 223
Jie Li

Time Petri Net-Based Software Reliability Analysis 231
Maoqiang Yang

Digital Evidence Protection Model with Batch-Verifying and Public Verifiability for Computer Forensics 237
Zhao-Hui Qi, Zhan-Feng Gao, Ying-Ming Shen, Bing-Xin Han

Finite Element Analysis of a Bus Skeleton and Structure Improvement 243
Fachao Jiang, Yuezhen Fan, Entuo Liu, Molin Wang

Design on a New Energy Lawn Mower: Fuel Cell Lawn Mower 249
Fachao Jiang, Entuo Liu, Molin Wang

Study on the SRI Sensing Characteristics of PFBG and TFBG 255
Rensheng Shen, Wen Lin, Xiangping Li, Xiao Zhang, Guotong Du

E-commerce Case Study of Fast Fashion Industry 261
Zhenxiang Wei, Lijie Zhou

Rapid Impulsive Noise Denoising in Range Images 271
Jian Wang, Zhen-qiang Yao, Lin Mei, Yao-jie Zhu, Yuan Yao, Yang Zhang

Study on Race and Hazard of Combinational Logic Circuit 281
Wang Xiuling, Wang Haocheng

Approach to Analyze Time Series Similarity Pattern Mining Based on Haar 287
Li Yi

Design of a Four-Annulus Integrated Disk Electrode for Bioelectrical Impedance Spectroscopy Measurement 297
Kening Wang, Xiongxing Zhang, Wei Wang

Development of Frequency and Power Control System Hardware-in-Loop Simulation Platform for Ship Power Plant 305
Sun Bin, Zeng Fan-ming, Zhang Wei-dong, Zhao Hua

Control of Sludge Recycle Flow in Wastewater Treatment Plants Using Fuzzy Logic Controller 313
Wangyani

Research and Application of Embedded Real-Time Operating System 319
Zhangbo

A PCA-Based Automated Method for Determination of Human Body Orientation 327
Weihe Wu, Aimin Hao, Wentao Wang, Yiqiang Wang

GPU-Based Large Seismic Data Parallel Compression 339
Kai Xie, H.Q. Yu, G.Y. Lu

Transform Optimization Based on Color Image Compression 347
Kai Xie, H.Q. Yu, G.Y. Lu

The Resilience Framework for Interdependent Infrastructure Systems Using the Dynamic Inoperability Input-Output Model 355
Wenping Xu, Zhiwu Wang, Mingsheng Hu, Liu Hong, Xueguang Chen, Shuliang Wang

Evolutionary Computation

Multi-source Separation Using over Iterative Empirical Mode Decomposition 363
Yutian Wang, Hui Wang, Qin Zhang

Paraphrase Based Similar Expression Generation 369
Wei He, Zhanyi Liu, Ting Liu

Parameters Allocation of Genetic Algorithm Based on RSSI of DSRC Communication 375
Chun-Liang Chen

Identification of FOG Multivariable Model Based on Fuzzy Relation 381
Zhao Xijing, Wang Lixin, He Zhikun, Li Rui

An Improved Ant Colony Algorithm Combined with Genetic Algorithm and Its Application in Image Segmentation 389
Zhou Haifeng

Paper's Indirect Strain Control Bases on Inverter 395
Tianli Li

Study on the Online Clustering Algorithm Based on Grid Structure 401
Wu Peng, Jia Jun, Zhu Tianxiang

The Measurement of Low-Level ^{220}Rn by Software Delayed Coincidence Method 407
Yan Yongjun, Lai Wei, Zhou Jianliang

An Optimization of Intelligent Fire Alarm System for High-Rise Building Based on ANASYS 415
Liu Keding

Coal Mine Safety Monitoring and Evaluation System Based on .Net 423
Zhang Chunzhi, Hu Guangzhou, Li Haoru, Wang Kejia

Research on Web Information Filter Model 431
Liu Yiqin

Power Flow and Power Flow Calculation Research in Power System 439
Li Jun

The Analysis and Design of Human Resources Information Sharing System in University 447
Jiang Yan

The Application Research of Multimedia Database in Innovative Venturing Course 453
Wu Binying, Chen Shiqiang, Fang Li

Design and Realization of WCDMA Project Simulation System 461
Yehui Liu, Huaiqun Wang

The Application of Computer Technology in the Interior Decoration Project 467
Qiu Yu, Qiu Jing

The Application of Image Recognition Technique in Candidate's Identity Authentication System 475
Gu Hongjiu

Research on the Security Risk Assessment and the Solutions of Cloud Computing 481
Wen Ding

The Study of Security Management Platform Technology and Methods Based on Cloud Computing 487
Zhang Jie

Data Mining in the Application of E-Commerce Website 493
Gu Hongjiu

Analysis of Security Mechanism in Web Database Application System 499
Juan Chen

The Research and Summary of Evolutionary Multi-objective Optimization Algorithm 505
Xu Jingqi

A Finite-Time Convergent Variable Structure Guidance Law with Impact Angle Constraint 513
Zuo-e Fan, Xian-jun Shi, Chang-peng Pan, Wen-jin Gu

Highway Traffic Automatic Detection System Based on Video and Image Processing 521
Jiang Ying

An Improvement to Leach-Based Routing Algorithm 527
Liu Re, Wang Xiu-Ping

Specifying API Trace Birthmark by Abstract Interpretation 535
Ying Zeng, Fenlin Liu, Jian Chen, Bin Yan

Extracting Semantic Information from Chinese Language Patent Claims . . . 547
Yong Tang, Shihan Yang, Jiwen Chai, Shanmei Liu

Sensitivity Analysis on the Influence Factors of Software Reliability Based on Diagnosis Reasoning 557
Li Qiuying, Li Haifeng, Wang Guodong

Study on Method of Layer Modelling in Geological Body Visualization 567
Niu Wenjie

Design and Realization of a Novel Low-Voltage High-Efficiency On-Chip Current Sensing Technique 575
Wei HaiLong, Geng ZengJian, Liu Zhi, Liao Xue

Construction of Mamdani Fuzzy Classifier Based on Genetic Algorithm 583
Zhou Weihong, Xiong Shunqing

Change Detection Using High Spatial Resolution Remotely Sensed Imagery 591
Zhang Ruihua, Wu Jin

Optimization of BP Neural Network Classifier Using Genetic Algorithm . . . 599
Zhou Weihong, Xiong Shunqing

Output-Feedback Stabilization of Discrete-Time Systems with Multiple Time-Varying Delays 607
Li Zhang, Fule Li

An Improved OTP Authentication System and Its Application in Mobile Commerce 615
Ye Xi

Research on ESC Hydraulic Control Unit Property and Pressure Estimation 625
Liu Wei, Ding Haitao, Guo Konghui

The Emotional Robot Model Based on Endocrine System 633
Huang Guorui

Research on the New Channel Estimation Algorithm for MIMO-OFDM System Based on Training Sequence 647
Zhigang Yang

The Research and Implementation of Technology of Generating Test Paper Based on Genetic Algorithm 657
Shao Ming-zhu, Li Wei-feng, Du Jing-yi

Intelligent Control Technology for Frequency Conversion Pump in Air Conditioning Systems 665
Shizhong Yang, Qingchang Ren

The Application and Development of Photoelectric Sensor 671
Zheng Chunjiao

Online Random Seeking in FLV Video Based on File Cutting 679
Zhihui Du, Wenjun Fan, Yunpeng Chai, Tianle Zhang

A Self-Adaptive Sliding Window Technique for Mining Data Streams 689
Yi Yang, Guojun Mao

The CAD Solutions in Highway Landscape Greening 699
Yin Guangli, Qian Shao

Using Genetic Algorithm to Improve Gray Watermark Algorithm 707
Jianbing Yi, Qiuen Xie

A New Secure Windows Terminal Based on Dual-Subsystem Architecture 715
Li XinMing, Li Yi, Zhang GongXuan, Wang XiaoMei

OsgAL-Based Sound Source Management in Simulation Games 723
Xin Li, Dehui Kong, Yong Zhang

An Automatic Extraction Method of Surveillance Visual Context 733
Haozhe Liang, Shukui Xu, Guohui Li

Improved Method of the PSF Parameters Estimation for Image Restoration 743
Zeng Liang, Zhou Yang, Yang Yingyuna

Image Enhancement Based on NSCT 753
Ma Changxia, Bi Ye, Jiang Hongan, Chen Yunping, Zhang Zhanqiang

Network Computation

Research on Design of Network Teaching System 761
Wenpeng Lu, Ruojuan Xue, Haixia Li

A Semantic Framework for Integration of Prison Internet 767
Peiliang Sun, Wei Wei, Ming Xia

Architecture of the Cloud Computing Platform for Enterprises 773
Shi Xingjun, Xuezhen Zhu

The Application of SSL Protocol in Computer Network Communication . . . 779
Tian Huan

Research on Wireless Network Security Model 785
Linmao, Wangyong

Application of Network Platform in EFL Writing Feedback 791
Kang Hui, Shi Ziming

The Bayesian Network and Trust Model Based Movie Recommendation System 797
Du Wei, Chen Junliang

The Network Security Management System Based on the Fuzzy Dynamic Evaluation 805
Deng Fan, Pei Tao, Ji Hongcheng

The Research and Design of IMS Network Management 813
Xia Ling

Requirements of the Times – The Collaboration of E-Government and E-Commerce 819
Xiao Shufeng

Study on the Development Path of China Modern Coal Logistics Based on Electronic Commerce 825
Jiang Lingli, Cen Kuang

Research on Flash Based Online Game 831
Liu Jinjun

Applicability Analysis and Classification in Network System Security 837
Luo Zhuojun

Design of Multimedia File Security Communication Platform Architecture Based on Web 841
Liu Yanshu, Yi Can, Li Zhiyong

Research on the Security Issues of Cloud Computing 845
Jia Weihua, Sun Shibing

Alarm Analysis on Intrusion Detection of Network Flow 849
Liu Yanshu, Cao Yujun, Lei Jiping

The Exploitation and Design of Enrollment and Employment Management System That Based on B/S 855
Li Bo

Research on Network Technology-Based Data Integration Model 861
Chen Min, Wang Jianping, Tan Shaosheng

Sentiment Analysis in Chinese BBS 869
Dai Lu, Ding Lixin

A Modified Transport Protocol for Heterogeneous Wireless Sensor Networks 875
Wu Yangbo, Zou Donglan, Li ShuLiang

Research on the Development and Application of Intelligent Network Teaching System 881
Li Fachun

Study on the Computer Networks System Safety Detecting Technology 887
Hu Liming

Algorithm Research of Network Course Design Based on Personal Learning 895
Li Fachun

Observation on the Wireless Sensor Router Algorithm Based on Mobile Agent 905
Pei Fang, Zhang Jie

Higher Vocational College Computer Culture Foundation Online Examination System Design 911
Yonghua Zhang, Lizhe Wang, Ji Li

MVDR Method for the Whole Conformal Arrays 917
Lin Zou, Zishu He

Research and Implementation of SOA-Based E-commerce Platform 923
Hu Liang

Production Dynamic Scheduling Method Based on Improved Contract Net of Multi-agent 929
Ling Qin, Shulin Kan

Research on Information Integrated Platform for Gear Enterprise Networked Manufacturing 937
Xiang-tao Chen, Xiang-yang He

A New 3D Statistical Method for Mobile Channel Model 945
Yihuai Yang, Gengfa Fang, Eryk Dutkiewicz, Yangqing Li

An Optimized Organizational Structure Model of Internet Security Management 953
Ran Lia, Shoulian Tang

Study on the Business Model of Tourism Electronic Commerce in China . . . 959
Jian Song

Semantic-Based Wireless Video Surveillance System 965
Haijun Gu, Wei Liu, Xiaohui Zhao

Ultra-Low Power CMOS Charge-Sensitive Preamplifier 975
Zhou Yunbo, Yu Zongguang, Yang Yu

Reduction of Power Consumption in Wireless Sensor Networks for Railway Disaster Prevention and Safety Monitoring System 983
Tao Wu, Yusong Yan, Xi Chen

The Integration Scheme of the Internet of Things and TD-SCDMA Network 993
Zhengyong Yu

Study on Product Data Structure and Realization of Model Based on PLM 1001
Xianghui Zhan, Xiaoda Li, Xiangyang Lui

The Research of DNS Safe Analysis and Defense Technology 1009
Shao Ming-zhu, Sun Bo, Xie Rui-yun

Study on the Dangerous Chemicals Transport Vehicles Monitoring System Based on RFID Technique 1015
Wang Jianxin, Wang Tong, Ma Yunfei, Tan Jialei

Implementation of Neural Network Backpropagation in CUDA 1021
Jinfeng Liu, Lei Guo

Research and Application of Serial Communication Modules Based on VB 1029
Zhang Tao

Design and Application of Export Model 1035
Lei Xiong

A Processing Pipeline for Virtual 3D Urban-Based Web3D Author Tools . . . 1039
Zhang Wei

Categorization of Medical Documents Using Hybrid Competitive Neural Network with String Vector, a Novel Approach 1045
Seyyed Mohammad Reza Farshchi, Mahdi Yaghoobi

Study of Computer Room Network Monitoring Management System 1055
Dianjun Xu

The Application of Online Traffic System on the Wireless Internet 1063
Zhiqiang Yao, Chiyuan Li, Ke Xu, Han Xia

A New Trusted Monitoring Algorithm to IA Based Distributed System 1071
Shao Liang, Wu Fang

A Study of the Impact of Electronic Processing Time on Performance in Optical Burst Switching Networks 1081
G.M. Li, Z.K. Sun, S.H. Geng

Study on Stayed-Cable Health Monitoring 1091
Z.K. Sun, G.M. Li, S.H. Geng

Research on the Information Library for Component Testing 1099
Ying Jiang, Lin Jin, Xiao-Dong Fu, Yu-Ping Liu

Author Index 1109

Advantages of Cross-Border Acquisition under International Financial Crisis

Zuoling Nie, Yuhai Su, and Chong-jin Wang

Shandong University of Finance and Economics, Jinan, P.R. China

Abstract. During the global financial crisis, the volume of Chinese FDI is enlarging rapidly and acquisition has already been our major entry mode. To understand this phenomenon, we constructed a market entry game model for a multinational enterprise. The finds include: (1) The enterprise prefers green-field investment when its export cost rises; (2) green-field indeed is a credible threat for other enterprises in the target market even if it is not a desirable option when the effective demand decreases. Therefore, rising of export cost and decreasing of effective demand provide more negotiating power to acquire the target firm.

Keywords: FDI, Acquisition, Game theory analysis.

1 Introduction

In 2008, international capital flow slowed down and the amount of cross-border direct investment shrank dramatically. Under the unsatisfactory background, China foreign direct investment (FDI) is experiencing flourish adversely. The total amount of FDI reaches 52.1 billion dollars in 2008, which almost doubles over that of 2007. In the first 3 quarters of 2009, the FDI total amount reaches 32.9 billion dollars which rises by 0.5% over the same period. Other interesting phenomenon is that acquisition has already become an important FDI mode. The FDI reaches 20.5 billion dollars through acquisitions in 2008, which accounts for 50.4% of the current total FDI in China [1]. The scale and mode of FDI is experiencing obvious change, and the economy crisis which was originated from U.S. should be the trigger event for these changes. Of course, it is difficult to find a sound theoretical explanation for these changes. We are trying to construct a game model to provide a theoretical analysis frame for this phenomenon.

2 Reviews of the Literatures

There is abundant research on the FDI modes. Dubin (1976) discovered that U.S. firms tend to favor green-field investments if the firm size was large, targeting a developing country and had previously acquired foreign experience. [2] However, Zejan(1990) noted that product line diversification encouraged, and industry growth discouraged cross-border acquisition. Furthermore he indicated that experience is

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1–6.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

insignificant and that takeovers become more common because of growing instability and uncertainty. Hennart and Park (1993) found that Japanese investors in the U.S. favor the use of acquisition rather then green-field investment if the target market is characterized by high scale economies and high concentration levels.[3]Focusing on Japanese manufacturing companies entering the United States, O´hUallacha´in and Reid (1997) discovered that Japanese firms broadly disperse, their acquisition of American targets while their green-field investments were mainly attracted by interregional distribution of domestic manufacturing. Thus, they conclude that agglomeration features tend to have an impact on the firms choice. [4]

Although the above mentioned empirical study is not based on the data of Chinese enterprise, for a certain degree, it can explain the change of China FDI scale and structure. Since 1978, Chinese enterprises experienced huge change, for example, some of them mastered developed management skills and set up sound operating organization, according to the study of Andersson and Svensson (1996), all this change can partly explain the latest development of acquisitions in China. Another example is, more and more Chinese enterprises realized the instability and uncertainty of FDI due to the global financial crisis, so it is not surprise to find that Chinese enterprises prefer to take the forms of cross-border acquisition according to the study of Zejan (1990). There are abundant of relevant empirical study to support the phenomenon in China. Xu(2009) introduced in detail the innovation of Chinese enterprises FDI entry mode and dynamic development process, and proposed the corresponding suggestion based on the analysis of shortcomings of the FDI innovative modes.

In this article, we take normative method to analyze the choice of FDI entry mode. Buckley and Casson(1998) showed that some extra expenses played important roles in the choice of entrance forms.[5] The idea of the article is structured as the following: to simulate the effect of financial crisis on the real economy, we adjust certain parameter to demonstrate the determinants of the FDI entry mode. We try to resolve the following 2 questions, one is why the scale of China FDI is enlarging rapidly under the background of global financial crisis, and the other is why acquisition has already been a major entry mode under the financial crisis.

3 A Market Entry Game Model

3.1 Assumption

In our model, we assume that there are two local firms (Firm 2, 3) in the host country, and a single multinational firm (Firm 1) that considers how to enter this country's market. Firm 1 has 3 options: acquisition, green-field investment, exporting goods. If the multinational proposes a merger, it makes a take-it-or-leave-it offer to one of the(identical) local firms. Since the local firms are ex ante symmetric, it does not matter which local firm will be the potential target. We will label the potential target as firm 2, leaving firm 3 as an independent producer in all scenarios. In case of an acquisition, only firm 1 survives, firm 2 becomes firm 1's division.

3.2 The Model

We construct a game model according to the above mentioned assumption: firm 1 offers a payment to acquire firm 2, firm 2 either accepts or rejects, the order of moves in the game is as follows: firm 1 makes local firm 2 an acquire proposal. Firm 2 either rejects or accepts this proposal. If the offer is rejected, firm 1 chooses whether to engage in green-field investment or to export. In case of an accepted acquisition, firm 1 determines cost-reducing investments for both divisions. Finally, all independent firms choose output levels as Cournot competitors. Note that letting firm 1 first make an acquisition proposal does not restrict its ability to choose green-field investment or exporting.

Due to quadratic, quasi-linear preferences in the host country, the inverse demand function is given by p=a−Q, with p denoting the equilibrium price for an aggregate supply of Q. The marginal cost of production without any cost saving by an acquisition is equal to c with c>a. However, if firm 1 serves the market by exports, an additional trade cost of size t per unit of exports arises, where t≤ (a−c) / 3. The latter assumption will guarantee that exporting yields non-negative profits. Furthermore, we assume that green-field investment requires a fixed cost of size F. We make the simplifying assumption that export and acquisition do not involve fixed costs. There does not involve other costs except for the above mentioned costs. In the following section we discuss in more detail about the three modes, and solve the equilibrium outputs and profits for each.

If the multinational exports, their production levels and profits are as following:

$$q_1^T = \frac{a-c-3t}{4}, q_2^T = \frac{a-c+t}{4}, q_3^T = \frac{a-c+t}{4} \tag{1}$$

$$\pi_1^T = \frac{(a-c-3t)^2}{16}, \pi_2^T = \pi_3^T = \frac{(a-c+t)^2}{16} \tag{2}$$

Where qi and π_i denote individual production and profits of firm i. The super-script T denotes the exporting regime.

Green-field investment, denoted by the superscript GF, allows the multinational to produce with marginal cost c in the host market. The individual equilibrium production levels of all three firms coincide, i.e.

$$q_1^{GF} = q_2^{GF} = q_3^{GF} = \frac{a-c}{4} \tag{3}$$

And the profits of the multinational differ only by the fixed cost F from the profits of the local firms:

$$\pi_1^{GF} = \frac{(a-c)^2}{16} - F, \pi_2^{GF} = \pi_3^{GF} = \frac{(a-c)^2}{16} \tag{4}$$

In case of an acquisition, denoted by the superscript M, the multinational is able to determine investments for both divisions1 and 2. On the product market, firm 1 (i.e., the parent firm) competes only with firm 3.Within this duopoly, the respective output and profit levels (gross of any investment costs) are equal to

$$q_1^M = \frac{a+c-2\zeta}{3}, q_3^M = \frac{a-2c+\zeta}{3}, \tag{5}$$

$$\pi_1^M = \frac{(a+c-2\zeta)^2}{9}, \pi_3^M = \frac{(a-2c+\zeta)^2}{9} \tag{2}$$

Where ζ is the marginal cost after investments I1 and I2 have been made by the respective divisions. Due to the synergy effect, the marginal cost of firm 1 after acquisition is smaller than that of independence. The marginal cost ζ is equal to

$$\zeta = c - \gamma(I_1 + I_2) \tag{7}$$

The parameter γ measures the general efficiency of investments. If δ measures the marginal cost of investment, according to equation (6) and (7), the profits including investment costs are equal to

$$\pi_1^M = \frac{\left[a-c+\gamma(I_1+I_2)\right]^2}{9} - \frac{\delta I_1^2}{2} - \frac{\delta I_2^2}{2} \tag{8}$$

And the first-order conditions yield the equilibrium investment levels

$$I_1^* = I_2^* = \frac{2(a-c)\gamma}{9\delta - 4\gamma^2} \tag{9}$$

Which implies equilibrium profits of the new firm is

$$\pi_1^{M^*} = \frac{(a-c)^2 \delta}{9\delta - 4\gamma^2} \tag{10}$$

Profits increase with the efficiency parameters γ and decrease with the cost parameter δ. Due to

$$\pi_1^{M^*}(I_1 = I_2 = 0) = \lim_{\delta \to \infty} \pi_1^{M^+} = \frac{(a-c)^2}{9} \tag{11}$$

There is a positive lower bound for the profit of the merged firm. Note also that this profit is gross profit, not taking into account the acquisition price of the target firm. The acquisition price, in turn, depends on the credibility of green-field in-vestment or exporting, respectively. In the acquisition regime, we do not take ac-count into the product and profit of firm 2, because the firm 2 has already been the division of firm 1 and get the offer price N as the compensation. Although the product decision of firm 3 has influence on the firm 1, as a profit maximization individual, the decision of firm 3 is determined by the market microstructure.

3.3 The Equilibrium Analysis of FDI Entry Mode

The profit indicates as the one period profit, if the market is assumed to survive T period, we should consider the net present cash flow, and then deduct F of N from it. We can use the game tree to show the above analysis process (see figure 1). Compared A_2 with A_3, we get the conditions that multinational prefer the mode of green-field investment to exporting into the foreign market.

$$F \le \bar{F} = \frac{3tcK(a-c)}{4(a+3c-3t)} \tag{12}$$

That is, when the fixed cost of green-field investment is enough small, the multinational prefers the green-field investment to exporting.

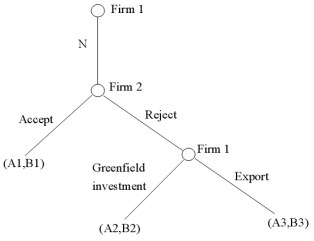


Fig. 1. The game tree

Where,

$$A_3 = \frac{a-c-3t}{4c}, B_3 = K\pi_2^T = K\frac{(a-c+t)^2}{16}$$

$$A_2 = \frac{(a-c)^2 K - F}{16} \Big/ \left(\frac{a-c}{4}cK + F\right), B_2 = K\pi_2^{GF} = K\frac{(a-c)^2}{16}$$

$$A_1 = \left[\frac{(a-c)^2 K - N}{9} \Big/ \frac{a-c}{3}cK + N\right] - N, B_1 = N$$

It is easy to find that the larger t is, the larger $\bar{F}$ is. In other words, the multinational more prefers green-field investment when the exporting cost is larger. We can get the following equation

$$\frac{\partial \bar{F}}{\partial a} = \frac{12tcK(4c-3t)}{4(a+3c-3t)^2} \tag{13}$$

When $t > 3c/4$, the first order differentiates of $\bar{F}$ relative to a is negative, which means that $\bar{F}$ negatively changes with a.

In the second phase of game, depending on the balance between the offering price N and the expected profit after the rejection, firm 2 makes decision whether to accept the acquisition offering from firm 1. If firm 1 chooses "green-field investment" in the third phase, the minimum offering price for firm 2 is N, $N \geq K\pi_2^{GF} = K(a-c)^2/16$; If firm 1 choose "exporting" in the third phase, then the minimum acceptance price is N', $N' \geq K\pi_2^T = K(a-c+t)^2/16$

In the first phase of the game, due to a>c, there is $N < N'$. For firm 1, if equation (12) is satisfied, "green-field investment" will be a credible threat for firm 2. According to this logic, in the first phase, firm 1 could choose one lower offering price, $N = K(a-c)^2/16$. The reason is that firm 2 will compare two variables, one is offering price and the expected profit it could get if firm 1 takes the mode of green-field investment after refection, $K\pi_2^{GF} = K(a-c)^2/16$; if equation (12) is not satisfied, to keep the success of the acquisition, the offering price must be larger than $K(a-c+t)^2/16$.

3.4 The Choice of FDI Entry Mode

How does the global financial crisis affect FDI? Firstly, the crisis caused the rise of unemployment in some countries and the trade protectionism begins to surge, as a consequence, the exporting cost (t) rises. The larger t has positive effect on B3 and negative effect on A3. However, the equilibrium of the game is dependent on the choice of firm 1, as long as equation (12) is satisfied, firm 1 will not choose export in the third phase. The rationale is "the larger t is, the larger $\overline{F}$ is", the multinational firm prefers green-field investment to exporting, and the credibility of threat is enhanced consequently.

Secondly, insufficient effective demand is the direct consequence of the global economy crisis. From the model, we know that the first order differentiate of $\overline{F}$ relative to a is negative if $t > 3c/4$. That means that $\overline{F}$ will become larger correspondingly when a becomes smaller. The equation (12) is satisfied if t and $\overline{F}$ become larger under financial crisis, in other words, "green-field investment" becomes a credible threat, and the offering price could be reduced to $K(a-c)^2/16$.

4 Conclusion

As the second large exporting country in the world, the exporting cost for Chinese firms will rise consequently. According to our analysis, to broaden international market, Chinese firms should shift to FDI under such background. So it is not surprise to find that China FDI shows dramatic high growth speed. Due to destroy of financial crisis, the effective demand in some target markets begins to fall down. Although green-field investment in the target market is not ideal, now it becomes a credible threat to the existing firms, at the same time, rise of the exporting cost makes the threat more credible, the acquisition price could decrease greatly correspondingly, which will provide larger negotiating space for the cross-border acquisition. Generally people contribute the rise of cross-border acquisition to the fall of asset price, while we provide another view to explain the phenomenon in our article by constructing the game model.

Acknowledgement. The research is financed by the science research team of Shandong "Tai Mount Scholar" in subject of Finance and Public Finance.

References

[1] United Nations Conference on Trade and Development, World Investment Report (2008)
[2] Dubin, M.: Foreign acquisitions and the spread of the multinational firm. Graduate School of Business Administration. Harvard University, Cambridge (1976)
[3] Hennart, J.F., Park, Y.R.: Greenfield vs. acquisition: The strategy of Japanese investors in the United States. Management Science 39(9), 1054–1070 (1993)
[4] OhÚallacháin, B., Reid, N.: Acquisition versus green field investment: The location and growth of Japanese manufacturers in the United States. Regional Studies (1997)
[5] Buckley, P.J., Casson, M.: Analyzing foreign market entry strategies: Extending the international approach. Journal of International Business Studies 29(3), 539–562 (1998)

A New Extension of LBP for Texture Classification

Kunlun Li and Pan Wang

College of Electronic and Information Engineering, Hebei University, Baoding, 071002, China
likunlun@hbu.edu.cn, wangpanhbu@126.com

Abstract. LBP is widely used for texture analysis because of its robustness to illumination and rotation changes and low computational complexity. LBP feather is obtained locally with the defect of losing global information. In this paper we proposed a novel solution, Tri-LBP, based on the combination of LBP, LBPV and LBP-HF according to Tri-training strategy. It contains both local and global feature, meanwhile, the semi-supervised learning of Tri-training can improve classification performance for little labelled sambles and large number of unlabelled ones. Our approach is evaluated by using texture images from the Outex database. The experiments show that the proposed method is effective for gray-scale and rotation invariant texture classification. The comparison with current methods illustrates that our proposal obtains a better performance. The application of Tri-training improved the classification accuracy obviously.

Keyword: Texture, LBP, Tri-training, Classification.

1 Introduction

Texture classification is a quite active research topic in computer vision and pattern recognition. One of the difficulties concentrates in image texture feature extraction. The texture image is bound with material itself, reflectivity, illumination, the camera and viewing position. Generally, Current main methods can be classified into four types: statistics, model, structure and signal processing methods.

Statistical methods are based on gray value of centre pixels and its adjacent area pixels, and obtain statistical characteristics of the areas texture. Model based methods suppose texture as distribution models varying with parameters, and take parameters as the texture characteristics. Structure methods, based on texton, take different types of texton, different direction and numbers as characteristics of the texture. Signal processing methods are different from the other three methods, instead of extracting texture feature from the image domain directly, but from the transform domain. The four basic methods have their own advantages and are appropriate for different situation respectively.

Local binary patterns (LBP), proposed by Ojala, take the symbol information of the distance between the neighborhood pixels as the texture characteristics [1,2]. Because of its robustness to illumination and rotation changes, low computational complexity,

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 7–12.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

LBP has many successful applications in texture classification, face recognition and automated cell phenotype image classification, etc.

However, LBP is just a local operator, which only considers some ideal situation and loses some important global information, and may fail to describe some texture patterns. Several variants have been proposed in the last few years to improve the distinguishing ability. Many scholars research and improve the traditional LBP in order to realize good classification results. This paper put forward a novel variant, Tri-LBP, which is based on Tri-training to combine local and global feathers. At the same time, the application of semi-supervised learning makes fully use of the labeled and unlabelled texture. The experiments show that classification accuracy of our method has signally improved.

2 LBP and Its Extensions

LBP is a simple and effective approach, it is robust to gray-scale and rotation variations, and discriminate a large range of textures efficiently. The basic idea is as follows.

Given a central pixel C in the image and texture is defined as:

$$T = t\left(g_c, g_0, g_1, \cdots, g_{P-1}\right) \tag{1}$$

where g_c is the gray value of the central pixel, g_P is the value of its neighborhood, P is the number of neighbors and R is the radius of the neighborhood.

At first, in order to achieve the invariant to the change of illumination, just the signs of the differences instead of their exact values are considered. At the same time, the invariance to the scaling of the gray scale is achieved. LBP code of the pixel is obtained by comparing its intensity with their neighbourhoods.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s\left(g_P - g_c\right)2^P \quad \text{where} \quad s\left(x\right) = \begin{cases} 1, if\ x > 0; \\ 0, otherwise \end{cases} \tag{2}$$

The second step is achieving rotation invariance. The same texture may produce different LBP codes under different rotation. According to Equation (3), LBP is transformed to the rotation invariant bins: LBPri.

$$LBP_{P,R}^{ri} = \min\left\{ROR\left(LBP_{P,R}, i\right) \middle| i = 0,1,\cdots,P-1\right\} \tag{3}$$

In order to improve the ability to discriminate, a uniformity measure U is introduced, which corresponds to the number of spatial transitions (bitwise 0/1 changes) in the 'pattern'.

$$U\left(LBP_{P,R}\right) = \left|s\left(g_{P-1} - g_c\right) - s\left(g_0 - g_c\right)\right| + \sum_{p=1}^{P-1}\left|s\left(g_{P-1} - g_c\right) - s\left(g_0 - g_c\right)\right| \tag{4}$$

So the final texture feature is defined as:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s\left(g_P - g_c\right) 2^P, U\left(LBP_{P,R}\right) \le 2 \\ P+1, otherwise \end{cases} \tag{5}$$

In the traditional LBP descriptor, calculation of the histogram assigns the same weight 1 to each LBP pattern without considering the LBP variance of the local region. Practically, the variance is important global information of the texture feature. Usually, the higher frequency of texture region is, the higher variance is. They contribute more to the discrimination of texture images. Therefore, LBPV make use of the variance as an adaptive weight to adjust the contribution of the LBP code in histogram calculation [3].

The tradition LBP histogram has proven to be a widely applicable image feature, meanwhile it exists the defect of losing the relative distribution of local orientations. LBP-HF constructs rotationally invariant features from the LBP histogram representing the whole region. Instead of computing invariant independently at each pixel location, the invariants are constructed globally to overcome the defect of losing the relative distribution. Meanwhile, it is invariant to rotations of input image and can describer texture more precisely [4].

3 Extensions Based on Tri-training: Tri-LBP

Tri-training is a new co-training algorithm proposed by Zhou et al. [5], which uses three classifiers to avoid the tedious cross validation process and has not other constraints about the attribute set and classifiers. Tri-training has two main techniques application. One is semi-supervised learning by using the limited labeled data to predict the unlabeled to expand the training set and another is ensemble learning by exploiting three different classifications and according to the same prediction results to label example and expand training sets. Tri-training combines those two gracefully and overcomes the defect of instability in semi-supervised learning.

The basic idea of Tri-training is follows. Assume *L* as the labeled example set and *U* as the unlabeled example set. Three classifiers are trained from *L*, namely *C1, C2, C3*. Suppose *D* is one example of *U* set. If *C2* and *C3* agree on the labeling of example *D*, then *D* can be labeled corresponding category for *C1*. So training set *S1* is expanded as $L \cup \{D|D \in U \text{ and } C2(x)=C3(x)\}$ and training C1 classification again. Similarly, the training set of *C2* and *C3* are expanded. Repeat these until *C1, C2* and *C3* do not change. Training process is over.

This paper utilizes several different algorithms to process texture images and three different sets of texture characteristic data is obtained. Three classifiers are obtained by training three different sets. Based on Tri-training integration strategy, the three classifiers are combined for texture image classification.

Original LBP have the defect of losing global spatial information, while global features carry some important texture information. LBPV combine globally rotation invariant matching with locally variant LBP features and exploits the local contrast information to improve the performance [3]. LBP-HF compute texture features from the histogram which represents the whole region and retains the relative distribution of local orientations [4]. In this paper, we propose a method to combine these three methods based on Tri-training strategy. The outline of our method is as follows.

Step1: The uniform LBP bins and rotation invariant bins are extracted by using the median value of the neighborhood. The conventional LBP features are named SET1. The LBPV and LBP-HF features are named SET2 and SET3 respectively.

Step2: According to tri-training, three classifiers are needed. In our method, we use the nearest neighborhood classifier with chi-square distance. Three nearest classifiers (named *C1, C2, C3*) are trained using the features extracted in Step1 respectively.

Step3: The classifier *C1* is initially trained. If the prediction of *C2* and *C3* is equal to one unlabeled data, label the data and add it to the training set of *C1*. Similarly, *C2* and *C3* training sets are extended respectively.

Step4: Repeat Step3 till all the training set is unchanged. If the time is more than 50, stop the repeating process.

4 Experimental Results

In order to verify the effectiveness of the proposed algorithm, we conducted experiments on a large and comprehensive texture databases: the Outex database [6]. Compared with the widely used Brodatz database, these image is includes 319 classes of textures captured using six spatial resolutions (100, 120, 300, 360, 500 and 600 dpi) and nine rotation angles (0, 5, 10, 15, 30, 45, 60, 75 and 90) under three illuminations, and can effectively verified the robustness to rotation and illumination variation of proposed methods.

Our method is verified from two different aspects: rotation invariant verification and discriminative nature verification. For simplicity, we set *P=8, R=1* in the experiments, although better accuracy could be achieved if these parameters are tuned appropriately.

Experimental environment: Intel Core 2.53 GHz CPU, 2 GB memory, WIN7 OS, MATLAB V7.8.

Experiment 1: Rotation Invariant Verification

In this experiment, we chose 30 classes texture from the Outex to verify the proposed. The total number of the images is 30*9*3=810. The training set for each class are under single 0^0 orientation. Hence, the rest of images build the testing set. The classification results compared with different operators are listed in Table 1.

Experiment 2: Illuminance Invariant Verification

In this experiment, we chose 60 class texture under single spatial resolutions from the Outex to verify the proposed. The total number of the images is 60*3*9=1620. The

training sets are all different orientation under 'inca' illumination and the test sets consist of the rest images. The classification results compared with different operators are listed in Table 1.

Table 1. The classification results

Method	Exp. 1/%	Exp. 2/%
LBPu2	61.63	73.31
LBPriu	82.07	64.84
LBP-HF	76.22	73.56
LBPV	70.59	74.71
Tri-LBP	84.37	89.59

Our experiments further proved that LBPriu has the better performance in rotation invariant than LBPu2 and LBPu2 has better discriminability than LBPriu. We chose the better one combine with LBPV and LBP-HF according to Tri-training. The results of two verification show our method can improved the classification accuracy obviously.

5 Conclusion

In this paper, we present a new texture classification method, Tri-LBP, which combines original LBP, LBPV and LBP-HF descriptor based on Tri-training. Tri-LBP combines local and global feathers, and the semi-supervised of Tri-training improves performances of classification. We evaluated our method both in illuminance invariant and rotation invariant. Experimental results validate this method can achieve higher classification accuracy than the other methods with a great robustness. According to Tri-training, our method can combine all sorts of texture description operator different from each other. In the further work, we can try other different fusion to fully describe various texture images with less information losing.

Acknowledgement. This work is supported by the National Natural Science Foundation of China under Grant (No.61073121), Natural Science Foundation of Hebei Province of China (No. F2009000215) and Biomedical Multidisciplinary Research Center of Hebei University (No.BM201102).

References

1. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distribution. Pattern Recognition 29(1), 51–59 (1996)
2. Ojala, T., Pietikainen, M., Maeenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987 (2002)

3. Guo, Z., Zhang, L., Zhang, D.: Rotation invariant texture classification using LBP variance (LBPV) with global matching. Pattern Recognition 43(3), 706–719 (2010)
4. Ahonen, T., Matas, J., He, C., Pietikäinen, M.: Rotation Invariant Image Description with Local Binary Pattern Histogram Fourier Features. In: Salberg, A.-B., Hardeberg, J.Y., Jenssen, R. (eds.) SCIA 2009. LNCS, vol. 5575, pp. 61–70. Springer, Heidelberg (2009)
5. Zhou, Z.H., Li, M.: Tri-training: Exploiting unlabeled data using three classifiers. IEEE Transactions Oil Knowledge and Data Engineering 17(11), 1529–1541 (2005)
6. Ojala, T., Mäenpää, T., Pietikäinen, M., Viertola, J., Kyllönen, J., Huovinen, S.: Outex–new framework for empirical evaluation of texture analysis algorithm. In: Proc. International Conference on Pattern Recognition, pp. 701–706 (2002)

Adaptive Backstepping Control for Nonlinear Systems Using Support Vector Regression

Liu Yinan*, Zhang Shengxiu, Cao Lijia, and Zhang Chao

Xi'an Research Inst. of Hi-Tech, Hongqing Town, Xi'an, 710025, PR China
lyn8307@gmail.com

Abstract. A nonlinear adaptive backstepping control approach is designed for a class of n-th order nonlinear systems. Support Vector Regression (SVR) is employed to adaptively approximate the unknown nonlinear functions composed of unknown uncertainties and disturbances. Unlike neural networks, no number of hidden units has to be determined for the controller and that no centers have to be specified for the Gaussian kernels when applying Mercer's condition. The curse of dimensionality is avoided in comparison with defining a regular grid for the centers in classical radial basis function networks. The closed-loop system is guaranteed to be bounded and tracking errors are also proved to converge exponentially to a small residual set around the origin by Lyapunov theory. Simulation results demonstrate the effectiveness of the approach proposed.

Keywords: adaptive control, backstepping, support vector regression, nonlinear system.

1 Introduction

In the past decades, numerous approaches have been proposed for the design of nonlinear control systems. Among these, adaptive control is a major design methodology that deals with modeling uncertainties in nonlinear (and linear) systems by on line tuning of parameters. And a lot of researchers have presented several important results on it [1-3]. But during these researches, some restrictions on the plants had to be made in order to guarantee the global stability, such as matching condition, extended matching condition, or growth conditions on system nonlinearities. In order to overcome these restrictions, adaptive backstepping was provided [4]. The essence of backstepping design is to select some appropriate functions of state variables recursively as pseudo control inputs for lower dimension subsystems of the overall system. It can guarantee global stabilities, tracking and transient performance for a broad class of strict-feedback system [5,6]. However, some drawbacks of the backstepping approaches are also exists, such as the nonlinear systems must be linearly parameterized or uncertain parameters and that the nonlinear functions must be exactly known [7,8].

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 13–23.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

In an attempt to remove the above drawbacks, neural networks for system identification or identification based control combined with backstepping have recently been proposed. It is now a well-established fact that a multilayer feedforward neural network or a radial basis function (RBF) neural network can approximate any well-behaved nonlinear function with suitable accuracy [9]. So with the help of neural networks, the linearity-in-the-parameter assumption of nonlinear function and the determination of regression matrices can be avoided [10,11].

On the other hand, kernel-based learning method such as support vector machine (SVM) or support vector regression (SVR) are featured as small samples learning, high generalization ability, and can effectively avoid overfitting phenomenon, the local minimum point, the 'dimension disaster' and so on [12-14]. In the case of RBF networks that mostly used combining with backstepping, one has a curse of dimensionality when one defines a regular grid for the centers (hidden units) in state space. In the SVM control case, the centers will follow from the optimal trajectory that one seeks. Furthermore, standard neural network approaches always work in a primal weight space, while in SVM methodologies the computations are done in a dual space such that the number of unknowns equals the number of training data points (and not the number of weights in the primal space, which can be infinite dimensional) [15].

In this paper, a nonlinear adaptive backstepping control approach using SVR is proposed. The rest of this paper is organized as follows. System formulation and preliminaries are given in Section 2. Section 3 proposed the nonlinear adaptive backstepping controller using SVR approximation. In Section 4, the stability analysis of closed-loop system is presents. The effectiveness of the adaptive backstepping control scheme using SVR proposed in this paper is demonstrated via an example in Section 5. Finally, some concluding remarks are given in Section 6.

2 System Formulation and Preliminaries

2.1 Problem Formulation

Let state variables $\boldsymbol{x}_1 \in \boldsymbol{R}^{n_1}$, $\boldsymbol{x}_2 \in \boldsymbol{R}^{n_2}, \cdots, \boldsymbol{x}_n \in \boldsymbol{R}^{n_n}$ and system input $\boldsymbol{u} \in \boldsymbol{R}^m$, with $m \geq n_n$ .Consider the following class of n-th order multiple-input multiple-output (MIMO) nonlinear systems with uncertainties and disturbances.

$$
\begin{aligned}
\dot{\boldsymbol{x}}_1 &= \boldsymbol{f}_1(\bar{\boldsymbol{x}}_1) + \Delta \boldsymbol{f}_1(\bar{\boldsymbol{x}}_1) + \boldsymbol{g}_1(\bar{\boldsymbol{x}}_1)\boldsymbol{x}_2 + \Delta \boldsymbol{g}_1(\bar{\boldsymbol{x}}_1)\boldsymbol{x}_2 + \boldsymbol{d}_1 \\
&\vdots \\
\dot{\boldsymbol{x}}_i &= \boldsymbol{f}_i(\bar{\boldsymbol{x}}_i) + \Delta \boldsymbol{f}_i(\bar{\boldsymbol{x}}_i) + \boldsymbol{g}_i(\bar{\boldsymbol{x}}_i)\boldsymbol{x}_{i+1} + \Delta \boldsymbol{g}_i(\bar{\boldsymbol{x}}_i)\boldsymbol{x}_{i+1} + \boldsymbol{d}_i \\
&\vdots \\
\dot{\boldsymbol{x}}_n &= \boldsymbol{f}_n(\bar{\boldsymbol{x}}_n) + \Delta \boldsymbol{f}_n(\bar{\boldsymbol{x}}_n) + \boldsymbol{g}_n(\bar{\boldsymbol{x}}_n)\boldsymbol{u} + \Delta \boldsymbol{g}_n(\bar{\boldsymbol{x}}_n)\boldsymbol{u} + \boldsymbol{d}_n \\
\boldsymbol{y} &= \boldsymbol{x}_1
\end{aligned}
\tag{1}
$$

where $\bar{\boldsymbol{x}}_i = [\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \cdots \quad \boldsymbol{x}_i]^{\mathrm{T}}$ are assumed to be available for measurement, $\boldsymbol{y}$ denotes the system output. The functions $\boldsymbol{f}_i(\cdot)$ and $\boldsymbol{g}_i(\cdot)$, $i = 1, 2, \cdots, n$, are smooth nonlinear functions that are assumed to be known, $\Delta\boldsymbol{f}_i(\cdot)$ and $\Delta\boldsymbol{g}_i(\cdot)$ are smooth nonlinear functions caused by both parametric and nonparametric uncertainties, $\boldsymbol{d}_i$ denote the disturbances.

To ensure controllability, we will invoke the following assumption.

Assumption 1: There exists a constant g_0, such that for $i = 1, 2, \cdots, n$, each function $\|\boldsymbol{g}_i(\cdot)\| \geq g_0$. where $\|\cdot\|$ denotes the 2-norm of a vector or a matrix.

The control objective is to design an adaptive control input $\boldsymbol{u}$ so that the output $\boldsymbol{y}$ follows a desired trajectory $\boldsymbol{y}_c$ with the constraint that all signals in the closed-loop system are semi-globally uniformly ultimately bounded. Generally, the assumption that $\boldsymbol{y}_c$ and its derivatives are all existent and bounded is required. If there is a n-th order system, $\boldsymbol{y}_c$ and its derivatives up to the $(n+1)^{th}$ order are all bounded is usually required [16].

Assumption 2: The desired trajectory $\boldsymbol{y}_c$ and its derivative $\dot{\boldsymbol{y}}_c$ are continuous and bounded.

In addition, to design the backstepping controller, another assumption is required.

Assumption 3 [17]: for $i = 1, 2, \cdots, n$, each function $\boldsymbol{f}_i(\cdot)$ and $\boldsymbol{g}_i(\cdot)$ and their first partial derivatives are continuous and bounded on any compact set $\boldsymbol{D}_i \subset \boldsymbol{R}^i$.

2.2 Support Vector Regression

In the system (1), there are some unknown uncertainties $\Delta\boldsymbol{f}_i(\cdot)$ and $\Delta\boldsymbol{g}_i(\cdot)$ and disturbances $\boldsymbol{d}_i$. They can be combined to form an unknown nonlinear function $\boldsymbol{\Delta}_i$ as follows:

$$\begin{aligned} \boldsymbol{\Delta}_i &= \Delta\boldsymbol{f}_i + \Delta\boldsymbol{g}_i\boldsymbol{x}_{i+1} + \boldsymbol{d}_i, \quad 1 \leq i \leq n-1 \\ \boldsymbol{\Delta}_n &= \Delta\boldsymbol{f}_n + \Delta\boldsymbol{g}_n\boldsymbol{u} + \boldsymbol{d}_n \end{aligned} \tag{2}$$

To identify $\boldsymbol{\Delta}_i$, $i = 1, 2, \cdots n$, the control design presented in this paper employs support vector regression (SVR). And the identification model $f(\boldsymbol{x})$ can be expressed as:

$$f(x_k) = \sum_{i=1,(i \in SV)}^{N_{SV}} \zeta_i \kappa(x_k, x_i) + b + \varepsilon(x_k) \tag{3}$$

where x_k is the input vector of SVR, ζ_i is the weight value, and the data points x_i which are corresponding to non-zero values of ζ_i are called support vectors, N_{SV} denotes the numbers of support vectors. b is the bias term, $\varepsilon(x_k)$ is the so-called SVR approximation error, $\kappa(x_k, x_i) = \phi(x_i)^T \phi(x_j) = \kappa_{ij}$ is a kernel function. In this study, the following Gaussian radial basis kernel function is used

$$\kappa(x_i, x_j) = \exp(-\frac{(x_i - x_j)^T (x_i - x_j)}{\upsilon^2}) \tag{4}$$

(3) could also be expressed as

$$f(\boldsymbol{x}) = \boldsymbol{W}_i^{*\mathrm{T}} \boldsymbol{\kappa}_i(\boldsymbol{x}) + \boldsymbol{\varepsilon}(\boldsymbol{x}) \tag{5}$$

where $\boldsymbol{W}_i^* = [\zeta_{i,1}, \zeta_{i,2}, \ldots \zeta_{i,L_i}, b_i]^{\mathrm{T}}$, $\boldsymbol{\kappa}_i = [\kappa(x, x_{i,1}), \ldots, \kappa(x, x_{i,L_i}), 1]^T$, L_i is the number of the support vectors, $\boldsymbol{\kappa}_i$ is the extended kernel function including the element 1, $x_{i,j}$ is the support vector, and x is the input for the SVR. The schematic of the SVR architecture can be given as Fig. 1.

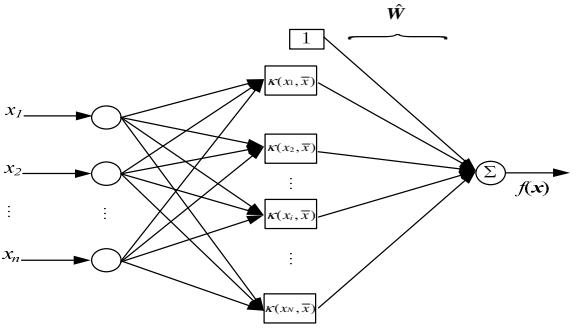


Fig. 1. Schematic of the SVR architecture

Theoretically, the SVR can approximate any continuous nonlinear function to any desired accuracy. And it satisfies that the approximation error $\|\varepsilon(\boldsymbol{x})\| \le \varepsilon_m$. To use SVR in backstepping, there is another assumption about weight vector $\boldsymbol{W}_i^*$.

Assumption 4: The weight vector $\boldsymbol{W}_i^*$ are bounded in the sense that

$$\left\|\boldsymbol{W}_i^*\right\|_F \le W_m \tag{6}$$

where W_m is known positive constant and $\|\cdot\|_F$ denotes the Frobenious norm of a matrix.

3 The Design of Adaptive Backstepping Controller

Step 1: Design a nominal control input $\boldsymbol{a}_1$. Recall first equation in (1) with (2), it can be written as:

$$\dot{\boldsymbol{x}}_1 = \boldsymbol{f}_1 + \boldsymbol{g}_1 \boldsymbol{x}_2 + \boldsymbol{\Delta}_1 \tag{7}$$

The tracking error vector is defined as $\tilde{\boldsymbol{x}}_1 = \boldsymbol{x}_1 - \boldsymbol{y}_c$. By treating $\boldsymbol{x}_2$ as a virtual control input and using the feedback linearization method, the nominal control input $\boldsymbol{a}_1$ is designed as follows:

$$\boldsymbol{a}_1 = -\boldsymbol{g}_1^{-1}(k_1 \tilde{\boldsymbol{x}}_1 + \boldsymbol{f}_1 + \hat{\boldsymbol{W}}_1^{\mathrm{T}} \boldsymbol{\kappa}_1 - \dot{\boldsymbol{y}}_c) \tag{8}$$

where k_1 denotes the designed positive constant; $\hat{\boldsymbol{W}}_1$ denotes the estimate of ideal weight matrix $\boldsymbol{W}_1^*$, and the estimate error is $\tilde{\boldsymbol{W}}_1 = \hat{\boldsymbol{W}}_1 - \boldsymbol{W}_1^*$. Select adaptive update law of SVR weight matrix as:

$$\dot{\hat{\boldsymbol{W}}}_1 = \boldsymbol{\Xi}_1 (\boldsymbol{\kappa}_1 \tilde{\boldsymbol{x}}_1^{\mathrm{T}} - \sigma_1 \hat{\boldsymbol{W}}_1) \tag{9}$$

where $\boldsymbol{\Xi}_1 \in \boldsymbol{R}^{l \times l}$ denotes the invertible positive gain matrix and $\sigma_1 > 0$ is a small design parameter called σ-modification coefficient.

Step i: Design a nominal control input $\boldsymbol{a}_i$, $i = 2, 3, \cdots n-1$. Consider

$$\dot{\boldsymbol{x}}_i = \boldsymbol{f}_i + \boldsymbol{g}_i \boldsymbol{x}_{i+1} + \boldsymbol{\Delta}_i \tag{10}$$

Define the tracking error vector $\tilde{\boldsymbol{x}}_i$ and the compensated tracking error signals $\boldsymbol{z}_i$ as:

$$\tilde{\boldsymbol{x}}_i = \boldsymbol{x}_i - x_{ic} \qquad \boldsymbol{z}_i = \tilde{\boldsymbol{x}}_i - \boldsymbol{\xi}_i \qquad i = 1, 2, \cdots n \tag{11}$$

where the variable $\boldsymbol{\xi}_i$ is the output of the following filter

$$\dot{\boldsymbol{\xi}}_i = -k_i \boldsymbol{\xi}_i + \boldsymbol{g}_i (\boldsymbol{x}_{i+1c} - \boldsymbol{a}_i) + \boldsymbol{g}_i^{\mathrm{T}} \boldsymbol{\xi}_{i+1} \tag{12}$$

with $\boldsymbol{\xi}_1(0) = 0$. For $i = n$, define $\boldsymbol{\xi}_n = 0$.

Then, the nominal control input $\boldsymbol{a}_i$ is designed as follows:

$$\boldsymbol{a}_i = -\boldsymbol{g}_i^{-1}(k_i \tilde{\boldsymbol{x}}_i + \boldsymbol{f}_i + \boldsymbol{g}_{i-1}^{\mathrm{T}} \boldsymbol{z}_{i-1} + \hat{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\kappa}_i - \dot{x}_{ic}) \tag{13}$$

where k_i denotes the designed positive constant; $\hat{\boldsymbol{W}}_i$ denotes the estimate of ideal weight matrix $\boldsymbol{W}_i^*$, and the estimate error is $\tilde{\boldsymbol{W}}_i = \hat{\boldsymbol{W}}_i - \boldsymbol{W}_i^*$. Select adaptive update law of SVR weight matrix as:

$$\dot{\hat{\boldsymbol{W}}}_i = \boldsymbol{\Xi}_i (\boldsymbol{\kappa}_i \tilde{\boldsymbol{x}}_i^{\mathrm{T}} - \sigma_i \hat{\boldsymbol{W}}_i) \tag{14}$$

where $\boldsymbol{\Xi}_i \in \boldsymbol{R}^{l \times l}$ denotes the invertible positive gain matrix and $\sigma_i > 0$ is σ-modification coefficient.

Step n: Design control input $\boldsymbol{u}$. Recall that

$$\dot{\boldsymbol{x}}_n = \boldsymbol{f}_n + \boldsymbol{g}_n \boldsymbol{u} + \boldsymbol{\Delta}_n \tag{15}$$

Tracking error vector is defined as $\tilde{\boldsymbol{x}}_n = \boldsymbol{x}_n - \boldsymbol{x}_{nc}$. The control input $\boldsymbol{u}$ can be designed as follows:

$$\boldsymbol{u} = -\boldsymbol{g}_n^{-1}(k_n \tilde{\boldsymbol{x}}_n + \boldsymbol{f}_n + \boldsymbol{g}_{n-1}^{\mathrm{T}} \boldsymbol{z}_{n-1} + \hat{\boldsymbol{W}}_n^{\mathrm{T}} \boldsymbol{\kappa}_n - \dot{\boldsymbol{x}}_{nc}) \tag{16}$$

where k_n , $\hat{\boldsymbol{W}}_n$, $\tilde{\boldsymbol{W}}_n$ are the same meaning as in **step i**. Select adaptive update law of SVR weight matrix like (14).

$$\dot{\hat{\boldsymbol{W}}}_n = \boldsymbol{\Xi}_n(\boldsymbol{\kappa}_n \tilde{\boldsymbol{x}}_n^{\mathrm{T}} - \sigma_n \hat{\boldsymbol{W}}_n) \tag{17}$$

where $\boldsymbol{\Xi}_n \in \boldsymbol{R}^{l \times l}$ denotes the invertible positive gain matrix and $\sigma_n > 0$.

4 Stability Analysis

In Section 3, the adaptive backstepping controller using SVR is designed. The stability of the close-loop system will be discussed in this section. The dynamics of the tracking error $\tilde{\boldsymbol{x}}_i$ and the dynamics of the compensated tracking error $\boldsymbol{z}_i$ will be given firstly.

4.1 Error Dynamics

The dynamics of tracking errors can be written as follows:

$$\begin{aligned}
\dot{\tilde{\boldsymbol{x}}}_1 &= \dot{\boldsymbol{x}}_1 - \dot{\boldsymbol{x}}_{1c} = \boldsymbol{f}_1 + \boldsymbol{g}_1 \boldsymbol{x}_2 + \boldsymbol{\Delta}_1 - \dot{\boldsymbol{x}}_{1c} \\
&= \boldsymbol{f}_1 + \boldsymbol{g}_1 \boldsymbol{a}_1 + \boldsymbol{W}_1^{\mathrm{T}} \boldsymbol{\kappa}_1 + \boldsymbol{\varepsilon}_1 - \dot{\boldsymbol{x}}_{1c} + \boldsymbol{g}_1(\boldsymbol{x}_2 - \boldsymbol{x}_{2c}) + \boldsymbol{g}_1(\boldsymbol{x}_{2c} - \boldsymbol{a}_1) \\
&= -k_1 \tilde{\boldsymbol{x}}_1 + \boldsymbol{W}_1^{\mathrm{T}} \boldsymbol{\kappa}_1 - \hat{\boldsymbol{W}}_1^{\mathrm{T}} \boldsymbol{\kappa}_1 + \boldsymbol{\varepsilon}_1 + \boldsymbol{g}_1 \tilde{\boldsymbol{x}}_2 + \boldsymbol{g}_1(\boldsymbol{x}_{2c} - \boldsymbol{a}_1) \\
&= -k_1 \tilde{\boldsymbol{x}}_1 + \boldsymbol{g}_1 \tilde{\boldsymbol{x}}_2 + \boldsymbol{g}_1(\bar{\boldsymbol{x}}_{2c} - \boldsymbol{x}_{2c}) - \tilde{\boldsymbol{W}}_1^{\mathrm{T}} \boldsymbol{\kappa}_1 + \boldsymbol{\varepsilon}_1 \\
\dot{\tilde{\boldsymbol{x}}}_i &= \dot{\boldsymbol{x}}_i - \dot{\boldsymbol{x}}_{ic} = \boldsymbol{f}_i + \boldsymbol{g}_i \boldsymbol{a}_i + \boldsymbol{\Delta}_i - \dot{\boldsymbol{x}}_{ic} + \boldsymbol{g}_i(\boldsymbol{x}_{i+1} - \boldsymbol{x}_{i+1c}) + \boldsymbol{g}_i(\boldsymbol{x}_{i+1c} - \boldsymbol{a}_i) \\
&= -k_i \tilde{\boldsymbol{x}}_i - \boldsymbol{g}_{i-1}^{\mathrm{T}} \boldsymbol{z}_{i-1} + \boldsymbol{W}_i^{\mathrm{T}} \boldsymbol{\kappa}_i - \hat{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\kappa}_i + \boldsymbol{\varepsilon}_i + \boldsymbol{g}_i \tilde{\boldsymbol{x}}_{i+1} + \boldsymbol{g}_i(\boldsymbol{x}_{i+1c} - \boldsymbol{a}_i) \\
&= -k_i \tilde{\boldsymbol{x}}_i - \boldsymbol{g}_{i-1}^{\mathrm{T}} \boldsymbol{z}_{i-1} + \boldsymbol{g}_i \tilde{\boldsymbol{x}}_{i+1} + \boldsymbol{g}_i(\boldsymbol{x}_{i+1c} - \boldsymbol{a}_i) - \tilde{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\kappa}_i + \boldsymbol{\varepsilon}_i \\
\dot{\tilde{\boldsymbol{x}}}_n &= \dot{\boldsymbol{x}}_n - \dot{\boldsymbol{x}}_{nc} = \boldsymbol{f}_n + \boldsymbol{g}_n \boldsymbol{u} + \boldsymbol{\Delta}_n - \dot{\boldsymbol{x}}_{nc} \\
&= -k_n \tilde{\boldsymbol{x}}_n - \boldsymbol{g}_{n-1}^{\mathrm{T}} \boldsymbol{z}_{n-1} + \boldsymbol{W}_n^{\mathrm{T}} \boldsymbol{\kappa}_n - \hat{\boldsymbol{W}}_n^{\mathrm{T}} \boldsymbol{\kappa}_n + \boldsymbol{\varepsilon}_n \\
&= -k_n \tilde{\boldsymbol{x}}_n - \boldsymbol{g}_{n-1}^{\mathrm{T}} \boldsymbol{z}_{n-1} - \tilde{\boldsymbol{W}}_n^{\mathrm{T}} \boldsymbol{\kappa}_n + \boldsymbol{\varepsilon}_n
\end{aligned} \tag{18}$$

Combine (11) and (12), the compensated tracking errors dynamics become:

$$\begin{aligned}
\dot{\boldsymbol{z}}_1 &= \dot{\tilde{\boldsymbol{x}}}_1 - \dot{\boldsymbol{\xi}}_1 = -k_1 \boldsymbol{z}_1 + \boldsymbol{g}_1 \boldsymbol{z}_2 - \tilde{\boldsymbol{W}}_1^{\mathrm{T}} \boldsymbol{\kappa}_1 + \boldsymbol{\varepsilon}_1 \\
\dot{\boldsymbol{z}}_i &= \dot{\tilde{\boldsymbol{x}}}_i - \dot{\boldsymbol{\xi}}_i = -k_i \boldsymbol{z}_i - \boldsymbol{g}_{i-1}^{\mathrm{T}} \boldsymbol{z}_{i-1} + \boldsymbol{g}_i \boldsymbol{z}_{i+1} - \tilde{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\kappa}_i + \boldsymbol{\varepsilon}_i \\
\dot{\boldsymbol{z}}_n &= \dot{\tilde{\boldsymbol{x}}}_n - \dot{\boldsymbol{\xi}}_n = -k_n \tilde{\boldsymbol{x}}_n - \boldsymbol{g}_{n-1}^{\mathrm{T}} \boldsymbol{z}_{n-1} - \tilde{\boldsymbol{W}}_n^{\mathrm{T}} \boldsymbol{\kappa}_n + \boldsymbol{\varepsilon}_n
\end{aligned} \tag{19}$$

4.2 Stability Properties

Now, consider the control Lyapunov function candidate:

$$V(t) = \frac{1}{2}\sum_{i=1}^{n}(z_i^{\mathrm{T}} z_i + tr(\tilde{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\Xi}_i^{-1} \tilde{\boldsymbol{W}}_i)) \tag{20}$$

The time derivative of the Lyapunov function is

$$\begin{aligned} \dot{V} &= \sum_{i=1}^{n}(z_i^{\mathrm{T}} \dot{z}_i + tr(\tilde{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\Xi}_i^{-1} \dot{\tilde{\boldsymbol{W}}}_i)) \\ &= -\sum_{i=1}^{n}(k_i \|z_i\|^2 - z_i^{\mathrm{T}} \tilde{\boldsymbol{W}}_i^{\mathrm{T}} \kappa_i + z_i^{\mathrm{T}} \boldsymbol{\varepsilon}_i + tr(\tilde{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\Xi}_i^{-1} \dot{\tilde{\boldsymbol{W}}}_i)) \\ &= -\sum_{i=1}^{n}(k_i \|z_i\|^2 + tr(\tilde{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\Xi}_i^{-1} (\dot{\hat{\boldsymbol{W}}}_i^{\mathrm{T}} - \boldsymbol{\Xi}_i \kappa_i z_i^{\mathrm{T}})) + z_i^{\mathrm{T}} \boldsymbol{\varepsilon}_i) \end{aligned} \tag{21}$$

Note the following inequality:

$$\begin{aligned} tr(\tilde{\boldsymbol{W}}_i^{\mathrm{T}} \boldsymbol{\Xi}_i^{-1} (\dot{\hat{\boldsymbol{W}}}_i^{\mathrm{T}} - \boldsymbol{\Xi}_i \kappa_i z_i^{\mathrm{T}})) &= tr(-\sigma_i \tilde{\boldsymbol{W}}_i^{\mathrm{T}} \hat{\boldsymbol{W}}_i) = tr(-\sigma_i \tilde{\boldsymbol{W}}_i^{\mathrm{T}} (\tilde{\boldsymbol{W}}_i + \boldsymbol{W}_i^*)) \\ &\le tr(-\sigma_i \tilde{\boldsymbol{W}}_i^{\mathrm{T}} \tilde{\boldsymbol{W}}_i + \frac{\sigma_i}{2} \tilde{\boldsymbol{W}}_2^{\mathrm{T}} \tilde{\boldsymbol{W}}_i) + \frac{\sigma_i}{2} \|\boldsymbol{W}_i^*\|_F^2 \\ &\le tr(-\frac{\sigma_i}{2} \tilde{\boldsymbol{W}}_i^{\mathrm{T}} \tilde{\boldsymbol{W}}_i) + \frac{\sigma_i}{2} W_m^2 \end{aligned} \tag{22}$$

And from Young's inequality, we have the following inequality.

$$z_i^{\mathrm{T}} \boldsymbol{\varepsilon}_i \le \|z_i\|^2 + \|\boldsymbol{\varepsilon}_i\|^2 / 4 \tag{23}$$

Combine Eqs. (22) and (23) with Eq. (21), the time derivative of the Lyapunov function could be written as:

$$\dot{V} \le -\sum_{i=1}^{n}(k_i^* \|z_i\|^2 + \frac{\sigma_i}{2} tr(\tilde{\boldsymbol{W}}_i^{\mathrm{T}} \tilde{\boldsymbol{W}}_i)) + \sum_{i=1}^{n}(\frac{\sigma_i}{2} W_m^2 + \|\boldsymbol{\varepsilon}_i\|^2 / 4) \tag{24}$$

where $k_i^* = (k_i - 1) > 0$.

Define $k = \min_{1 \le i \le n}(2k_i^*, \sigma^i)$, $c = \sum_{i=1}^{n}(\frac{\sigma_i}{2} W_m^2 + \|\boldsymbol{\varepsilon}_i\|^2 / 4)$. Finally, the following equation is obtained for the derivative of the chosen Lyapunov candidate function (20):

$$\dot{V} \le -kV + c \tag{25}$$

Eq. (25) implies that $\dot{V} < 0$, when $V \le c / 2k$. Multiplying Eq. (25) by e^{kt} yields

$$\frac{d}{dt}(V(t)e^{kt}) \le ce^{kt} \tag{26}$$

Integrating both sides of (26) over $[0,t]$, we obtain

$$0 \le V(t) \le \frac{c}{k} + [V(0) - \frac{c}{k}]e^{-kt} \tag{27}$$

Therefore, all signals of the closed-loop system are uniformly ultimately bounded. Furthermore, it means that if the designed positive constants k_i are chosen suitably, tracking errors will converge exponentially to a small residual set around the origin.

5 Simulation Studies

In this section, a simple simulation is presented to show the effectiveness of the approach proposed above. The model of the system is given as

$$\begin{aligned} \dot{x}_1 &= 0.5x_1 + (1 + 0.1x_1^2)x_2 \\ \dot{x}_2 &= x_1 x_2 + [2 + \cos(x_1)]u \\ y &= x_1 \end{aligned} \tag{28}$$

where x_1 and x_2 are states, and y is the output of the system, respectively. The initial condition is $x_0 = [x_{10}, x_{20}]^T = [1,0]^T$ and the desired reference signal of the system is $y_d = \sin(t)$.

First, in order to obtain the original parameters of SVR, the input-output data pairs are accumulated for 20 seconds and the sampling rate is set to 0.01s. The design parameters C, ε, σ for SVR are set to $1000, 0.01, 2$. Then the design parameters of the controller are $k_1 = k_2 = 2.5$, $\Xi_1 = \Xi_2 = diag\{2\}$, $\sigma_1 = \sigma_2 = 0.4$.

Fig. 2 shows the result of applying controller proposed above for tracking desired signal y_d with no uncertainties. As shown in it, the system output shows the excellent tracking performance. And Fig. 3 shows the trajectory of the controller.

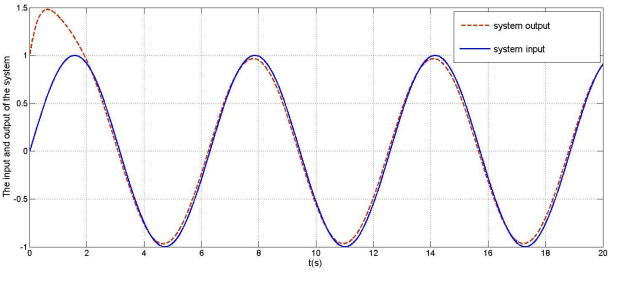


Fig. 2. The input and output of the system

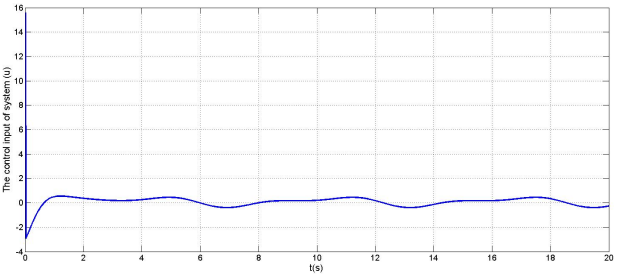


Fig. 3. The trajectory of the controller

In order to validate the effectiveness of the proposed adaptive control approach, the uncertainties and disturbances are added to the system and the model of system as in the following manner:

$$\begin{aligned} \dot{x}_1 &= 0.5x_1 + (1+0.1x_1^2)x_2 + \Delta \times \sin(\omega t) \\ \dot{x}_2 &= x_1 x_2 + [2+\cos(x_1)]u + \Delta \times \sin(\omega t) \\ y &= x_1 \end{aligned} \tag{29}$$

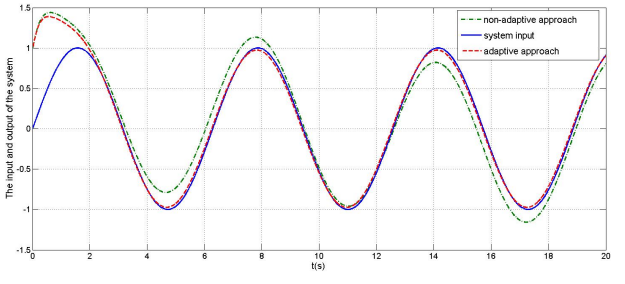


Fig. 4. The input and output of the system

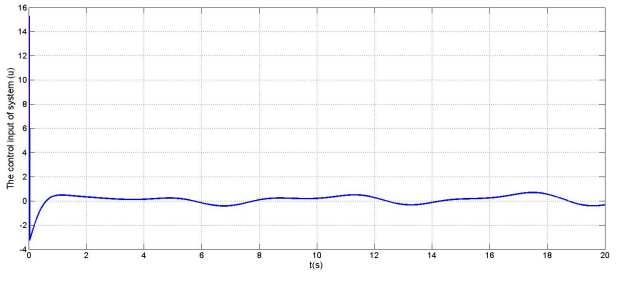


Fig. 5. The trajectory of the controller

where ω is set to 0.3, Δ ranges from 0 to 0.3. Fig. 4 shows the input and output signals (adaptive approach and non-adaptive approach). The results demonstrate that SVR using the adaptation algorithm in give a more precise output than the non-adapted case. Fig. 5 shows the trajectory of the controller.

6 Conclusion

This paper has been concerned with designing a nonlinear adaptive backstepping control system capable of tracking desired trajectory while uncertainties and disturbances existing in system model. SVR is employed to adaptively approximate the unknown nonlinear functions composed of unknown uncertainties and disturbances. According to stability analysis using Lyapunov function, the closed-loop system is guaranteed to be bounded and tracking errors are also proved to converge exponentially to a small residual set around the origin. The effectiveness of the proposed control approach is demonstrated in the tracking problem of a nonlinear model with uncertainties and disturbances.

References

[1] Nam, K., Arapostations, A.: A model reference adaptive control scheme for pure-feedback nonlinear systems. IEEE Transactions on Neural Networks 33, 803–811 (1998)
[2] Labiod, S., Boucherit, M.S., Guerra, T.M.: Adaptive fuzzy control of a class of MIMO nonlinear systems. Fuzzy Sets and Systems 151, 59–77 (2005)
[3] Yucelen, T., Calise, A.J.: Kalman Filter Modification in Adaptive Control. Journal of Guidance, Control and Dynamics 33(2), 426–439 (2010)
[4] Kanellakopoulos, I., Kokotovic, P.V., Morse, A.S.: Systematic Design of Adaptive Controllers for Feedback Linearizable Systems. IEEE Transaction on Automatic Control 36(1), 1241–1253 (1991)
[5] Wang, C., Ge, S.S.: Adaptive Backstepping Control of Uncertain Lorenz System. International Journal of Bifurcation and Chaos 11(4), 1115–1119 (2001)
[6] Khalil, H.K.: Nonlinear Systems, 3rd edn. Prentice-Hall, Upper Saddle River (2002)
[7] Kwan, C., Lewis, F.L.: Robust backstepping control of nonlinear systems using neural networks. IEEE Transaction on Systems, Man and Cybernetics 30, 753–765 (2000)
[8] Knohl, T., Unbehauen, H.: ANNNAC—extension of adaptive backstepping algorithm with artificial neural networks. Control Theory and Applications 147, 177–183 (2000)
[9] Schilling, R.J., Carroll, J.J.J., Al-Ajlouni, A.F.: Approximation of Nonlinear Systems with Radial Basis Function Neural Networks. IEEE Transactions on Neural Networks 12(1), 1–15 (2001)
[10] Zhang, T., Ge, S.S., Hang, C.C.: Adaptive neural network control for strict-feedback nonlinear systems using backstepping design. Automatica 36, 1835–1846 (2000)
[11] Niu, Y., Lam, J., Wang, X.Y., et al.: Adaptive H_∞ Control Using Backstepping Design and Neural Networks. Transactions of the ASME 127, 478–485 (2005)
[12] Vapnik, V.N.: The nature of statistical learning theory. Springer, New York (1995)
[13] Smola, A.J., Schölkopf, B.: A tutorial on support vector regression. Statistics and Computing 14, 199–222 (2004)

[14] Basak, D., Pal, S., Patranabis, D.C.: Support Vector Regression. Neural Information Processing-Letters and Reviews 11(10), 203–224 (2007)
[15] Suykens, J.A.K., Vandewalle, J., Moor, B.D.: Optimal control by least squares support vector machines. Neural Networks 14, 23–35 (2001)
[16] Shi, H.: A Novel Scheme for the Design of Backstepping Control for a Class of Nonlinear Systems. Applied Mathematical Modelling 35(4), 1893–1903 (2011)
[17] Farrell, J.A., Polycarpou, M., Sharma, M., et al.: Command Filtered Backstepping. IEEE Transaction on Automatic Control 54(6), 1391–1395 (2009)

The Application of Data Mining Technology in the Breakdown Maintenance of Warship Equipment

Hua Rong, Mingrong Chen, and Zhen Wang

Department of Operational Commands, Dalian Naval Academy, Dalian 116018, China

Abstract. Along with the development of the Navy, warship equipment breakdown maintenance and support become more difficult. Data mining technology was developed to solve the problems in warship equipment breakdown maintenance. The paper adopts the algorithm of fuzzy association rule, establishes the framework of equipment data mining system, analyses the fault causes, creates the fault phenomenon database, equipment fault database, and fault path database. The research is valuable for increasing the samples of fault and maintenance parts automatically and improving the lack of slow and inaccuracy of fault samples location.

Keywords: data mining technology, warship equipment, breakdown maintenance.

1 Introduction

Navy must have three capabilities in the modern naval combat missions. Firstly, Navy has the combat capability of inshore defense and offshore offense and defense to guarantee the security of the national territory, territorial waters and exclusive marine economic zones. Secondly, Navy has the combat capability of safeguarding the national marine rights and interests with ability of the control of air and sea, to guarantee the security of maritime economic transport channel and marine resource development project. Thirdly, Navy has the function of strategic deterrence at sea and maritime strategy to combat mobile base, and has the ability to maintain the country's overall interests and the strategic position. For promoting the three great abilities, we must develop a corresponding development of weapons and equipment. Because of modern equipment have high technical content, higher requirements for maintenance personnel, maintenance time and maintenance techniques have been put forward. And due to the particularity of surface warships' combat zone, a greater difficulty of maintenance of equipment will be increased. This paper adopts the data mining technology to construct an open platform for increasing the samples of the fault and maintenance parts automatically, and solve the problem that fault samples location are not quick and accurate enough.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 25–29.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 The Structural Framework of Data Mining System of Some Certain Equipment

There are several techniques used in the equipment breakdown maintenance, such as the fault tree coding technology, data mining technology, intelligent diagnosis of expert system technology and equipment fault data preprocessing technology [1-3]. This paper takes certain warship equipment as an example and discusses the data mining technology in the application of breakdown maintenance of warship equipment. The system uses the association rules technology to build the data mining system of equipment. The structural framework of the conceptual model can be designed by the user query interface, query collaborative machine, database management module, knowledge base management module, management module of the model base, pre-processing module of knowledge discovery base, knowledge evaluation module and the conclusion and interpretation modules. The system structure is shown in Figure 1.

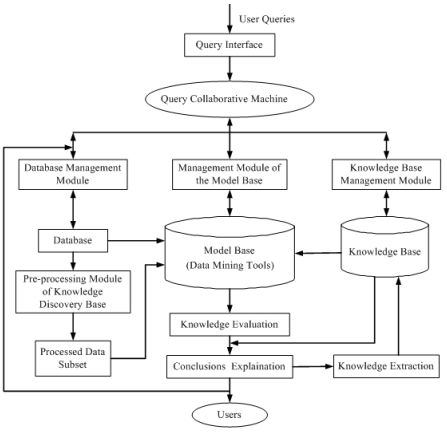


Fig. 1. The structural framework of data mining system of some certain equipment

3 The Algorithm of Fuzzy Association Rule Based on Tree

3.1 The Process of Data Mining

Data mining system usually consists of four parts: data preparation module, the pattern mining module, the results of expression and users interaction modules. Among them the user interaction runs through the whole process of data mining [4]. The entire data mining process is divided into the following steps, such as: 1) Understanding problems; 2) Data collection and filtering; 3) Data selection and transformation; 4) Data mining Model assessment; 5) Model assessment; 6) Knowledge representation.

3.2 The Establish of Equipment Fault Database

Association rules mining usually aims at database or Data Warehouse (DW). In order to adapt to the mining of fault association rules and find fault knowledge, a suitable database should be built. According to the needs of different fault knowledge, a variety of different database can be constructed out. A corresponding fault database will be built in this article, which focuses on the link between the fault reason and phenomenon associated with knowledge, and association knowledge of system-level fault of equipment.

(1) Fault reason and phenomenon database

In a large and complex equipment system, a fault is often corresponding to many reasons, and these reasons are concurrent. Many of the same equipment failure can constitute the fault database (shown in Table 1), through the connection operation of the various fault data sheet.

Table 1. Fault cause and the phenomenon database

Fault Number	Phenomenon 1	Phenomenon 2	…	Phenomenon n	Cause 1	Cause 2	…	Cause n
0001	1	0	…	1	1	0	1	1
0002	1	0	…	0	1	1	1	0
…	…	…	…	…	…	…	…	…
m	1	0	…	1	1	1	0	1

For the description of the fault data, the element of database "1" indicates that this failure has occurred or arisen, "0" is on the contrary. The following database elements have the same meaning.

In order to establish the fault data sheet for some non-Boolean element values, fault data is converted to the values in the interval [0, 1].

(2) Equipment fault database

Failures of equipment, composed of systems, are sometimes occurred with the association of other failures, so these associated failures are often caused by a number of subtle reasons and it is difficult to found by man. In order to find the relationship between fault knowledge, the fault database shown in Table 2 can be established which all of the associated equipment as a transaction when the fault occurs.

Table 2. Fault equipment database

Fault Number	Equipment 1	Equipment 2	Equipment 3	…	Equipment n
0001	1	0	1	…	1
0002	1	0	0	…	0
…	…	…	…	…	…
m	1	1	1	…	0

4 The Mining Based on the Association Rules of Multi-Level Failure

The fault diagnosis system is a complex system, which has several subsystems and adopts the multi-level fault association rules. And many failures in each subsystem have the mutual influences and restrictions and the rule knowledge of failure are implied in the different abstract levels of the database. In order to found out the associated rules of different levels, the equipment system is divided into three levels, such as: the Equipment System Layer, Subsystem Layer, Phenomenon Layer and Cause Layer, and it should paid attention that the Phenomenon Layer and the Cause Layer belong to the same level of different concepts.

In order to mine the association rules of multi-level fault, the corresponding relational database can be established according to the conceptual level. The representation of levels can use the tabular form. For reducing the data redundancy, this paper makes full use of the mature SQL query optimization technique of relationship database and connects these to generate a fault data sheet, which contains the failure properties of all levels.

5 The Analysis of Obtaining the Failure Knowledge

It can get the following three kinds of knowledge through the mining of the above fault database, which based on the association rule, such as the knowledge about the fault reason and phenomenon, the fault knowledge of fault equipment and the knowledge about the choice of Spare Parts.

5.1 The Knowledge about the Fault Reason and Phenomenon

The associated knowledge can be gained by the mining of the fault cause and phenomena database shown in Table 1, which based on the Association Rules, such as: Fault(X, “Phenomenon 1”)∧ Fault(X, “Phenomenon 2”) Cause(X, “Cause 1”)

Where: X represents the equipment variables. The association rules expresses that the possibility of fault Phenomenon 1 and fault Phenomenon h went wrong simultaneously accounting for all faults is $a\%$, and at the same time the possibility of the fault Cause 1 also occurred is $b\%$. Fault experts can find out what makes these three appear at the same time in such a high probability according to the knowledge, thus more deeper understanding about the mechanism of faults can be gained. And the associated rules can be used as fault knowledge into the knowledge base of fault diagnosis expert system and guidance of the fault diagnosis.

5.2 The Fault Knowledge of Fault Equipment

The associated knowledge can be gained by the mining of the fault database shown in Table 2, which based on the Association Rules, such as: Fault(X, “Equipment 1”) Fault(X, “Equipment 2”)

Where: X represents the equipment variables. The association rules expresses that the possibility of Equipment 1 went wrong accounting for all faults is $c\%$ and the possibility of Equipment 1 and Equipment 1 went wrong simultaneously is $d\%$. So the possible failure positions and failure causes can be analyzed according to the probability of fault occurrence value, and the associated faults, which is difficult to find, can be eliminated by the knowledge.

6 The Application of the Fast and Accurate Positioning Method in Data Mining

The fault position, determined by association rules according to the fault phenomenon, has certain probability of misjudgment. In order to determine the fault positions quickly and accurately, the system bring the judgment results about fault position of failure phenomena dynamically into the loop, through the establishment of the fault path database mode and adopting an open platform. As a result the probability of judging the fault positions rapidly and accurately is increased. The database is shown in Table 3, and this table can be filled in automatically, but also in artificially.

Table 3. Fault path database

Fault Number	Subsystem 1	Subsystem 2	Subsystem 3	…	Subsystem n
0001	Position 1	Position 1	Position 1	…	Position 1
0002	Position 2	Position 2	Position 2	…	Position 2
…	…	…	…	…	…
M	Position n	Position n	Position n	…	Position n

References

[1] Liu, W.-C., Shi, H.-J., Ma, S.-Q.: Algorithm of weight fuzzy association rules. Computer Engineering and Design 31(16), 3654–3657 (2010)

[2] Shen, P.-F., Wang, X.-W., Geng, Z.-G., et al.: Research on Method Sorting of Data Mining. Value Engineering 26, 146–147 (2010)

[3] Huo, W.-G., Shao, X.-L.: A Fuzzy Associative Classification Method Based on Multi-Objective Evolutionary Algorithm. Journal of Computer Research and Development 48(4), 567–575 (2011)

[4] Zheng, S.-M., Shao, R.-M., Miao, Z.: Data Mining Architecture Based on Grid. Command Information System and Technology 2(5), 41–47 (2011)

Intrusion Detection Based on Immune Principles and Fuzzy Association Rules

Zhang Lei, Meng Lingrui, and Hou Chunjie

Henan University of Science and Technology, School of Electronic Information Engineering

Abstract. In this paper, a new intrusion detection method based on immune principles and fuzzy association rules is proposed. The proposed method uses fuzzy association rules for building fuzzy classifiers, which is also the detection engine of the intrusion detection system. A novel immune-inspired algorithm is proposed for mining fuzzy association rule set, in which the fuzzy sets corresponding to each attribute and the final fuzzy rule set can be directly extracted from a given data set. The KDD-99 dataset is used to evaluate the performance of the proposed algorithm and compared with other relevant intrusion detection methods. The results show the detection performances of the proposed algorithm are comparable with other relevant intrusion detection systems.

Keyword: Detection, Immune principles, Fuzzy association rules.

1 Introduction

Intrusion detection techniques can be categorized into anomaly detection and misuse detection. Signature-based misuse detection techniques are one kind of most widely used intrusion detection methods [1]. Data mining techniques, such as artificial neural network, fuzzy systems, are another kind of common intrusion detection techniques widely used in intrusion detection community [2-3]. The advantage of these intrusion detection system based on data mining techniques is that these system have the self-adaptability and self-learning ability, and they can discover useful detection rules or knowledge which are easily understood and used for users.

Fuzzy systems based intrusion detection systems have been proposed and successfully used in many application areas [4-6]. In [4], a fuzzy genetics-based learning method was introduced on intrusion detection problem, which hybridizing the approximate reasoning method of fuzzy systems with the learning capabilities of genetic algorithms. Three kinds of genetic fuzzy systems were proposed in [5] based on Michigan, Pittsburgh and iterative rule learning (IRL) approaches. Another fuzzy system based intrusion detection system is also proposed in [6], which is based on iterative rule learning using a fuzzy rule-based genetic classifier.

In this paper, a new intrusion detection method based on immune principles and fuzzy association rules is proposed. The proposed method uses fuzzy association rules for building fuzzy classifiers, in which the fuzzy sets corresponding to each attribute and the final fuzzy rule set can be directly extracted from a given data set.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 31–35.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Fuzzy Association Rules and Fuzzy Classification

A fuzzy association classification rule has the following form:

$$R_i: \ if \ X \ \text{is} \ A \ \text{then} \ C_i \ \text{with} \ (Sig_i, Cer_i). \tag{1}$$

where $'X \ \text{is} \ A'$ is called the antecedent of the rule R_i and $'C_i'$ is called the consequent of R_i. $Sig(R_i)$ and $Cer(R_i)$ are the significance factor and the certainty factor of the rule R_i.

$$Sig(R_i) = \frac{\sum_{t_i \in T} (\prod_{j=1}^{n} m'_{a_j}(t_i[a_j]))}{T_N} \tag{2}$$

Where T is the giving dataset and T_N is the number of transaction in T.

$$m'_{a_j}(t_i[a_j]) = \begin{cases} m_{a_j}(t_i[a_j]), & \text{if } m_{a_j} \geq th_m \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Suppose that we have induced L disjoint fuzzy rule sets, $RS=(RS_1, \cdots, RS_L)$, such that each rule set RS_i $(1 \leq i \leq L)$contains the rules describing the patterns observed in the class i.

Considering a transaction t, the firing strength for the rule R_i (defined in equation 1) is formulated as follows.

$$fs(t, R_i) = \prod_{j=1}^{n_i} m_{a_j}(x_j) \tag{4}$$

Where n_i is the number of attributes appeared in the rule R_i. Then we can compute the matching measure of a transaction t and a fuzzy rule set RS_i , which is calculated as follows.

$$mm(t, RS_i) = \sum_{j=1}^{N_i} fs(t, R_j) \times Cer(R_j) \tag{5}$$

The class label for the transaction t is the class having the maximum matching measure, which is determined as follows.

$$C_label = \arg\max_{i=1}^{L} mm(t, RS_i) \tag{6}$$

3 The Proposed Intrusion Detection Model

In our previous literature [7], an algorithm based on natural immune principles was proposed for determining fuzzy sets and corresponding membership functions. In this study, we also adopt this method. A novel immune-inspired algorithm is proposed for mining fuzzy association rule set in this section. The detailed steps of the proposed algorithm are as follows.

1. Generate the initial population.

A inter coding scheme is considered, and then each antibody in the population has the following form.

$$(a_1, a_2, \cdots, a_n), \quad a_i \in \{0,1,2,\cdots,m\}, i = 1,2,\cdots,n$$

Correspond to a fuzzy association rule as follows:

$$R_i: \textit{ if } x_1 \text{ is } m_{a_1} \text{ and } x_2 \text{ is } m_{a_2} \cdots \text{and } x_n \text{ is } m_{a_n} \text{ then } C_i \text{ with } (Sig_i, Cer_i). \tag{7}$$

2. Selection of the antibody with high affinity for clone.

The affinity of the antibody Ab_i is defined as the value of the *certainty factor* of association rule R_i. A given number antibodies with higher affinity are selected for clone. The density of an association rule in the current population is also considered in order to keep population diversity effectively. This concept reflects whether the number of rules similar to this rule is many enough.

3. Cloning process.

The cloning process is carried out such that the number of clones of an antibody is directly proportional to its affinity and the cell proliferation will be inhibited for high density individuals. When the density *of an* antibody *is higher than a predefined threshold, then its* number of clones is directly set to the minimum otherwise the number of clones *is proportional to its affinity.*

4. Mutation operation.

The mutation rate of an antibody is inversely proportional to the affinity of the cell. Concretely speaking, when an antibody is selected to execute mutate operation, one gene is changed by a new inter selected randomly from its range.

5. Memory selection.

The memory selection process adds some new higher confidence rules into the memory population. At the same time some redundant rules are deleted from the memory population.

6. Updating the antibody population for the next iteration.

A proportional of antibodies with lower affinity are replaced by the new produced ones after the mutation process based on their affinity. Thus we can get the antibody population for the next iteration.

7. Termination test.

The stopping condition is the predefined maximum number of iteration and the coverage constraints, which are that the percentage of instances classified correctly by the current memory population surpasses a predetermined threshold.

4 Performance Evaluations

The KDD-99 dataset is used to evaluate the proposed algorithm for intrusion detection [8]. Usually the performances of intrusion detection systems are evaluated in terms of false positive rate and detection rate. The false positive (negative) rate is the number of normal (attack) connections labeled as attack (normal).

Table 1 provides a summary of performance comparison for some relative detection approaches which are also trained on the '10%KDD' dataset and tested using the 'Corrected (Test)' data [9]. As observed in Table 1, the detection rate and false positive rate of our algorithm are competitive to other relative approaches.

Table 1. Detection performance comparison for different approaches

algorithms	SVM	H-SOM	Clustering	ABC[9]	The proposed algorithm
Detection rate (%)	98	90-91.5	93	91	94.5
False positive rate (%)	10	7.6-14.5	10	3.3	5.2

Three different data distribution of training dataset are adopted in the training process. In the first case, all the data in the 10% KDD are used for training; in the second case, only the normal data in the 10% KDD are used for training; in the last case, the normal data and the attack data are in equal portions, each 50%.

From Table 2, we can see that the detection performances on the normal data for the three cases are all satisfying. Especially the results of the 'Dos' category are more prominent. However, the performance for the attack data 'U2R' and 'R2L' categories are relatively lower. This is because most of these attacks are content based which means the attack is manifested in the packet payload.

Table 2. Detailed detection performance comparison for different training datasets

	Case1	Case2	Case3
Normal	92.9%	89.7%	91.8%
Dos	95.8%	95.9%	96.0%
Probe	79.2%	88.6%	80.7%
U2R	21.1%	21.1%	20.8%
R2L	12.6%	19.7%	15.2%

Next, the detection performance are tested in two cases: all the 'Corrected' dataset set is used as the test dataset and the remaining records by excluding unknown (unseen) attacks from the 'Corrected' dataset is used as the test dataset. As can be seen from Table 3, the detection performance in the first case are relatively poor than in the second case. The reason is that the unseen attack data are not included in the test dataset and this at least doesn't reduce performance.

Table 3. Detailed detection performance comparison of different test datasets

	All 'Corrected' dataset	excluding unknown (unseen) attacks
Normal	92.9%	94.5%
Dos	95.8%	96.6%
Probe	79.2%	79.6%
U2R	21.1%	21.8%
R2L	12.6%	12.7%

5 Conclusion

In this paper, a novel intrusion detection method based on immune principles and fuzzy association rules is proposed. An immune-inspired algorithm is used for mining fuzzy association rule set from the given data set directly, and generates the most appropriate fuzzy association rule set as the descriptive model for normal data and attack data. The KDD-99 dataset is used to evaluate the performance of the proposed algorithm, and the simulation results are comparable to the performance of other relevant algorithms.

References

[1] Lee, W., Stolfo, S., Mok, K.: A data mining framework for building intrusion detection models. In: Proceedings of the IEEE Symposium on Security and Privacy, pp. 120–132 (1999)

[2] Florez, G., Bridges, S.M., Vaughn, R.B.: An improved algorithm for fuzzy data mining for intrusion detection. In: Proceedings of the North American Fuzzy Information Processing Society Conference NAFIPS, New Orleans, LA, pp. 27–29 (2002)

[3] Tajbakhsh, A.: Design and implementation of an Intrusion Detection Systems using data mining techniques. M.Sc. Thesis. Amirkabir University of Technology (2006)

[4] Saniee, A., Habibi, M.J., Lucas, C.: Intrusion detection using a fuzzy genetics-based learning algorithm. Journal of Network and Computer Applications 30, 414–428 (2007)

[5] Saniee, M., Mohamadi, H., Habibi, J.: Design and analysis of genetic fuzzy systems for intrusion detection in computer networks. Expert Systems with Applications 38, 7067–7075 (2011)

[6] Ozyer, T., Alhajj, R., Barker, K.: Intrusion detection by integrating boosting genetic fuzzy classifier and data mining criteria for rule pre-screening. Journal of Network and Computer Applications 30, 99–113 (2007)

[7] Zhang, L., Li, R.: An Algorithm for Mining Fuzzy Association Rules Based on Immune Principles. In: Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering, pp. 1285–1289 (2007)

[8] Stolfo, S.J., et al.: KDD (1999), `http://www.kdd.ics.uci.edu/databases/kddcup99/kddcup99.html`

[9] Kayacik, H.G., Zincir-Heywood, A.N., Heywood, M.I.: On dataset biases in a learning system with minimum priori information for intrusion detection. In: Proceedings of the 2nd Annual Conference on Communication Networks and Services Research, pp. 181–189 (2004)

Temperature Model for FOG Zero-Bias Using Gaussian Process Regression

He Zhikun, Liu Guangbin, Zhao Xijing, and Yang Jian

Xi'an Research Institute of High-Tech, Xi'an, Shaanxi 710025, P.R. China
hezhikun0@sina.com

Abstract. As the environmental temperature variation is one of the most important factors influencing the zero-bias of Fiber Optic Gyro (FOG), the temperature drift characteristic of FOG zero-bias was analyzed and its model method was studied. A new approach based on Gaussian process regression (GPR) was introduced to deal with the complex nonlinearity of FOG zero-bias against temperature. Through some small training set, the nonlinear mapping relationship between FOG zero-bias and temperature could be established accurately. Compared with Least Square primary and quadratic regression methods, it is shown that the model built by the new approach is able to obtain higher accuracy and the fitting error is reduced by an order of magnitude. Meanwhile, the case study shows that the GPR model is feasible, easy to be implemented, and very attractive for a wide application.

Keywords: Gaussian Process Regression (GPR), Zero-Bias, Temperature Model, Nonlinear.

1 Introduction

Fiber optic gyroscope (FOG) is an inertial instrument which is based on Sagnac effect and has no rotating part. It has a lot of unique advantages such as small size, low operating costs, large dynamic range, long lifetime, quick start-up and strong shock load capability. So it has been widely used both in military and civil areas for inertial navigation and guidance, attitude determination and stabilization, strapdown rate gyro compass and vehicles position and orientation [1]. To be adaptive for using in these various areas, FOG is required to have large operating temperature range; it ordinarily is -40℃ ~ 60℃. But the core components of FOG such as fiber optic sensing coils, light sources and couplers are sensitive to temperature. When the operating temperature varies, thermal-induced non-reciprocity noise will be included in the FOG output [2]. It results in zero-bias and scale-factor drifting [3]. In general, there are two ways to reduce the bad effect caused by temperature varying. One is to control the inner temperature of FOG to keep constant. The other, which is easier to be realized than the former, is to model and compensate for FOG temperature drift in software [4-5].

Affected by many factors, FOG possesses very complex nonlinear temperature characteristic, which can't be described exactly by conventional methods such as linear polynomial fitting. Up to the present, many researchers have done much work to

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 37–45.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

solve it and there have been many advanced methods to be adopted. For example, some machine learning methods like Neural Network (NN) [6-7] and Support Vector Regression machine (SVR) [8] have been used and there have obtained many valuable research results. However, these methods have some flaws. NN, which based on empirical risk minimization rule, only guarantees that the estimated error for training points is minimal. It is easy to trap at a local minimum of the chosen optimization criterion during the learning procedure when a gradient descent algorithm is used. And its generalization ability is restricted. Based on statistical learning theory and structure risk minimization rule, SVR learns the structure of target model automatically, has better generalization ability and overcomes essentially some flaws of NN [9]. But SVR still has some problems. For example, it is difficult to choose kernel function and obtain its hyperparameters. In addition, the prediction of SVR has no clear probabilistic interpretation.

Gaussian Process is a new approach in machine learning area, which has been developed rapidly over last two decade. It is based on the strict theory of statistical learning and is applicable well for some complex problems in high dimension, small sample, or nonlinear systems. It also has good generalization ability. Compared to NN and SVR, Gaussian Process has a lot of advantages such as being easy to implement, flexible to nonparameter infer, adaptive to obtain hyperparameters, and so on. Moreover, Gaussian Process provides a probabilistic approach to learning in kernel machines, and produces a prediction with a clear probabilistic interpretation. Therefore, Gaussian Process has many successful applications and becomes one focus in machine learning area [10-11].

In this paper, a new temperature model for FOG zero-bias using Gaussian process regression (GPR) method is proposed. The rest of this paper is organized as follows. The principle of GPR is introduced in Section 2. Section 3 analyzes the zero-bias model of FOG and gives the modeling process using GPR method. In Section 4, the effectivity of the new approach and temperature model are demonstrated and verified via an example. Finally, some significative conclusions are given in Section 5.

2 The Principle of Gaussian Process Regression [12-15]

We have a training set $D=\{(\mathbf{x}_i, y_j) \mid i=1,\cdots,n\}=\{X,\mathbf{y}\}$, where $\mathbf{x}_i \in R^d$ denotes an input vector of dimension d and $y_i \in R$ denotes a corresponding scalar output. $X=[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n]$ denotes a matrix of dimension $d\times n$ which aggregates the column vector inputs for all n cases and $\mathbf{y}$ aggregates the n outputs. The task of GPR is to learn the mapping relationship between inputs X and outputs $\mathbf{y}$, i.e. $f(\cdot): R^d \mapsto R$, and predict the most possible output value $f(\mathbf{x}_*)$ given the new test point $\mathbf{x}_*$. The implement of GPR can be divided into two steps: prediction and training step.

2.1 The Prediction Step of GPR

We can interpret the learning principle of GPR both in the weight space and function space where it obtains identical results. With the function-space view mainly, we define a Gaussian Process to describe the distribution over functions, and do inference directly in function space.

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution. It is completely specified by its mean function and covariance function as

$$\begin{cases} m(\mathbf{x}) = \mathbf{E}[f(\mathbf{x})] \\ k(\mathbf{x},\mathbf{x}') = \mathbf{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))] \end{cases} \tag{1}$$

where $\mathbf{x},\mathbf{x}' \in R^d$ denotes any random variable. So we can define the Gaussian process as $f(\mathbf{x}) \sim GP(m(\mathbf{x}), k(\mathbf{x},\mathbf{x}'))$. For notational simplicity, we usually pretreat the data to make the mean function to be zero.

For linear regression problems, we can define the model as

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{w}, \qquad y = f(\mathbf{x}) + \varepsilon \tag{2}$$

where $\mathbf{x}$ denotes an input vector, $\mathbf{w}$ denotes a weights vector of the linear model, f denotes the function value and y denotes the observed value which is polluted by additivity noise. We further assume that this additivity noise follows an independent, identically distributed Gaussian distribution with zero mean and variance σ_n^2, i.e. $\varepsilon \sim N(0,\sigma_n^2)$.

For more general nonlinear regression cases, we can project their inputs into high dimensional feature space using "kernel trick" and apply the linear model in new space. For example, the vector input $\mathbf{x}$ can be projected into N dimensional space from d dimensional space where $N > d$. So we can define the nonlinear regression model as

$$f(\mathbf{x}) = \phi(\mathbf{x})^T\mathbf{w}, \qquad y = f(\mathbf{x}) + \varepsilon \tag{3}$$

In the Bayesian formalism, we need to specify a prior distribution over the weights vector $\mathbf{w}$ which denotes our prior knowledge about the weights vector $\mathbf{w}$ before obtaining the observations. So we further assume that this weights vector $\mathbf{w}$ follows a Gaussian distribution with zero mean and variance Σ_p, i.e. $\mathbf{w} \sim N(0,\Sigma_p)$.

Combining $f(\mathbf{x}) = \phi(\mathbf{x})^T\mathbf{w}$ and $\mathbf{w} \sim N(0,\Sigma_p)$, we can obtain

$$\begin{cases} \mathbf{E}[f(\mathbf{x})] = \phi(\mathbf{x})^T\mathbf{E}[\mathbf{w}] = 0 \\ \mathbf{E}[f(\mathbf{x})f(\mathbf{x}')] = \phi(\mathbf{x})^T\mathbf{E}[\mathbf{w}\mathbf{w}^T]\phi(\mathbf{x}') = \phi(\mathbf{x})^T\Sigma_p\phi(\mathbf{x}') \end{cases} \tag{4}$$

From above formula, we can see that $f(\mathbf{x})$ and $f(\mathbf{x}')$ follows a joint Gaussian distribution with zero mean and covariance $\phi(\mathbf{x})^T\Sigma_p\phi(\mathbf{x}')$, i.e. the function output values

$f(\mathbf{x}_1), f(\mathbf{x}_2), \cdots, f(\mathbf{x}_n)$ corresponding to any n inputs follows a joint Gaussian distribution. We define covariance function or kernel function as $k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \Sigma_p \phi(\mathbf{x}')$. As Σ_p is positive definite, it exists $\Sigma_p^{1/2}$ to satisfy $(\Sigma_p^{1/2})^2 = \Sigma_p$. Let $\psi(\mathbf{x}) = \Sigma_p^{1/2} \phi(\mathbf{x})$, we obtain a dot product representation of kernel function as $k(\mathbf{x}, \mathbf{x}') = \psi(\mathbf{x})^T \cdot \psi(\mathbf{x}')$.

Thus the prior distribution of observed value $\mathbf{y}$ becomes

$$\mathbf{y} \sim N(0, K(X, X) + \sigma_n^2 I_n) \tag{5}$$

and the joint prior distribution of observed value $\mathbf{y}$ and prediction f_* becomes

$$\begin{bmatrix} \mathbf{y} \\ f_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I_n & K(X, \mathbf{x}_*) \\ K(\mathbf{x}_*, X) & k(\mathbf{x}_*, \mathbf{x}_*) \end{bmatrix}\right) \tag{6}$$

where $K(X, X) = K_n = (k_{ij})$ denotes the $n \times n$ symmetric and positive definite covariance matrix whose element $k_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ measures the relativity between $\mathbf{x}_i$ and $\mathbf{x}_j$, $K(X, \mathbf{x}_*) = K(\mathbf{x}_*, X)^T$ denotes the $n \times 1$ covariance matrix between the inputs of training set and the test point $\mathbf{x}_*$, $k(\mathbf{x}_*, \mathbf{x}_*)$ denotes the variance of the test point $\mathbf{x}_*$, and denotes identity matrix of dimension n.

Then we can obtain the joint posterior distribution of prediction as

$$f_* \mid X, \mathbf{y}, \mathbf{x}_* \sim N(\bar{f}_*, \mathrm{cov}(f_*)) \tag{7}$$

$$\text{where } \bar{f}_* \triangleq \mathbf{E}[f_* \mid X, \mathbf{y}, \mathbf{x}_*] = K(\mathbf{x}_*, X)[K(X, X) + \sigma_n^2 I_n]^{-1} \mathbf{y} \tag{8}$$

$$\mathrm{cov}(f_*) = k(\mathbf{x}_*, \mathbf{x}_*) - K(\mathbf{x}_*, X)[K(X, X) + \sigma_n^2 I_n]^{-1} K(X, \mathbf{x}_*) \tag{9}$$

Therefore the mean and variance of the prediction corresponding to the test point $\mathbf{x}_*$ are respectively $\hat{\mu}_* = \bar{f}_*$ and $\hat{\sigma}_{f_*}^2 = \mathrm{cov}(f_*)$.

As the requirement that the covariance of Gaussian Process method must be positive definite in any finite input set is consistent with the property of kernel in Mercer's theorem, the covariance function and kernel function are equivalent and eq. (8) can become

$$\hat{\mu}_* = \sum_{i=1}^{n} \alpha_i k(\mathbf{x}_i, \mathbf{x}_*) \tag{10}$$

where $\alpha_i = \alpha = [K(X, X) + \sigma_n^2 I_n]^{-1} \mathbf{y}$. From eq. (10), it is shown that the mean of prediction is a linear combination of kernel functions $k(\mathbf{x}_i, \mathbf{x}_*)$. When data is projected into high dimensional feature space, its nonlinear relationship becomes linear, and the complex nonlinear problems become linear problems which are easier to solve.

2.2 The Training Step of GPR

GPR can choose various covariance functions. Some common functions are shown in the following:

(1) Linear Covariance Function: $k(\mathbf{x},\mathbf{x}') = \mathbf{x}^T \Lambda^{-2} \mathbf{x}'$ (11)

(2) Polynomial Covariance Functions: $k(\mathbf{x},\mathbf{x}') = \sigma_f^2 (\mathbf{x}^T \mathbf{x}' + c)^d$ (12)

(3) Squared Exponential Covariance Function (SE):

$$k(\mathbf{x},\mathbf{x}') = \sigma_f^2 \exp(-\frac{1}{2}(\mathbf{x}-\mathbf{x}')^T \Lambda^{-2} (\mathbf{x}-\mathbf{x}')) \tag{13}$$

(4) Rational Quadratic Covariance Function (RQ):

$$k(\mathbf{x},\mathbf{x}') = \sigma_f^2 \left[1 + \frac{(\mathbf{x}-\mathbf{x}')^T \Lambda^{-2} (\mathbf{x}-\mathbf{x}')}{2\lambda}\right]^{-\lambda} \tag{14}$$

(5) The Matérn Class of Covariance Functions:

$$\begin{aligned} &k_{v=d/2}(\mathbf{x},\mathbf{x}') = \sigma_f^2 f_d(r_d) \exp(-r_d), r_d = \sqrt{\frac{d}{l^2}(\mathbf{x}-\mathbf{x}')^T (\mathbf{x}-\mathbf{x}')} \\ &d = 1,2,3 \qquad f_1(t) = 1, f_3(t) = 1+t, f_5(t) = f_3(t) + \frac{t^2}{3} \end{aligned} \tag{15}$$

where $\Lambda = diag(l^2)$, l denotes the characteristic length-scale, and σ_f^2 denotes the signal variance.

The parameters aggregation $\theta = \{\Lambda, \sigma_f^2, \sigma_n^2, c, \lambda\}$ is called hyperparameters, and the training step of GPR is also the optimization of hyperparameters. We can obtain the optimal hyperparameters by Maximum Likelihood Estimate method. Built the negative log marginal likelihood function $L(\boldsymbol{\theta}) = -\log p(\mathbf{y} \mid X, \boldsymbol{\theta})$ firstly and its partial derivatives with respect to the hyperparameters. Then minimize the partial derivatives to obtain the optimization of hyperparameters based on optimization algorithm such as conjugate gradient and Newton's method. The negative log marginal likelihood function $L(\boldsymbol{\theta})$ and its partial derivatives are shown in the following:

$$L(\boldsymbol{\theta}) = -\log p(\mathbf{y} \mid X, \boldsymbol{\theta}) = \frac{1}{2}\mathbf{y}^T C^{-1} \mathbf{y} + \frac{1}{2}\log|C| + \frac{n}{2}\log 2\pi \tag{16}$$

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i} = \frac{1}{2} tr\left(C^{-1}\frac{\partial C}{\partial \theta_i}\right) - \frac{1}{2}\mathbf{y}^T C^{-1} \frac{\partial C}{\partial \theta_i} C^{-1} \mathbf{y} = \frac{1}{2} tr\left((\alpha\alpha^T - C^{-1})\frac{\partial C}{\partial \theta_i}\right) \tag{17}$$

where $C = K_n + \sigma_n^2 I_n$, $\alpha = (K + \sigma_n^2 I_n)^{-1} \mathbf{y} = C^{-1}\mathbf{y}$.

After obtaining the optimization of hyperparameters, we can compute the prediction $\hat{\mu}_*$ and its variance $\hat{\sigma}_{f_*}^2$ corresponding to the test point $\mathbf{x}_*$ using eq. (16) and eq. (17).

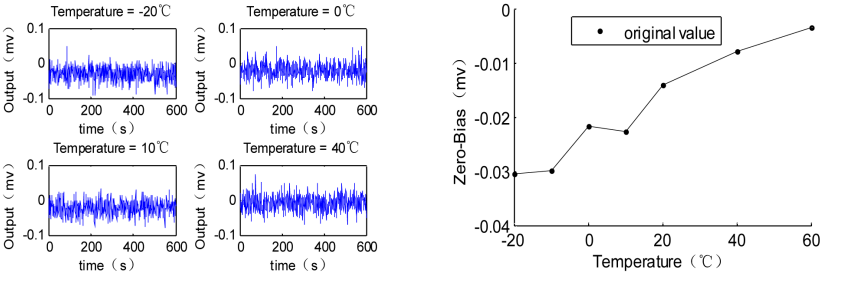
3 Temperature Model for FOG Zero-Bias

With the extension of research, the variation of environmental temperature becomes one of the most important factors to influence the zero-bias of FOG. It is significative for reducing the influence of FOG drift to build an appropriate compensation model about the nonlinear relationship between zero-bias and variation of environmental temperature. Usually, the zero-bias can be denoted as $B_0 = f(T,t)$ by taking temperature and time into account synthetically, where T denotes temperature factor and t denotes time factor. But through long-time tests, we can find that the influence of time factor to zero-bias is little. So the zero-bias can be denoted as $B_0 = f(T)$ for simplify. The purpose of this paper is to learn the mapping relationship $f(\cdot)$ using GPR. The process is shown in the following:

(1) Build the training set $D = (T, B_0)$ according to the experiment data (T_i, B_{0i}), where $i = 1,2,\cdots,n$.
(2) Choose an appropriate covariance function and set the initial value of hyperparameters.
(3) Train the GPR model using the training set, i.e. obtain the optimal hyperparameters by minimizing the negative log marginal likelihood function.
(4) Compute the prediction corresponding to any temperature point using eq. (16).

4 Simulation Test and Result

The aim of this section is to verify the effectiveness of GPR method to compensate the FOG zero-bias relative to environmental temperature. The data from a certain type FOG were sampled and investigated. The FOG was placed in a temperature control unit on a test platform. The temperature of this unit was adjusted to -20℃, -10℃, 0℃, 10℃, 20℃, 40℃, 60℃ in turn and the FOG drift data were collected with 1Hz. The sampling time at every temperature point was ten minutes. The part of original data is shown in Figure 1(a).



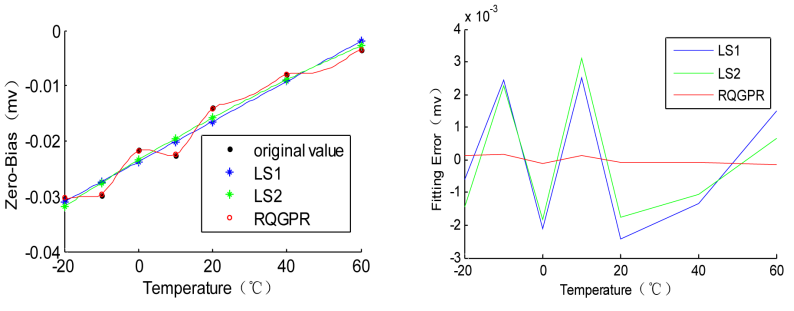
(a) The original data curve (partly) (b) The original points of FOG zero-bias

Fig. 1. The original data

To reduce the influence of FOG random drift and improving calibrated precision, it is usually to measure repetitiously and take its average as FOG zero-bias. We can obtain different zero-bias value at different temperature point, which is shown in Figure 1(b).

From Figure 1(b), we can find that the influence of temperature factor on FOG zero-bias is remarkable and the relationship between temperature and zero-bias is nonlinear obviously.

In this paper, we choose the Rational Quadratic Covariance Function and set the initial value of hyperparameters as $\ln l = 0$, $\ln \sigma_f = 0$, $\ln \lambda = 0$. Conjugate gradient method is used to obtain the optimal hyperparameters. For comparison, we also use Least Square primary Regression (LS1) and Least Square quadratic Regression (LS2) methods to build the nonlinear relationship. The fitting results are shown in Figure 2(a). And the fitting errors are shown both in Figure 2(b) and Table 1.



(a) The fitting results (b) The fitting error curves

Fig. 2. The fitting results

Table 1. The fitting errors of three methods

Method / Temperature	LS1	LS2	GPR
-20℃	-0.5912 e-03	1.4298 e-03	1.2681 e-04
-10℃	2.4520 e-03	2.2742 e-03	1.5503 e-04
0℃	-2.1048 e-03	-1.8102 e-03	-0.9550 e-04
10℃	2.5251 e-03	3.1035 e-03	1.4632 e-04
20℃	-2.4334 e-03	-1.7596 e-03	-0.9349 e-04
40℃	-1.3487 e-03	1.0496 e-03	-0.7900 e-04
60℃	1.5010 e-03	0.6715 e-03	-1.5835 e-04

From Figure 2(b) and Table 1 as above, we can find that the fitting error of GPR method is reduced by an order of magnitude compared to LS1 and LS2 methods. It is shown that the model established by GPR method is able to obtain higher accuracy and describe better the nonlinearity of FOG zero-bias against temperature.

5 Conclusion

The relationship between temperature and zero-bias of FOG, which can't be described exactly by the traditional methods, is complicated nonlinear. In this paper, Gaussian process regression approach is used to deal with this complex problem. Compared with the Least Square primary and quadratic regression methods, the new model established by Gaussian process regression approach is able to describe the nonlinearity of FOG zero-bias against temperature better and its output have higher accuracy with the fitting error reduced by an order of magnitude. Meanwhile, the case study also shows that the GPR model is feasible, easy to be implemented, and very attractive for a wide application.

References

[1] Wang, W.: The fiber optic gyroscope inertial navigation system. China Astronautic Publishing House, Beijing (2010)
[2] Sudhakar, D., Gregg, K., Chellappan, N.: Strategic interferometric fiber-Optic gyroscope for ballistic missile inertial guidance. In: AIAA Guidance, Navigation and Control Conference and Exhibit, Honolulu, Hawaii, August 18-21, pp. 1–8 (2008)
[3] Gu, Q.T.: A New Method for System Identification in the Presence of Colored Noise. Control Theory and Applications 7(3), 56–62 (1990)
[4] Stéphane, B., Michel, J.F.D., Gordon, S.K.: Noise analysis of an air-core fiber optic gyroscope. IEEE Photonics Technology Letters 19(9), 1520–1522 (2007)
[5] Stéphane, B., Hyang, K.K., Michel, J.F.D., Gordon, S.K.: Reduced thermal sensitivity of a fiber-optic gyroscope using an air-core photonic-bandgap fiber. Journal of Lightwave Technology 25(3), 861–865 (2007)
[6] Feng, L.S., Nan, S.Z., Jin, J.: Research on modeling and compensation technology for temperature errors of FOG. Journal of Astronautics 27(5), 939–941 (2006)
[7] Fan, C.L., Jin, Z.H., Tian, W.F., et al.: Temperature drift modeling of FOGs based on grey RBF neural network. Measurement Science and Technology 15(1), 119–126 (2004)
[8] Vapnik, V.N.: Statistical learning theory. John Wiley &Sons, New York (1998)
[9] Jiang, W., Yi, G.X., Zeng, Q.S.: Temperature drift prediction for MEMS gyroscopes based on PSVR. Journal of Chinese Inertial Technology 17(3), 360–365 (2009)
[10] Xiong, Z.H., Zhang, J.C., Shao, H.H.: GP-based soft sensor modeling. Journal of System Simulation 7(4), 793–794, 800 (2005)
[11] Jack, M.W., David, J.F., Aaron, H.: Gaussian process dynamical models for human motion. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(2), 283–298 (2008)

[12] Rasmussen, C.E., Williams, C.K.I.: Gaussian processes for machine learning. MIT Press (2006)
[13] Sofiane, B.B., Amine, B.: Gaussian process for nonstationary time series prediction. Computational Statistics & Data Analysis 47, 705–712 (2004)
[14] Lawrence, N.: Gaussian process latent variable models for visualization of high dimensional data. In: Advances in Neural Information Processing Systems, NIPS (2003)
[15] Lawrence, N.: Probabilistic non-linear principal component analysis with Gaussian process latent variable models. Journal of Machine Learning Research (JMLR) 6, 1783–1816 (2005)

Pulse Wave Averaging Computation and Clinical Verification

Li Yang and Yunan Hu

Department of Control Engineering, Naval Aeronautical Engineering University, China

Abstract. The cerebrovascular dynamics parameters play an important role in the cerebral circulation. According to the electric circuit graph theory and hemodynamic equation, establish nonlinear network model of the cerebral circulation, then use the averaging method to find an approximate solution about this model. We uses this method in the cerebrovascular equivalent network, the solution may help explain the development processes of venous diseases. Simulation shows that the computing result brings into correspondence with the blood flow Q change of the pathological clinical observation.

Keywords: Pulse Wave, Averaging Computation, Clinical Verification.

1 Introduction

Pulse wave propagation characteristics are closely linked with the hemodynamic parameters of the blood circulation system. Changes in pulse wave characteristics are an important basis to evaluate the physiological and pathological state of the human cardiovascular system. Professor Luo Zhichang, who used the existing two-chamber model of elastic wave pulse to extract the characteristics of K which represents changes of the pulse wave's area. K value reflects the characteristic quantities which changes in the amount of area of the pulse wave. This paper proposed a solving method that cardiovascular function parameters K-value will be calculated based on the averaging method calculated pulse wave, according to changes of the area and waveform of pulse in different physiological and pathological conditions, make the K value to combine with the pathological and diagnostic analysis [1-3].

2 Hemodynamic Network Modeling and Averaging Computation

According to the aforementioned study, in order to solve the pulse wave of blood circulation network diagram, first, analyze its network modeling. In order to build blood circulation network's model, at first establishes the dynamic equation of one blood vessel branch. For simplicity, we make the following assumptions: Al. the blood is incompressible; A2.the temperatures in all branches are identical. Under assumptions Al and A2, one branch of the blood network is described with the following equations[4-5]:

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 47–51.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

$$T_j \frac{dQ_j}{dt} = -R_j \left| Q_j \right| Q_j + H_j \Rightarrow T\dot{Q} = -Q_D^2 R + H \tag{2}$$

where Q_j is flow through a branch j, R_j are hydrodynamic resistances, H_j are pressure drops of the branches, $T_j=\rho l_j/S_j$ are inertia coefficients, j=1,…,n and n is the number of network branches (excluding the generator branch). $T = diag\{T_j\}, R = col\{R_j\}$ and

$$Q_D^2 = diag\{Q_j \left| Q_j \right|\} \tag{3}$$

Let n_c denote the number of nodes. Then $l=n-n_c+1$ is the number of links (excluding the generator branch) and n-l is the number of tree branches.

Like an electrical network, a fluid network must satisfy Kirchhoff's current law, i.e., the flow out of any node is equal to the flow into that node. Mathematically, Kirchhoff's current law for fluid flow networks can be expressed as[12]:

$$E_{Qin}\begin{bmatrix} Q_{in} \\ Q \end{bmatrix} = 0 \quad \text{or} \quad \sum_{j=1}^{n} E_{Qij} Q_j + e_{Qin_i} Q_{in} = 0, \ i = 1, \cdots, n-l \tag{4}$$

where n-l+1 is the number of nodes (of which one is a "reference" node), Q is a vector of flows, $E_{Qin} = [e_{Qin} \ E_Q]$, and E_Q=[E_{Qij}] is a full rank matrix of order (n-l)×n where E_{Qij}=1 if branch j is connected to node i and the flow goes away from node i, E_{Qij}=-1 if it goes into node i, E_{Qij}=0 if branch j is not connected to node i; e_{Qin} is an (n-l)×1 vector such that, if the generator is connected to node i and the flow goes away from node i then e_{Qini}=1, if the flow goes into node i then e_{Qini}=-1, and e_{Qini}=0 if the generator is not connected to node i.

Similarly, the network satisfies Kirchhoff's voltage law, i.e., the sum of the pressure drops around any loop in the network must be equal to zero, or mathematically

$$E_H H = 0, \ \text{or} \ \sum_{j=1}^{n} E_{Hij} H_j = 0, \qquad i = 1, \cdots, l, \tag{5}$$

where H_j is the pressure drop of the branch j, H is a vector of pressure drops, E_H=[E_{Hij}] is an l×n mesh matrix, in which each mesh (loop) is formed by a link and a unique chain in the tree connecting the two nodes of the link. The elements of E_{Hij} are defined as follows: E_{Hij}=1 if branch j is contained in mesh i and has the same direction, E_{Hij}=-1 if branch j is contained in mesh i and has the opposite direction, E_{Hij}=0 if branch j is not contained in mesh i.

In order to establish a dynamic model of minimal order, one has to find independent variables as states of the system. We take the flows of link (co-tree) branches as state variables. If regards one time heartbeat as one period T, decomposes the blood pressure wave f(t) into each kind of simple harmonic wave combination, namely:

$$Q_{in}(t) = Q_0 + \left(\sum_{k=1}^{n} a_k \sin(\frac{2\pi k}{T} t + \phi_k) \right) = Q_0 + \sum_{k=1}^{n} a_k \sin(k\omega t + \phi_k) \tag{6}$$

For convenience of analysis, we label the link branches (except the generator branch) from 1 to *l*. Define

$$Q = \begin{bmatrix} Q_c \\ Q_a \end{bmatrix}, \quad H = \begin{bmatrix} H_c \\ H_a \end{bmatrix} \tag{7}$$

so that Q_c and H_c vectors describe flow and pressure drop, respectively, in the links, excluding the generator branch, and Q_a and H_a vectors describe them in the tree branches.

The matrices E_H and E_Q in can be split into blocks

$$E_H = \begin{bmatrix} E_{Hc} & E_{Ha} \end{bmatrix} \tag{8}$$

$$E_{Qin} = \begin{bmatrix} e_{Qin} & E_{Qc} & E_{Qa} \end{bmatrix} \tag{9}$$

where [6-7]

$$E_{Qa} = I_{(n-l)\times(n-l)}, E_{Hc} = I_{l\times l}, E_{Ha} = -E_{Qc}^T \tag{10}$$

Hence, the structure of the network can be expressed in the matrix form as

$$E = \begin{bmatrix} 0 & I & -E_{Qc}^T \\ e_{Qin} & E_{Qc} & I \end{bmatrix} \tag{11}$$

Furthermore,

$$T = \begin{bmatrix} T_c & 0 \\ 0 & T_a \end{bmatrix}, R = \begin{bmatrix} R_c^T & R_a^T \end{bmatrix}^T \tag{12}$$

Fluid circulation through the network of network modeling, according to the aforementioned study, using the average method can solve the flow waveform, and then find its pulse wave flow waveform, namely:

$$\overline{Q}_c(t) = Q_{c0} - \left(\sum_{k=1}^{n} \frac{a_k^2}{4} \right) V^{-1} U + B_c \left(\sum_{k=1}^{n} a_k \sin(k\omega t + \phi_k) \right) \tag{13}$$

$$\overline{Q}_a(t) = (-E_{Qc} Q_{c0} - e_{Qin} Q_0) + \left(\sum_{k=1}^{n} \frac{a_k^2}{4} \right) E_{Qc} V^{-1} U + B_c \left(\sum_{k=1}^{n} a_k \sin(k\omega t + \phi_k) \right) \tag{14}$$

where $T_0(T) = T_c + E_{Qc}^T T_a E_{Qc}$, $U(R,T,E) = col\{B_{ci}^2 R_{ci}\} - E_{Qc}^T col\{B_{ai}^2 R_{ai}\}$, $V(R,E,Q_0) = diag\{Q_{c0i} R_{ci}\} - E_{Qc}^T W$, $W = \{E_{Qcij}(-E_{Qci} Q_{c0} - e_{Qin_i} Q_0) R_{ci}\}_{(n-l)\times l}$, and Q_{c0} denotes *l*-dimensional solution of quadratic equation, that is:

$$Q_{c0D}^2 R_c - E_{Qc}^T diag\{(E_{Qci} Q_{c0} + e_{Qin_i} Q_0)^2\} R_a = 0 \tag{15}$$

such that V is nonsingular and $-T_0^{-1} V$ is Hurwitz. Then for a given $Q_0 > 0$, for sufficiently small a and sufficiently large ω the solutions of the system (1)~(6) locally exponentially converge to an $O(1/\omega + a^4)$ neighborhood.

3 Cerebral Circulation K Value Computing and Verification

Generally we measured radial artery pulse wave, because of its high flow, easy to measure, but considering that the circulatory system is large and complex, its model is difficult to solve, because the arterial pulse wave is constant in the transmission cycle, Therefore, we use the above method to cerebral circulation network to be solved. Anatomy of the cerebral circulation is shown in Figure 1, the cerebral circulation equivalent plane structures shown in Figure 1[2-3].

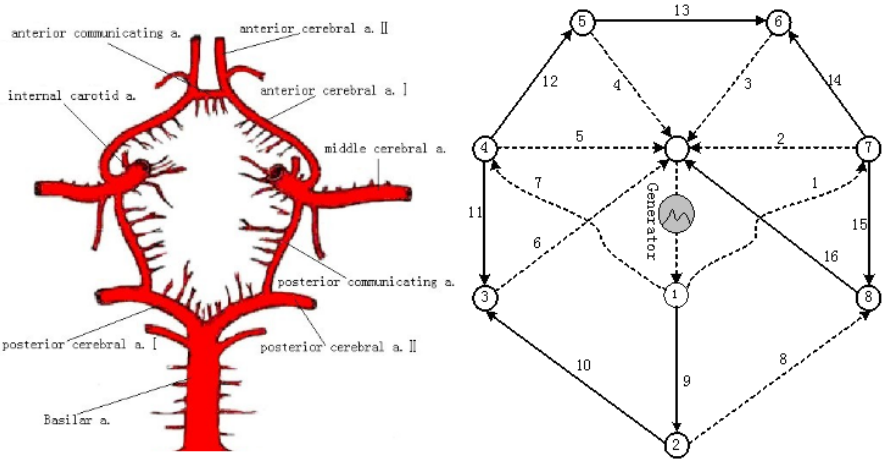


Fig. 1. The network equivalent plane diagram of the circle of Willis

We knew l,d_1,ρ,of the cerebral circulation network from paper[2], obtains T from $T=\rho l/S$ and R from $R=1.63\ l/\mathrm{D}^4$. We may obtain Q_{c0} equation set from type (19), this equation only has numerical solution, but does not have the exact solution, uses the genetic algorithm to solve the iterative solution.

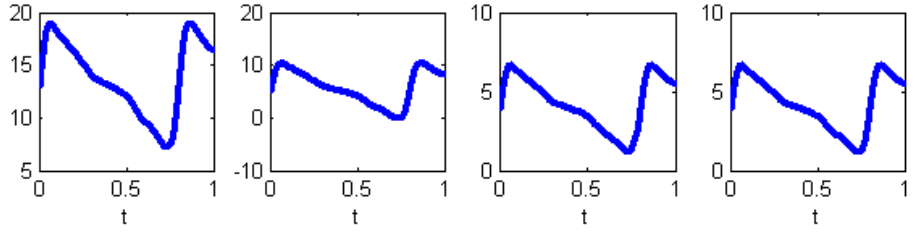


Fig. 2. The Cerebral circulation pulse wave with the normal human

We solve the cerebral circulation blood flow Q with the normal person, its computing simulation result is shown in figure 2. From figure 2, the normal cerebral circulation network 10 times harmonic waveform are basically the same with healthy

middle-aged in figure 2, the K value is 0.356, the K value is 0.356, and K is clinically consistent with 0.34 ~ 0.39, it is shown in figure 3[1].

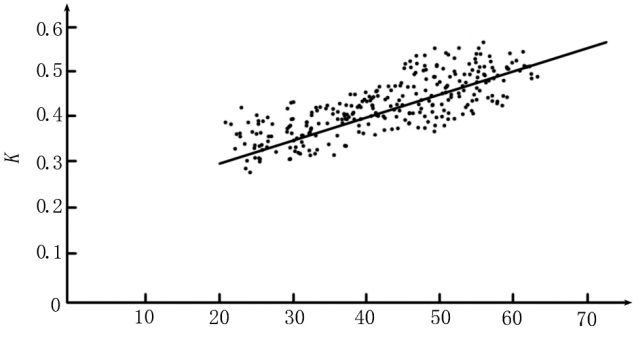


Fig. 3. Maps of K values in the different age groups

4 Conclusion

In summary, the pulse waveforms extracted the K value by averaging computation, although it does not fully reflect subtle changes in the pulse curve that contains all the local physiological and pathological significance, but it does represent some important physiological parameters in the human blood circulatory system, such as peripheral vascular resistance, blood viscosity and so on. Considering the characteristic information to reduce only one characteristic quantity K, it is easy to remember, a clear physiological significance, and change very regular, easily accepted by clinicians, so it can be used as an important cardiovascular physiological parameters of the clinical examination.

References

[1] Luo, Z., Zhang, S., Yang, W.: A research on characteristic information of pulse wave. Journal of Beijing Polytechnic University 22(1), 71–79 (1996)
[2] Ding, G.-H., Yao, W., Wang, Y.-B.: A Hemodynmic Model and A Mathematical Method to calculate the Dynamics indexes for Cerebral Circulation. Journal of Hydrodynamics, Ser. B 9(4), 71–78 (1997)
[3] Hillen, B., Hoogstraten, H.W., Post, L.: A mathematical model of the flow in the circle of Willis. Journal of Biomechanics 19(03), 187–194 (1986)
[4] Koroleva, O.I., Krstic, M.: Averaging analysis of periodically forced Fluid flow networks. Automatica 41(1), 129–135 (2005)
[5] Hu, Y., Koroleva, O.I., Krstic, M.: Nonlinear control of mine ventilation networks. Systems and Control Letters 49(4), 239–254 (2003)
[6] Khalil, H.K.: Nonlinear Systems. Prentice Hall, New Jersey (1987)
[7] Desoer, C.A., Kuh, E.S.: Basic Circuit Theory. McGraw-hill, Columbus (1969)

Conditional Fault Diagnosis of Bubble Sort Graphs under the PMC Model

Shuming Zhou[1,3], Jian Wang[2], Xirong Xu[2], and Jun-Ming Xu[3]

[1] Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou, Fujian 350007, China
[2] Department of Computer Science, Dalian University of Technology, Dalian, 116024, China
[3] Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, 230026, China

Abstract. As the size of a multiprocessor system increases, processor failure is inevitable, and fault identification in such a system is crucial for reliable computing. The fault diagnosis is the process of identifying faulty processors in a multiprocessor system through testing. The conditional diagnosability, which is a new metric for evaluating fault tolerance of such systems, assumes that every faulty set does not contain all neighbors of any processor in the systems. This paper shows that the conditional diagnosability of bubble sort graphs B_n under the PMC model is $4n$-11 for $n\geq 4$, which is about four times its ordinary diagnosability under the PMC model.

Keywords: Conditional diagnosability, Fault diagnosis, Bubble sort graphs.

1 Introduction

With the rapid development of multi-processor systems, fault diagnosis of interconnection networks has become increasingly prominent. As a significant increase in the number of processors, processor failure is inevitable. In order to ensure the stable running of the systems, we must find out the faulty processors and repair or replace them. System-level diagnosis, as a powerful tool, has been widely used. The basic idea is to design an effective algorithm to find out faulty processors through a comprehensive analysis of test results which are stimulated by adjacent processors. This method does not have to use special equipment to test. Most of the recent research efforts in system-level diagnosis have focused on enhancing the applicability of system-level diagnosis-based approaches to practical scenarios such as VLSI testing, diagnosis of interconnection networks employed in parallel computers [1]. The PMC model, proposed by Preparata et al. [2] for dealing with the system's self-diagnosis, assumed that each node can test its neighboring nodes, and test results are "faulty" or "fault-free". Under this model, the diagnosability of an interconnection network is the maximum number of faulty nodes in the system that can be identified. To grant more accurate measurement of diagnosability for a large-scale processing system, Lai et al. [1] introduce the conditional diagnosability of a system under the PMC model, which suppose the probability that all adjacent nodes of one node are faulty simultaneously

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 53–59.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

is very small. That is, conditional diagnosability is the diagnoability under the condition that all adjacent nodes of any node can't be faulty simultaneously. They further showed that the conditional diagnosability of Q_n is 4(n-2)+1 for $n \geq 5$. The conditional diagnosabilitities of matching composition networks[3,4], Shuffle-cubes [5], folded hypercubes[6], augmented cubes [7] are obtained in succession. This paper establishes the conditional diagnosability of the bubble-sort graph B_n under the PMC model. The remainder of this paper is organized as follows. In Section 2, we introduce some terminology and preliminaries used through this paper. Section 3 concentrates on the conditional diagnosability of B_n.

2 Terminologies and Preliminaries

A multi-processor system, whose topological structure is an interconnection network, can be modeled as a simple undirected graph $G = G(V, E)$, where a vertex $u \in V$ represents a processor and an edge $(u,v) \in E$ represents a link between vertices u and v. If at least one end of an edge is faulty, the edge is said to be faulty; otherwise, the edge is said to be fault-free. The connectivity of a graph G, denoted by $\kappa(G)$, is the minimum number of vertices whose deletion results in a disconnected graph or a trivial graph. The components of a graph G are its maximal connected subgraphs. A component is trivial if it has no edges; otherwise, it is nontrivial. The neighborhood set of the vertex set $X \subset V(G)$ is defined as $N_G(X) = \{ y \in V(G) \mid \exists x \in X$ such that $(x, y) \in E(G)\} - X$. For convenience, |S| denotes the number of elements in the set S. And we also use $|G|$ to represent the number of vertices in the graph G.

The PMC model requires that u and v can test each other for any edge $(u,v) \in E$. When u tests v, we call u as testing node, and call v as tested node. The test output is 0 (or 1) which implies that v is faulty (or faulty-free). $\sigma(u,v)$ denotes the output of u testing v. And it is assumed that the test outputs are correct if the testing node is fault-free; otherwise the outputs are unreliable.

The collection of all outputs is called the syndrome σ. For a given syndrome σ, a subset of vertices $F \subset V(G)$ is said to be consistent with σ if the syndrome σ can be produced from the situation that, for any $(u,v) \in E$ such that $u \in V\text{-}F$, $\sigma(u,v) = 1$ if and only if $v \in F$. It means that F is a possible set of faulty nodes. Since test output produced by a faulty node is unreliable, a given set F of faulty nodes may produce different syndromes. On the other hand, different faulty sets may produce the same syndrome. Let $\sigma(F)$ represent the set of all syndromes that could be produced by F. Two distinct sets F_1 and F_2 of $V(G)$ are said to be distinguishable if $\sigma_{F_1} \bigcap \sigma_{F_2} \neq \varnothing$; otherwise, F_1 and F_2 are indistinguishable. We say that (F_1, F_2) is a distinguishable pair if $\sigma_{F_1} \bigcap \sigma_{F_2} = \varnothing$; otherwise, ($F_1$, F_2) is an indistinguishable pair. We also use $F_1 \Delta F_2 = (F_1 - F_2) \bigcup (F_2 - F_1)$ to denote the symmetric difference of F_1 and F_2.

Definition 1. [1,2] A system G is said to be t-diagnosable if, a given syndrome can be produced by a unique faulty set, provided that the number of faulty nodes does not

exceed t. The largest value of t, for which a given system G is t-diagnosable, is called the diagnosability of system G, denoted as t(G).

Lemma 1. [1,2] For any two distinct sets F_1 and F_2 of V, (F_1, F_2) is a distinguishable pair iff there exist a vertex $u \in V(G)-(F_1 \bigcup F_2)$ and a vertex $v \in F_1 \Delta F_2$ such that (u,v) $\in$ E.

So, if two sets F_1 and F_2 are indistinguishable, then there is no edge between $F_1 \Delta F_2$ and $V(G)-(F_1 \bigcup F_2)$.

Definition 2. [1] A faulty set $F \subset V(G)$ is called a conditional fault-set, if $N(v) \not\subset F$ for any vertex v $\in$ V(G).

Definition 3. [1] A system G is said to be conditionally t-diagnosable, if for any two distinct conditional fault-sets F_1 and $F_2 \subset V(G)$ with $|F_1| \leq t, |F_2| \leq t$, ($F_1$, F_2) is a distinguishable pair. The largest value of t which makes system G is conditionally t-diagnosable is called the conditional diagnosability of system G, denoted as $t_c(G)$.

Lemma 2. [1] A system G is t-diagnosable under the PMC model, if f for any F_1 and $F_2 \subset V(G)$, $F_1 \neq F_2$ with $|F_1| \leq t, |F_2| \leq t$, ($F_1$, F_2) is a distinguishable pair.

Lemma 3. [1] A system G is t-diagnosable under the PMC model, if and only if for an indistinguishable pair of sets F_1 and $F_2 \subset V(G)$, it implies that $|F_1| > t$ or $|F_2| > t$.

Lemma 4 [1] Let $G(V,E)$ be a multi-processor system, and (F_1, F_2) be an indistinguishable conditional pair with $F_1 \neq F_2$, then the following conditions hold:

(1) $|N(u) \bigcap (V - (F_1 \bigcup F_2)| \geq 1$ for $u \in (V - (F_1 \bigcup F_2)$;
(2) $|N(v) \bigcap (F_1 - F_2)| \geq 1$ and $|N(v) \bigcap (F_2 - F_1)| \geq 1$ for $v \in F_1 \Delta F_2$.

Let (F_1, F_2) be an indistinguishable conditional pair, and let $S = F_1 \bigcap F_2$. By observation, every component of G-S is nontrivial. Moreover, we have

(1) for each component C_1 of G-S, if $C_1 \bigcap (F_1 \Delta F_2) \neq \varnothing$, then $d(C_1) \geq 1$ for $v \in C_1$;
(2) for each component C_2 of G-S, if $C_2 \bigcap (F_1 \Delta F_2) \neq \varnothing$, then $d(C_2) \geq 2$ for $v \in C_2$.

Network reliability is one of the major factors in designing the topology of an interconnection network. The hypercube and its variants were the major class of networks. The n-star graph (S_n for short) is an attractive alternative to the hypercube [8]. The bubble-sort graphs similar to the n-star graph, which belongs to the class of Cayley graphs, have been attractive alternative to the hypercubes. They have some good topological properties such as highly symmetry and recursive structure. In particular, the n-dimensional bubble-sort graph B_n is vertex transitive, while it is not edge transitive [9]. The connectivity of B_n is n-1 and the diameter is $n(n-1)/2$. It was shown that finding a shortest path in B_n can be accomplished by using the familiar bubble-sort algorithm [8]. Algorithms for node-to- node disjoint paths, pancyclicity and hamiltonian laceablity in B_n are obtained in [10,11,12]. Now, we introduce the bubble-sort

graphs. An n-dimensional bubble-sort graph is (n-1)-regular and symmetric. It has $n!$ nodes and $(n-1)n!/2$ edges while its connectivity and diameter are n-1 and $(n-1)n/2$, respectively.

Definition 4. An n-bubble-sort graph B_n has $n!$ nodes. Each node has a unique address, which is a permutation of n symbols $i=1,2,...,n$. A node that has an address $u=u_1u_2...u_n$ is adjacent to node whose address is $u^i=u_1u_2...u_{i-1}u_{i+1}u_i...u_n$ with $1\le i\le n-1$.

A very important property of the bubble-sort graph is its recursive structure [2]. We decompose B_n into n subgraphs B_n^i ($i=1,2,...,n$) such that each B_n^i fixes i in the last position of the label strings which represents the vertices, and so B_n^i is isomorphic to B_{n-1}. Let $S_i=S\bigcap B_n^i$ for $i=1,2,...,n$. For $1\le i\le n$, the ith element of the label of vertex u in B_n is represented by $u[i]$. An edge e=xy is called a pair-edge if $x[n]=y[n]$ and $x[n-1]=y[n-1]$. e=$x'y'$ is called the coupled pair-edge corresponding to e=xy, where $x'[n]=x[n-1]$, and $y'[n]=y[n-1]$. We call two edge, xx' and yy', the coupler or two pair-edge e=xy and e'=$x'y'$.

3 Conditional Diagnosability of B_n

We decompose B_n into n subgraphs B_n^i ($i=1,2,...,n$) such that all vertices of B_n^i have the same last bit i of the label strings, and so $B_n^i\cong B_{n-1}$. Let S be a faulty set of $V(B_n)$. Denote $A_1=\{B_n^i \mid B_n^i$ contains at least n-2 nodes in $S\}$, and $A_2=\{B_n^i \mid B_n^i$ contains at most n-3 nodes in $S\}$. We also denote A_2 the subgraph of B_n induced by the union of subgraphs in A_2. It is easy to check that the following lemma holds.

Lemma 5. A_2-S is connected.

Lemma 6. $t_C(B_n)\le 4n-11$.

Proof. Let e=xy be one pair-edge with the coupled pair-edge e'=$x'y'$. These two pair-edges with their coupler constitute a cycle of length of 4. Obviously, $|N_{B_n}(x,y,x',y')|=4(n-3)$. Let $F_1=N(x,y,x',y')\bigcup\{x,y\}$, $F_2=N(x,y,x',y')\bigcup\{x',y'\}$. It is easy to check that F_1 and F_2 are two indistinguishable conditional fault-sets, and $|F_1|=|F_2|=4(n-1-2)+2=4n-10$. Thus, $t_C(B_n)\le 4n-11$.

Now, we show the conditional diagnosability of B_n is 4n-11 for n≥5. Let F_1 and $F_2\subset V(G)$, $F_1\ne F_2$, (F_1, F_2) is a distinguishable pair for n≥5. We shall show our result by proving that either with $|F_1|\ge 4n-10$ or $|F_2|\ge 4n-10$.

Lemma 7. For any two indistinguishable conditional fault-sets F_1 and F_2 in B_n with $n\geq 5$, which satisfies $F_1 \neq F_2$, we have either $|F_1| \geq 4n-10$ or $|F_2| \geq 4n-10$.

Proof. Since (F_1, F_2) is an indistinguishable conditional-pair, there exists no edge between $F_1 \Delta F_2$ and $B_n-(F_1 \cup F_2)$ by Lemma 1. Define $S = F_1 \cap F_2$ and let $B_n[F_1 \Delta F_2]$ be the subgraph of B_n induced by the vertex set $F_1 \Delta F_2$. We choose a maximal component C in $B_n[F_1 \Delta F_2]$ when $B_n[F_1 \Delta F_2]$ is not connected; otherwise, let $C = B_n[F_1 \Delta F_2]$. By lemma 4, we have $|C| \geq 4$. Thus, we only need to prove $\lceil |C|/2 \rceil + |S| \geq 4n-10$, which implies that $|F_1| \geq 4n-10$ or $|F_2| \geq 4n-10$.

We decompose B_n into n subgraphs B_n^i $(i = 1, 2, ..., n)$ such that each B_n^i fixes i in the last position of the label strings which represents the vertices. We also let $S_i = S \cap B_n^i$.

If $|S| \geq 4n-12$, then we have $\lceil |C|/2 \rceil + |S| \geq 4n-10$ for $|C| \geq 4$, so the lemma holds. Now, we only consider the situation $|S| \leq 4n-13$.

Let $A_1 = \{B_n^i \mid B_n^i$ has at least n-2 nodes in $S\}$, and $A_2 = \{B_n^i \mid B_n^i$ has at most n-3 nodes in $S\}$. Obviously, A_1 has at most three elements by the fact that $4(n-2) > 4n-12$.

If $A_1 = \varnothing$, by lemma 5 we have $|C| \geq (n-3)[(n-1)!-(n-3)] > 2(4n-12)$ for $n \geq 5$. Thus, we have $\lceil |C|/2 \rceil + |S| \geq 4n-10$. Now, we consider $A_1 \neq \varnothing$ as follows.

Case 1. $C \cap (A_2 - S) \neq \varnothing$.

Since A_1 has at most three subgraphs, we have $|C| \geq (n-3)[(n-1)!-(n-3)] \geq 2(4n-10)$ for $n \geq 5$; and we arrive at the result.

Case 2. $C \cap (A_2 - S) = \varnothing$.

Obviously, $C \subset A_1$, $N_{A_2}(C) \subset S \cap A_2$ by the maximality of C. Now, we divide this case into three subcases below.

Subcase 2.1. There is exactly one subgraph, say X, in A_1.

Since every vertex of $C \cap X$ has exactly one neighbor outside of X, we have $N_{B_n-X}(C) \subset S \cap A_2$; and so $|N_{B_n-X}(C)| = |C|$. Obviously, $N_X(C) \subset S \cap X$. Since $|C| \geq 4$, let T be a path of length three in C. Obviously, $N_X(C) \supseteq N_X(T) - (C-T)$, and so $|N_X(C)| \geq |N_X(T)| - |C-T|$.

Since $S = (S \cap X) \cup (S \cap A_2)$, we have $|S| = |S \cap X| + |S \cap A_2| \geq |N_X(C)| + |N_{B_n-X}(C)| \geq 4n-12$. Thus, we have $\lceil |C|/2 \rceil + |S| \geq 4n-10$.

Subcase 2.2. There are exactly two subgraphs, say X and Y, in A_1.

We assume, without loss of generality, that $|C\cap X|\geq|C\cap Y|$. For any $x\in C\cap X$, $|N_C(x)|\geq 2$ by Lemma 4 and every vertex of B_n^i has exactly one neighbor outside of this subgraph, then we have $|C\cap X|\geq|C\cap Y|\geq 2$. Due to $N_X(C\cap X)\subset S\cap X$, and $C\cap X$ has at least two vertices, $|S\cap X|\geq|N_X(C)|\geq 2(n-2)-2$. Similarly, we have $|S\cap Y|\geq|N_Y(C)|\geq 2(n-2)-2$. Since every pair of $S\cap X$, $S\cap Y$, $S\cap A_2$ are disjoint, we have $S=(S\cap X)\cup(S\cap Y)\cup(S\cap A_2)\supseteq N_X(C\cap X)\cup N_Y(C\cap Y)\cup S\cap A_2$. Then we have $|S|\geq|N_X(C\cap X)|+|N_Y(C\cap Y)|+|S\cap A_2|\geq 2[2(n-2)-2]+|S\cap A_2|$.

Thus, we have $\lceil|C|/2\rceil+|S|\geq 4n-10$.

Subcase 2.3. There are exactly three subgraphs, say X, Y, Z, in A_1.

We assume, without loss of generality, that $|C\cap X|\geq 2$ by the fact that $|C|\geq 4$. We have $|F\cap X|\geq|N_X(C\cap X)|\geq 2(n-2)-2$. Since every vertex of B_n^i has exactly one neighbor outside of this subgraph, and every pair of $S\cap X$, $S\cap Y$, $S\cap Z$, $S\cap A_2$ are disjoint, we have $S=(S\cap X)\cup(S\cap Y)\cup(S\cap Z)\cup(S\cap A_2)\supseteq N_X(C\cap X)\cup(C\cap Y)\cup(C\cap Z)\cup S\cap A_2$. Then we have $|S|\geq|N_X(C\cap X)|+|C\cap Y\|C\cap Z|+|S\cap A_2|\geq 2(n-2)-2+2(n-2)+|S\cap A_2|\geq 4n-10$.

Thus, we have $\lceil|C|/2\rceil+|S|\geq 4n-10$.

By Lemma 6 and 7, we have

Theorem 1. The conditional diagnosability of bubble sort graph B_n under the PMC model is $t_C(B_n)=4n-11$ ($n\geq 5$).

Theorem 2. The conditional diagnosability of bubble sort graph B_4 under the PMC model is $t_C(B_4)=5$.

Proof. Let F_1 and F_2 be two distinguishable fault-sets in B_4. Denote $S=F_1\cap F_2$ and $C\subset F_1\Delta F_2$ be a connected component in B_4-S. By Lemma 4, we have $|C|\geq 4$.

If $|S|\geq 4$, then $\lceil|C|/2\rceil+|S|\geq 6$. Now we suppose that $|S|\leq 3$. In the worst case, B_4-S has two components, one of which is an isolated vertex u, then we have $C=B_4-S-\{u\}$, which implies $|C|=20$. So we have $\lceil|C|/2\rceil+|S|\geq 10$. Thus $t_C(B_4)\geq 5$, while $t_C(B_4)\leq 5$, hence $t_C(B_4)=5$.

References

1. Lai, P.-L., Tan, J.J.M., Chang, C.-P., Hsu, L.-H.: Conditional diagnosability measures for large multiprocessor systems. IEEE Transactions on Computers 54(2), 165–175 (2005)
2. Preparata, F.P., Metze, G., Chien, R.T.: On the connection assignment problem of diagnosable systems. IEEE Transactions on Computers 16, 848–854 (1967)
3. Xu, M., Thulasiraman, K., Hu, X.: Conditional Diagnosability of Matching Composition Networks Under the PMC Model. IEEE Transactions on Circuits and Systems-II: Express Briefs 56(11), 875–879 (2009)
4. Zhu, Q.: On conditional diagnosability and reliability of the BC networks. Journal of Supercomputing 45(2), 173–184 (2008)
5. Xu, M., Hu, X., Shang, S.: The conditional diagnosability of shuffle-cubes. Journal of System Science and Complexity 23, 81–90 (2010)
6. Zhu, Q., Liu, S.-Y., Xu, M.: On conditional diagnosability of the folded hypercubes. Information Sciences 45, 174–184 (2008)
7. Chang, N.-W., Hsieh, S.-Y.: Conditional Diagnosability of Augmented Cubes Under the PMC Model. IEEE Transactions on Dependable and Secure Computing 9(1), 46–60 (2012)
8. Akers, S.B., Kirshnamurthy, B.: A group-theoretic model for symmetric interconnection networks. IEEE Transactions on Computers 38(4), 555–565 (1989)
9. Lakshmivarahan, S., Jwo, J., Dhall, S.K.: Symmetry in interconnection networks based on Cayley graphs of permutation groups: A survey. Parallel Computing 19(4), 361–407 (1993)
10. Suzuki, Y., Kaneko, K.: An Algorithm for Disjoint Paths in Bubble-Sort Graphs. Systems and Computers in Japan 37(12), 27–32 (2006)
11. Araki, T., Kikuchi, Y.: Hamiltonian laceability of bubble-sort graphs with edge faults. Information Sciences 177(13), 2679–2691 (2007)
12. Kikuchi, Y., Araki, T.: Edge-bipancyclicity and edge-fault-tolerant bipancyclicity of bubble-sort graphs. Information Processing Letters 100, 52–59 (2006)

Process Simulation of Micro Device with Virtual Reality Technology

Zheng Liu[1] and Hua Chen[2]

[1] School of Mechatronic Engineering, Xi'an Technological University, Xi'an China
[2] Xi'an Technological University, Xi'an China
zheng.liumail@gmail.com,
chenhua126@163.com

Abstract. The process design of micro device is more intuitive and efficient with the fabricating feature technology. To improve manufacturability of the feature-based process model, the process simulating method is presented, which is based on 3D modeling and compatible with virtual reality environment. Firstly, the relationship between the simulating model and the feature modeling is presented. Secondly, the simulating framework is constructed with both surface and bulk micro mechanizing technology. With java3D technology, the feature-based models in both design and process stages are intuitively represented. Moreover, it provides the linkage for the feature model data structure and virtual reality configuration. Finally, the 3D simulating model is combined with the virtual reality environment.

Keywords: Process simulation, Virtual reality, Feature modeling, Micro device.

1 Introduction

The virtual reality technology provides effective way in intuitive and interactive environment construction. Recently, it is used in various stages of micro device design. To improve efficiency, a physical cell method is presented to aid in the assembly of micro devices, which is interfaced with the virtual reality environment to improve the micro assembly activities [1]. On the other hand, the micro fabricating simulation has the same intuitive requirement with the micro assembly. To support the Top-Down design flow, some researches are implemented within the system level simulation. These simulations are generally provided at abstraction levels, such as VHDL-AMS method for coupled distributed physical models [2]. Moreover, the feature technology provides contribution for the consistency of system level simulation and process level design [3]. However, for the physical simulation, there are still problems to solve. Among them, the intuitive and efficient modeling method is the key link. The three dimensional dynamic CA method increases the simulation speed. The simulating software based on it provides the 3D visualizing environment for the micro machining processes [4,5]. In addition, the JAVA3D technology possesses the born advantages in designing the distributed virtual geographic

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 61–65.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

environment, which is the foundation to construct the cooperative design environment for the micro device design to improve the manufacturability [6].

To enhance the virtual design method of micro device, efforts have been made to improve the simulation process. Because of the different modeling method between the fabricating features and the simulating representation, how to build architecture to interface the feature modeling and the simulating process is the key problem. This paper is intended to address the key issues for constructing the 3D simulating method in accordance with the virtual reality configuration by means of JAVA3D technology.

2 Design and Simulation Framework for Feature Modeling

To improve the design efficiency, feature modeling framework is constructed in both the design and the process stages. The process level modeling is shown in Fig. 1. The fabricating processes simulation and the virtual reality support environment are important sections for the process level design. In this paper, the deepened parts are discussed.

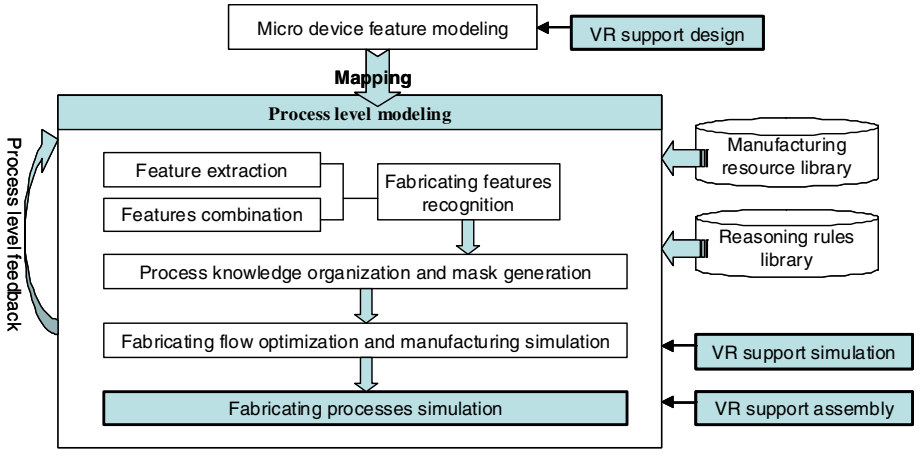


Fig. 1. The feature-based process level modeling framework

Above all, the designing feature model of micro device is mapped into the process model that is organized with fabricating features. The mapping strategies involve the feature extraction, the features combination and the fabricating features recognition. With the Ontology-based knowledge representing way, the process framework is constructed. As one attribute of the etching feature, the mask is generated from the information of solid model that is generated with operating algorithm. Because of the multiple mapping results, the vinous fabricating flows are generated. The optimization is the rule-based reasoning process with the manufacturing resource library and reasoning rules library support. After the optimization of the fabricating flow, the simulation is performed, which is also feature based and 3D visualized.

Throughout the design stages, virtual reality environment is supportive to improve the intuitive design.

3 Simulating Method for Bulk and Surface Micromachined Device

The simulation involves bulk micromachining and surface micromachining method. The relation between the simulation and the 3D modeling is illustrated with Fig. 2. The solid model representation is based on JAVA3D technology, which constructs the 3D supporting platform with virtual reality environment. All the fabricating features and simulating operations are constructed on the platform. The simulating processes for bulk micromachining involve substrate generation, bonding operation, etching operation and lithography operation. For surface micromachining, deposition operation and sacrificial layer generation are also simulated, which is generally geometric simulation. While, for bulk etching simulation, physical fabricating simulation is necessary to illustrate the real fabricating situation.

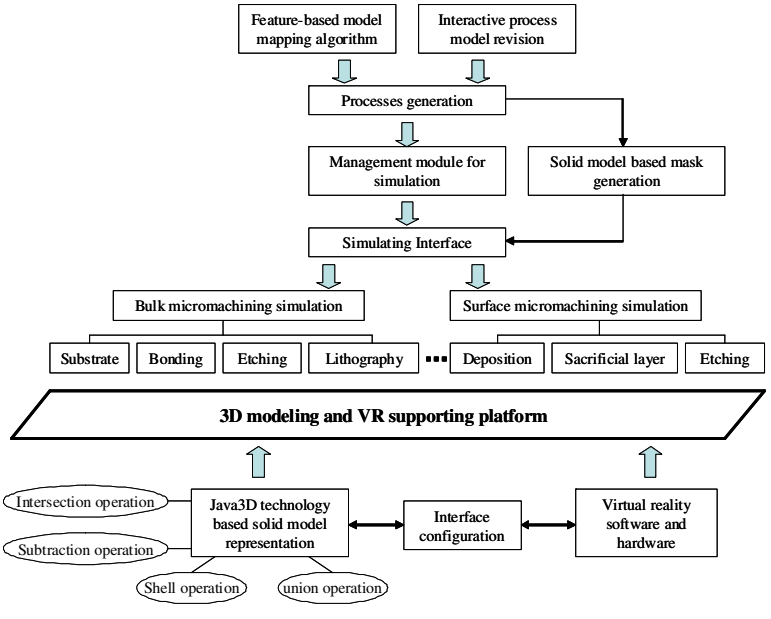


Fig. 2. The simulation process with 3D modeling and VR supporting platform

4 The Wet Etching Simulation Based on Java3d Technology

The wet etching process is simulated with physical parameters, which are dynamic virtual fabricating processes. Each step is modelled with both mathematic algorithm

and 3D representation. On one hand, the cellular automata method is adopted to calculate the situation of fabrication, and on the other, the JAVA3D technology is adopted to construct the 3D visualization.

4.1 Feature Modeling With Java3d Data Structure

JAVA3D API is the interface defined for 3D display. The relationship between feature modeling and JAVA3D representation is shown in Fig. 3. The virtual universe is on the top level, which is built by the SimpleUniverse class. It is the basis of the scene graph. Each Local below has one or more of BranchGroup nodes. The constructing principle for shape is the same, which contains an Appearance and an Array. The features of CSG tree model are connected with the JAVA3D shapes to realize the 3D visualization.

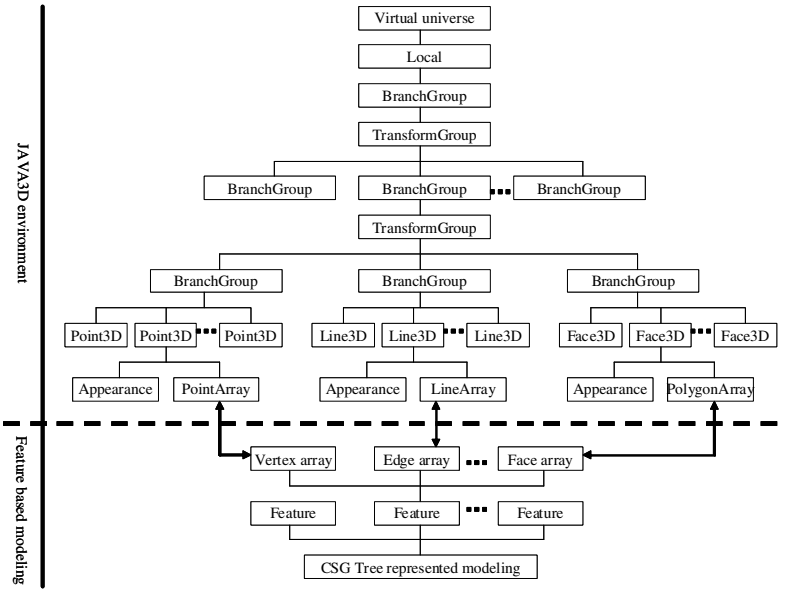


Fig. 3. The relationship between feature modeling and JAVA3D representation

4.2 Simulating Method in 3D Environment

The cellular automata method can be applied in physical simulating process of micro device, such as silicon anisotropic etching simulation. The cellular automata method is a system, in which the space, time and states are discrete. Each variable takes only finite number of states. With this method, the complex physical system is simulated with simplified rules.

5 The Interface Construction for VR Environment

To develop VR environment, the interface for input and display devices is required to design firstly. The common VR application must track the user's head position and orientation to generate the head position within the virtual scene graph. In addition, other parts of the body (hands, arms or legs) are also tracked to interact with the objects in the virtual scene. Java3D API integrates the data of head tracker into the graphics information generated. It works by the view model technology, which connect the physical information of users with the virtual environment and make benefit for the interface construction.

6 Conclusion

In this paper, the method to simulate the fabricating processes of micro device in a intuitive 3D way is suggested. With this method, the feature based modeling method of both design and process stage is combined with the physical simulating process in the virtual reality environment to improve manufacturability. The extension of this work can lead to a compatible feature-based design framework for micro device design. In addition, JAVA3D technology provides benefit for the prospective corporative design theory in distribute 3D environment construction. More work is underway to improve the efficiency of simulating process and the fabricating features representation.

Acknowledgments. This work was financially supported by the Science and Technology Development Plan Foundation of Shaanxi Province (No. 2011K07-11), Scientific Research Program Funded by Shaanxi Provincial Education Department (Program No. 11JK0864), President Fund of Xi'an Technological University (No. XAGDXJJ1007) and Shaanxi Major Subject Construction Project.

References

1. Cecil, J., Gobinath, N.: Development of a virtual and physical work cell to assemble micro-devices. Robotics and Computer-Integrated Manufacturing 21, 431–441 (2005)
2. Chapuis, Y.A., Zhou, L., Fujita, H., Hervé, Y.: Multi-domain simulation using VHDL-AMS for distributed MEMS in functional environment: Case of a 2D air-jet micromanipulator. Sensors and Actuators A: Physical 148, 224–238 (2008)
3. Liu, Y., Jiang, P., Zhang, D.: 3D-feature-based structure design for silicon fabrication of micro devices. Microsystem Technologies 13, 701–714 (2007)
4. Zhu, Z., Liu, C.: Micromachining process simulation using a continuous cellular automata method. Machinery Design & Manufacture 9, 252–261 (2000)
5. Zhou, Z., Huang, Q., Li, W.: Application of cellular automata method in process simulation of MEMS. Journal of Transducer Technology 23, 77–80 (2004)
6. Zhang, J., Gong, J., Lin, H., Wang, G., Huang, J., Zhu, J., Xu, B., Teng, J.: Design and development of Distributed Virtual Geographic Environment system based on web services. Journal of Engineering Graphics 177, 3968–3980 (2007)

Intelligent Control of Beer Fermentation Temperature

Tianli Li

College of Electrical and Information Engineering, Shaanxi University of Science and Technology, China

Abstract. Aiming at the nonlinear and the characteristics of the temperature for beer fermentation process. A kind of human-simulated intelligent control is putted forward. This theory is applied in the about 60 beer fermentation tank control, the actual results found human-simulated intelligent control can be a good way to solve the temperature control, meet the technological requirements, produce high-quality products.

Keywords: beer fermentation, temperature, large inertia, nonlinear, intelligent control.

1 Introduction

Fermentation is a key link for brewing process,it is a complex biochemical reaction process,it is an important index beer quality formed step.The yeast fermentation process is the conversion process for wort extract,most of the (sugar) (occupy 2/3) change into ethanol and carbon dioxide ,this process will produce the massive hot.For example, 11BX wort for 250 cubic meters with cans of volume wort fermentation, When real fermentation degree aim at 65%, the heat generated by the 10485.5 MJ, for in this process,saccharification formed of fermentable wort sexual sugar,in yeast,under the action of generating alcohol and a series of beer flavor substances.The content of various flavor substances, endowed with beer, make the particular flavor, that is to say, fermentation decided the beer taste and quality stability[1]. Beer fermentation process divides into mainly fermentation (reduction)process,cooling process,well done (the wine) after process,its fermentation temperature process curve (see figure1). Fermentation temperature process curve is the process ensure complete fermentation[2]. Control temperature is essentially to control heat transfer,namely control with the outside world of heat exchange between fermentation. Automatic control task is to make sure that the fermentation process curve through the control Settings in three sections of independent 3 pick the refrigerant medium flow[3].

2 Human-Simulated Intelligent Control

2.1 The Necessity of Human-Simulated Intelligent Control

The volume of large cone bottom fermentation tank approximately 200-800 cubic meters. Controlled object has a large inertia, large delay properties, the change of

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 67–71.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

the temperature is a slow process. Due to the traditional PID control in different fermentation temperature stage parameters vary greatly, it is difficult to determine[4]. Because of its role in complicated cases, a long time factors affecting, and have a lot more factors can not have been measured, so it cans not establish the mathematical model, and the control effect is not ideal. Temperature control error is in 2 degrees Celsius,it cannot meet the technological requirements. But experienced operator can controll good control according to his own experience, We choose human-simulated intelligent fuzzy control[5], based on the experiences of those based on fuzzy control theory, the changes and according to actual error of the characteristics of temperature variations pick. the control performs device to use pneumatic butterfly valve.

2.2 Control Algorithm Design

We design the control thoughts,according to the mechanism of beer fermentation、fuzzy control theory and artificial control experience[6]:

(1) When the error e>=0 or -0.05<=e<0, the valve is closed to let fermentation tank warming.
(2) When the error change increases (Δe<0), extende the time of open valve.
(3) When the error change decrease (Δe>0), shorten the time of open valve.
(4) We adjuste valves open time,when a given value less than the adjustable amount (Δe<-0.05), according to the error size. We designs the following algorithm according to error e and error change:

$U_1(i)$, $U_2(i)$, $U_3(i)$ -- Respectively represent upper, middle and lower three sections of refrigerant valve control signal[7].

A,B,C,D-- Representing different process section of time (see figure 1)

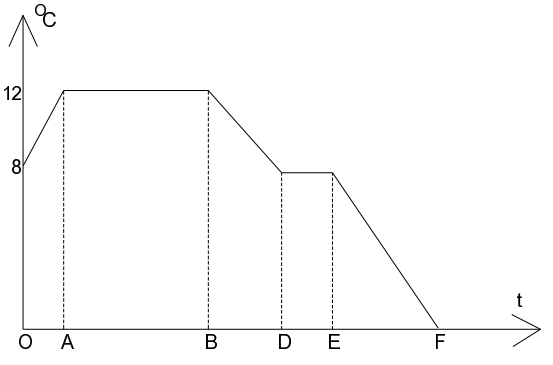


Fig. 1. Beer fermentation temperature process curve

$U_j(i)=1$, show upper valve open（J=1）middle valve open（J=2）lower valve open（J=3）$U_j(i)=0$, show upper valve shut（J=1）middle valve open（J=2）lower valve open (J=3);

(1) IF e>=0 OR -0.05<=e<0 THEN $U_1(i)=U_2(i)=U_3(i)=0$ $t_1(i)=t_2(i)=t_3(i)=0$

(2) IF 0<=t<=A AND e(i)<=-0.5 THEN $U_1(i)=U_2(i)=1,U_3(i)=0$ $t_1(i)=t_2(i)=T,t_3(i)=0$

-0.5<e(i)<-0.1 THEN $U_1(i)=U_2(i)=1,U_3(i)=0$

$t_1(i)=kt_1(i-1),t_2(i)=k_0t_2(i-1),$ $t_2(i)=0$

-0.1<e(i)<=-0.05 THEN $U_1(i)=U_2(i)=1,U_3(i)=0$

$t_1(i)=k_1t_1(i-1),t_2(i)=k_2t_2(i-1),$ $t_2(i)=0$

-0.1<e(i)<=-0.05 AND Δe<0 THEN $U_1(i)=U_2(i)=1,U_3(i)=0$

$t_1(i)=k_3t_1(i-1),t_2(i)=k_4t_2(i-1),t_2(i)=0$

(3) IF A<t<=B OR C<t<=D OR t>=E AND e(i)<=-0.5 THEN

$U_1(i)=U_2(i)=0,U_3(i)=1$ $t_1(i)=t_2(i)=0,t_3(i)=T2$

-0.5<e(i)<=-0.1 THEN $U_1(i)=U_2(i)=0,U_3(i)=1$ $t_1(i)=t_2(i)=0,t_3(i)=k_5t_3(i-1),$

-0.1<e(i)<=-0.05 AND Δe>0 THEN $U_1(i)=U_2(i)=0,U_3(i)=1$

$t_1(i)=t_2(i)=0,t_3(i)=k_6t_3(i-1)$

-0.1<e(i)<=-0.05 AND Δe<0 THEN $U_1(i)=U_2(i)=0,U_3(i)=1$

$t_1(i)=t_2(i)=0,t_3(i)=k_7t_3(i-1)$

(4) IF A<t<=C OR D<t<=E AND e(i)<=-0.5 THEN

$U_1(i)=U_2(i)=U_3(i)=1$ $t_1(i)=t_2(i)=t_3(i)=T3$

-0.5<e(i)<=-0.1 THEN $U_1(i)=U_2(i)=U_3(i)=1$ $t_1(i)=k_8t_1(i-1),t_2(i)=k_9t_2(i-1),$

$t_3(i)=k_{10}t_3(i-1)$

-0.1<e(i)<=-0.05 AND Δe>0 THEN $U_1(i)=U_2(i)=U_3(i)=1$

$t_1(i)=k_{11}t_1(i-1),t_2(i)=k_{12}t_2(i-1),$ $t_3(i)=k_{13}t_3(i-1)$

-0.1<e(i)<=-0.05 AND Δe<0 THEN

$U_1(i)=U_2(i)=U_3(i)=1t_1(i)=k_{14}t_1(i-1),t_2(i)=k_{15}t_2(i-1),$ $t_3(i)=k_{16}t_3(i-1)$

Explain：

(1) $U_1(i),U_2(i),U_3(i)$ --This control output.

$t_1(i)$ $t_2(i)$ $t_3(i)$ ； $t_1(i-1)$ $t_2(i-1)$ $t_3(i-1)$ ； --This and the last control output time;

$k\ldots k_{16},T1,T2,T3$ --According to experience setting parameters;

(2) Due to different, pick itself and equipment conditions and yeast vigor different fermentation 、 Refrigerant, pressure, temperature and other factors, $k\ldots k_{16},T1,T2,T3$ In setting parameters, these can appear multiple parameters, we suse different data access in different circumstance;

(3) According to experience, in different fermentation stage, we open different valve: Lord leaven period in the valve open, cooling period is opened entirely, heat preservation period the valve open.

2.3 Practical Application

(1) Controlled object: 60 360 cubic meters fermentation tank;

(2) Controlled object characteristics:

①60 fermentation tank are partial manufactured, so the upper, middle and lower refrigerant distribution is different, the upper, middle and lower temperature sensor installation position is different, refrigerant pipe diameters for different: 44 cans ,16 cans for DN65;

②60 fermentation tank use the same refrigerant tube, and its pressure and temperature difference is big;

③the same cans of the amount of time into different fermentation.

3 Actual Effect

3.1 Achieving Physical Index（Not Including Sensor Factors）

Temperature measurement error：≤± 0.1 ℃;
Temperature control error：≤± 0.3 ℃;
Pressure measurement error：≤± 0.1Kpa;
Pressure control error：≤± 0.02 Mpa;
Operating curve is smoothly, normally.

3.2 Achieving Physiochemical Index

1 The control effect

(1) The ensures of control accuracy requirements diacetyl required temperature reduction.
(2) The control effect is guaranteed the seamer of tank pressure constant after. Cooling process temperature uniformity make fermentation process produces CO_2 got fully stable since fall, the wine CO_2 content and to improve the beer kill mouth force.
(3) Intelligent temperature control avoided blindness for operation pick, guarantee the temperature uniformity, process parameters gets smooth realization.
(4) Application of the system make workers from heavy manual labor.Meanwhile, the system reduced the wine loss because of manual measuring temperature. This system makes the same can group at the same time to achieve different process parameters, producing different varieties of products. This system can be realized by different program control the development of new varieties of beer.

Table 1. Comparison table of main physiochemical index before and after the project implementation

before the project implementation			after the project implementation		
cans	diacetyl(mg/l)	Co_2 (%)	cans	diacetyl(mg/l)	Co_2 (%)
2#	0.09	0.35	2#	0.05	0.46
3#	0.10	0.36	3#	0.05	0.45
18#	0.08	0.36	18#	0.04	0.46
22#	0.09	0.40	22#	0.05	0.46
25#	0.09	0.35	25#	0.04	0.46
27#	0.11	0.35	27#	0.04	0.45
35#	0.07	0.38	35#	0.05	0.44
36#	0.09	0.37	36#	0.05	0.45
42#	1.09	0.55	42#	0.05	0.46
44#	0.09	0.54	44#	0.05	0.46

2 Main physical and chemical index contrast see table 1

4 Conclusion

The control system is applied in beer temperature control system, it overcomes disadvantage of traditional control, it is better to solve the fermentation temperature control.

References

[1] Peng, X., Dan, Y.-H.: Human simulated intelligent control of magnetic levitation systems. In: Proceedings of the World Congress on Intelligent Control and Automtion, pp. 1376–1381 (2010)
[2] Zhong, L., Liu, H.-Y., Wang, F., Xu, L.-X.: Overview of human-simulated intelligent control method and its application. International of Modelling Identification and Control 9(1-2), 161–167 (2010)
[3] Li, X.-B., Wang, H.-B.: Intellgent predicting control of vacuum freeze-drying temperature. Computer Engineering and Applications 46(30), 241–244 (2010)
[4] Wai, R.J., Ma, K., Lee, J.D.: Cascade direct adaptive pendulum servomechanism. IEEE Transaction on System, Man and Cyberrnetics 38(2), 439–454 (2008)
[5] Lv, H.-L., Duan, P.-Y., Jia, L.: One novel fuzzy controler design for HVAC system. In: Chinese Control and Decision Conference, pp. 2071–2076 (2008)
[6] Zhu, W.-M., Chen, Y.-H., Sha, L.-M.: Simulation study of central air conditioning controller. Industrial Instrumentation & Automation (2), 57–59 (2008)
[7] Jiang, L., Xiao, J., Huang, J.-C.: Multi-sampling rate fuzzy optimal control of nonlinear system. Control and Decision 23(4), 382–387 (2008)

Discussion on Enterprise Emergency Management Decision Support System

Wenchang Liu

School of Management, Liaoning University of Technology, China

Abstract. This article in research on emergency management theories, on the basis put forward specific content of emergency management and decision support systems, and is based on their impact time is divided into three classes of strategic, tactical, and operational, in accordance with the decision support system requirements is a key technology platform design and main features of the system and workflow design.

Keywords: Decision Support Systems, Emergency Management, Emergency Decision.

1 Introduction

For the uncertainty, complexity and urgency of emergency decision-making problems, decision makers depend on the computer system more than conventional decisions. But in determining the division of the EDSS and decision makers in emergency decision process, one have to follow the principle of man-machine capacity limitation, relying too much on decision makers or the EDSS will reduce the efficiency of emergency decision-making plan. EDSS should be used to assist rather than replace decision makers, through system's ability to extend decision makers' rational ability, and strengthen the decision makers' endurance to emergency, and decision makers will concentrate on using experience, intuition and enlightenment to judge and choose the key problems. By full preparation to alleviate the real-time task load.

2 Emergency Decision-Making Tasks and Classification

According to the time long to short of decisions' effects, it can be divided into three layers: strategy, tactics and operation. Strategy: Prevention strategies, such as anti-seismic standards, fire control standards of building regulations etc, they concrete answer what kind of laws, policies and institutions can alleviate the degree of emergency events harm; Emergency supplies and facilities planning, according to local emergency demand information, planning for supplies and emergency contingency ability to plan the address and facilities; Planning the post-disaster long-term reconstruction. Hazard risk assessment, estimating to probability of occurrence of accident and degree of undesirable consequence; Detailed plans' formulation,

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 73–77.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

including detail measures of all kinds of response from stem to stern, the specific performance is all kinds of person complete specific activities with a variety of materials after disasters as time goes by; Drill planning, according to the procedures of the scheduled plan to train and improve related personnels' emergency response ability; In addition, to win the struggle with disasters, not only depends on the construction of professional staff, but on the support of emergency supplies. So to keep different place different kinds of emergency supplies inventory level is necessary, which involves emergency supplies' inventory management. Operation: Real-time monitoring and warning of enterprise hazards, adopt various technical means to gain the information of hazards and surroundings, and gives corresponding early warning information according to the assessment process; Emergency resources' distribution plan, including in practical response and recovery stage, according to the disaster's demand information and combining the resources information to realize the distribution plan of materials, equipment and personnel; Response plan, with distribution scheduling program needs the optimization formulate specific scheduling solution in a certain time and transportation constraints; The affected personnel salvage and retreat scheme, according to the damage degree of disasters and the environmental conditions to provide specific paitient assistance program and ordinary personnel evacuation plan.

3 Design of Emergency Decision Support System

3.1 Overall Structure Design

In recent years, domestic enterprise security work patterns are widely adopted to the international mode, means and methods of emergency management have been greatly improved and enhanced, this article mainly give a brief design aimming at the emergency management in safety accidents. According to the analysis of enterprise's characteristics and accident consequences, enterprise emergency decision support systems is built with the purpose of: 1) to prevent accidents, real-time monitoring major hazard to nip the accident in the bud as far as possible; 2) when the accident happens, measures must be taken to halt the accident, to prevent accident deterioration and control accident with shortest period, and to keep state property and people's life free from or get less damage. The enterprise emergency decision support system based on system's main purpose is to use the computer to compiled and store according to original data of the identified major hazards and hazards around and the emergency relief scheme that can be implemented after the accident. Once the accident happends, data will be collect and input immediately, the accident emergency decision-making group will be provided with the best emergency decision-making plan , through computer's logic computation and reasoning, which provide support for commanders' decision of emergency rescue. The overall structure of enterprise emergency decision support system is as figure 1 shows.

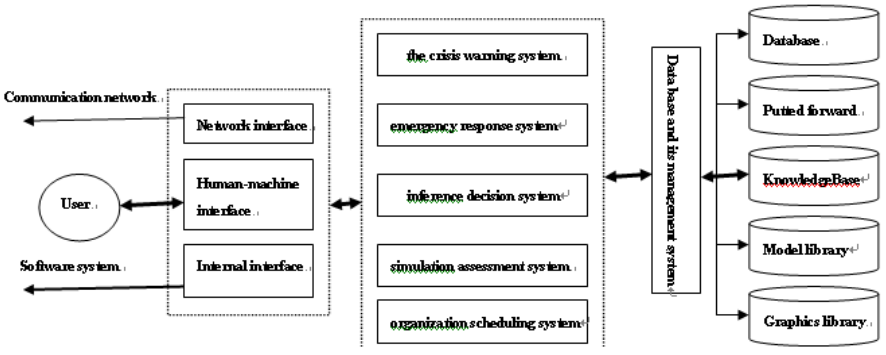


Fig. 1. Emergency decision support system overall structure

1. interface system. Interface system consists of man-machine interface, network interface and in-house software interface. Man-machine interface establishes the interaction between the user and the system, including the user input crisis information and system output preplan information. In most cases, the crisis information mastered by user is limited, and most of the information is fuzzy, uncertain, incomplete, which requires the man-machine interface strong information identifying and processing ability.
2. the crisis warning system. Used for collecting information pretreatment and discriminant warning before enterprise crisis , there are three roles. One is the information collection, including data collection, knowledge acquisition, policy simulation and monitoring analysis and forecasting. Two is information evaluation, inputing information through the artificially or automatic system, warning crisis by judging crisis evaluation index. Three is information preparation , pretreating crisis information to ensure that the decision makers can get briefly, comprehensive information as far as possible.
3. emergency response system. Used in preliminary stage, it can instantly bring storaged emergency information to the decision makers and crisis management personnel, order relevant departments and personnel mobilize imminently, take actions according to advance emergency plan or temporary arrangement. The decision makers should make an emergency decision immediately. Emergency information includes: briefly description of crisis situation, the emergency plan or temporary instructions, contact relations between the relevant decision-making departments and personnel, the resources of an organization and availability, etc. Main function of the system is to provide decision makers the most important information with concise way, in order to avoid human error.
4. inference decision system . Inference decision systems is to analyze the crisis and put forward relevant suggestions by applying the current cases and rules, through two aspects to help decision makers to design an emergency plan: One is to extract similar case from case base and refer the solution; Two is to absorb expert experience from knowledge base, follow the emergency rule to design the emergency plan. For uncertainty of the crisis situation, reasoning mechanism needs to take fuzzy logic or default logic.
5. simulation assessment system . After amergency decision scheme generation, using model to simulate decision-making plan's effect, estimate decision-making

risk, then evaluate different schemes according to pre-decided standards, determine the ultimate solution.

6. organization scheduling system. Calculating labor power, material, equipment and expenses of implementing decisions by using the model and providing the dispatching scheme of labor power, material resources and capital.
7. information base and its management system. The information base is a general term for databse, case base, knowledge base and graphics library. Information management system consist of each library management system, respectively carry out routine operation to realize the consistency of the information base. To improve the efficiency of the system, various information bases can be divided into general information base and special information base.

3.2 Enterprise Emergency Decision Support System's Main Function Design

1. Fundamental information management. Basic information management is mainly that system administrators manage and maintain the basic information, such as, major hazard, emergency resources, geographical information, emergency plan, knowledge base, sensor information, video monitoring information, incident response level classification standard information, statistical reports and emergency production scheduling scheme

2. alarm receipt management . The main function of alarm receipt management is recording and processing accident alarm information. When accident happens, the general process is the reporter or sensors alarms its branch control room, total control room and security department at the same time. On-duty person records alarm information according to the relevant accident information reported by the reporter or sensors. On-duty person decided the response level and whether to receipt alarm according to the influence degree combined with emergency treatment process in the enterprise emergency plan. According to the accident emergency disposal principle that "grading response, apanage primarily", branch control room touched the alarm first when accident happened. Factory group level decided whether to receipt alarm and its way according to accident type and size.

3. Remote video review . Emergency decision-making system uses software interface provided by invoking enterprises Video Surveillance System and Switches the video stream displayed in the client to realize remote video review. When the accident happens, after the authorization by system, we can remotely review video image information of the accident scene. Through this function, Emergency command room according to the big screen image display of total control room ,will spread the video streams needed to the client under the control of the out leadership. Out leadership can loge into emergency rescue information system by the Internet, know the accident situation and command the emergency rescue.

4. emergency auxiliary decision . Emergency auxiliary decision functions realize supporting auxiliary decision of emergency rescue process in big and important accident by reviewing The scene of the accident video, reference case and plan, logic analysis model, statistical calculation model and so on. specific functions include.

5. Emergency communication platform . Emergency communications platform is mainly integrated enterprise fixed telephone, mobile phone, depending on enterprise interfaces provided by switches. It provides a unified emergency communication platform for major accident emergency rescue. After the accident, emergency

command staff can make Mass text messages and notification and real-time voice communication with other personnel through fixed phone or mobile phone. The system can also make automatic call after call failed. Emergency communications platform can hold telephone conference at any time of accident emergency rescue.

6. Log management . System log management function realizes the automatic recording of process of system's emergency rescue, Alarm information records, accident preliminary classification result, text messages, information and reply condition, the scene of the accident positioning information, enterprise emergency production scheduling solution information, the consequences of the accident simulation results, evacuation information and emergency resource scheduling information. Meanwhile, the system provides log maintenance function for the user.

4 Problems Needed Further Discussion and Research

The establishment of disaster model under the complex terrain or obstacles diffusion, fire, explosion and their superposition will need to be researched. Full 3d accident warning and emergency GIS needs to study. The realization of the linkage function between Real-time control system and this system, such as: the real time control of the fire engine's global positioning, safety monitoring system for real-time control of signal collection etc. With the support of this system platform, achieving these real-time control information reaction and handling is very convenient, but needs site personnel's close cooperation. Compared with the emergency response system and reasoning decision-making system of decision support system, The structure of this system is not comprehensive in the enterprise applications. It remains to be further research and the structure of the system remains to be strengthened.

References

[1] Luo, X.: Study on Information Mechanisms of Enterprise Crisis Management. Science Press (January 2009)
[2] Cao, L.: Study on the Preparation of Enterprise Crisis Management. Market Modernization (19) (2007)
[3] Wang, W., Zhan, J., Chen, Y.: A Brief Talk on the Construction of Enterprise Crisis Management. Modern Economic Information 02 (2009)
[4] Zhu, L., Zhu, F.: Establishment of Pre-Warning System Model for Crisis Management in Enterprises. China Soft Science 11 (2004)
[5] Wang, D.: Study on Pre-Warning Enterprise Crisis Management Model. Shangdong Peoples Publishing House
[6] Liu, Y.: Precaution Mechanism and Mode Discussion of Tarim Oil Field's Crisis Management. Southwest Petroleum University (2005)
[7] Hu, Y.: Path Analysis of Constructing Enterprise Crisis Management System. Northern Economy and Trade 07 (2008)

Shallow Parsing of Chinese Based on HMM Model*

Zheng Weifa[1] and Xie Wenliang[2]

[1] Information Office, Guangdong University of Business Studies, Guangzhou 510320, China
[2] Editorial Office, Journal of Guangdong University of Business Studies, Guangzhou 510320, China

Abstract. Complete parsing is difficult to meet the need of precision and recall rate in Chinese. To address this problem, a new model for shallow parsing of Chinese is presented in this paper. We adopt Church theory and carry on Chinese phrases recognition based on HMM; improve the precision rate of sentences separation by improving the observance probabilities of HMM model and making use of the context information of the Chinese sentences. At the same time, by studying the rules of Chinese sentence, we extract some rules useful for ambiguity elimination. The experimental result indicates that the model based on HMM has high precision and recall rate.

Keywords: HMM, Shallow Parsing, Natural Language Processing.

1 Introduction

Shallow parsing (also partial parsing, or chunk parsing) is a kind of language processing policy that emerge recent year in the natural language processing area. Shallow parsing and complete parsing are two different methods used for parsing. Complete parsing needs to analyze the whole structure of a sentence in order to reveal the meaning of the sentence uttermost; while shallow parsing only need to divide the sentence into some smaller elements and recognize some of the relatively simple element of the structure, such as non-recursion noun phrase, verb phrase, et al, as oppose to relation of the whole elements of the sentence. These recognized units usually are called chunk. The result of shallow parsing are not a complete syntax tree, rather it is just a subgraph of a complete syntax tree. If attachment were added to the units, a whole syntax tree can be constructed. So the main task of shallow parsing is sentence elements recognition and parsing, thus we simplified sentence parsing in some extent. There are two major research methods in shallow parsing:

(1) Statistic method, such as Church[2], applied HMM method to convert the sentence border recognition problem into the problem of inserting border between the words label.
(2) Regular method, such as Abney[3], used finite-state cascade to simplified sentence syntax parsing.

* Supported by Foundation for Distinguished Young Talents in Higher Education of Guangdong, China Number:LYM09084.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 79–86.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

HMM (Hidden Markov Model)[4] is a statistics technology developed from audio recognition, it provides technology for auto construction of recognition system based on probability of the training data. In reference [2,136-143], Church used HMM to recognize the non-recursion noun phrase. He transformed the phrase border recognition problem into the problem of inserting left border of NP("[") and right border of NP ("]") into pairs of phrase labels. In this paper, we applied the thinking of Church algorithm by considering the concretely connected structures of Chinese sentence. In computing the observance probability of HMM model, we take context information into consideration; at the sometime, we extract rules sets, limit and optimize ambiguity problem according to the feature structure of Chinese sentence.

2 Thinking and Algorithm of HMM

2.1 Basis Thinking of HMM

HMM model includes two levels: observable layer and hidden layer. The hidden layer is a Markov process, or a finite state machine, which has transfer probability in every state transfer process. In syntax parsing, observable layer is a sequence of feature words while the hidden layer is a sequence of type of phrase. Training data is the text corpus that labeled with the types of the phrase. The trained HMM is an automatic syntax parser which can label the types of a phrase that has already separated from a Chinese sentence. Low level syntax parsing can be describe as follows:

for any given text that has already separated into phrase and labeled with types of part of speech, such as $O = w_1 / v_1\ w_2 / v_2 \ldots w_i / v_i \ldots w_n / v_n$, we want to find out a sequence of phase $S = s_1 s_2 \ldots s_i \ldots s_n$, so that $P(S \mid O)$ gets the max value which demonstrate the best phrase component of O. Among them, O is a sentence to be parsing; w_i is a phrase that has already separated from the sentence, v_i is the type of part of speech, s_i is the type of phrase that can be described as following:

$$v_i \in (a, b, c, d, e, h, i, j, k, l, n, o, p, q, r, t, u, v, w, y),$$

$$s_i \in \{ap, dp, pp, tp, sp, mp, mcp, np, vp\}.$$

$$s_i \in \{ap, dp, pp, tp, sp, mp, mcp, np, vp\}.$$

2.2 Basic Algorithm of HMM

(1) The learning of HMM

The parameter set of HMM model can be denoted by

$$\lambda = (\pi, A, B). \tag{1.1}$$

Through the training, we can get the model λ.

Suppose that we can get enough training simple and compute the numbers of state that start from S_i in a certain sequence, we denote the number as $Init(i)$; at the same time, we can compute the frequency that S_i state shifting to S_j state, denoted by C_{ij}; in S_j state particular, the numbers of the parts of speech set $V_k = v_l v_{l+1}$ are released, denoted by $E_j(k)$. Let N be the number of the states in the model, which means that N is the number of the type of phrase. M is the observant value of the model which means M is the parts of speech strings that makeup the phrase. Different states consist of different parts of speech strings which means M is vary from one another. As a state viable, the appearance probability of phrase type at any moment depends on the previous phrase's type.

$$\pi = \{\pi_i\}$$

$$\pi_i = P(q_1 = s_i) = \frac{Init(i)}{\sum_{j=1}^{N} Init(j)}, 1 \le i, j \le N \tag{1.2}$$

$$A = \{a_{ij}\}$$

$$a_{ij} = P(q_t = s_j | q_{t-1} = s_i) = P(S_j \mid S_i) = \frac{C_{ij}}{\sum_{k=1}^{N} C_{ik}}, 1 \le i, j \le N \tag{1.3}$$

$$B = \{b_j(V_k)\}$$

$$b_j(V_k) = P(v_l v_{l+1} \mid S_j) = \frac{E_j(k)}{\sum_{k=1}^{M} E_j(k)}, 1 \le j \le N, 1 \le k \le M \tag{1.4}$$

(2) Decoding Problem

Given a sequence $O = O_1 O_2 ... O_M$ and a model $\lambda = (\pi, A, B)$, we want to find a state sequence $S = s_1 s_2 ... s_n$ that can best describe sequence O in some means. Suppose at *t* moment, the appearance probability of phrase type depends on the previous phrase's type, we have $P(S_i) = P(S_i \mid S_{i-1})$. According to Maximum Likehood Estimate (MLE) method and Bayes rule, we get the probability of state shifting:

$$P(S_i) = P(S_i \mid S_{i-1}) \quad 1 \le i \le N$$

$$P(S_i \mid v_l v_{l+1}) = \frac{P(S_i) * P(v_l v_{l+1} \mid S_i)}{\sum_{n=1}^{N} (P(S_i) * P(v_l v_{l+1} \mid S_i))}$$

$$= (a_{i-1,i} * b_i(V_k)) / \sum_{i=1}^{N} (a_{i-1,i} * b_i(V_k))$$

$$1 \le i \le N, \quad 1 \le k \le M \tag{1.5}$$

$P(S_i \mid v_l v_{l+1})$ is the possible probability of the parts of speech string $v_l v_{l+1}$ that belongs to S_i.

Let $P(S^* \mid v_l v_{l+1}) = \max(P(S_1 \mid v_l v_{l+1}),..., P(S_i \mid v_l v_{l+1}),..., P(S_n \mid v_l v_{l+1}))$

From the meaning of probability, we can consider that the parts of speech strings $v_l v_{l+1}$ belongs to phrase S^*.

(3) Chunk separation

We have discussed how to decide the phrase state that the adjacency parts of speech strings belong to from the perspective of probability. Yet we haven't decided whether or not that the adjacency parts of speech strings belong to the same chunk. Using Church thinking, we can transform the problem of phrase border recognition into the problem of inserting '[' and ']' into the pairs of parts of speech. If not taken null phrase (such as '[]') and embedded phrase (such as '[[' ']]'), et al., into consideration, there are only four circumstances that a pair of parts of speech label exists: (1) [; (2)]; (3)][; and (4) null(no border). The last circumstance can be further separated into two conditions: (a) no border but within the phrase (I); (b) no border but beyond the phrase (O). Thus any pair of parts of speech label exists only 5 circumstances: (1) [; (2)]; (3)][; (4) I ;(5) O.

Let $P(v_l, v_{l+1})$ be the probability of the phrase that made up by two adjacency parts of speech. $B = b_1 b_2 ... b_i ... b_n$ is a string of information that describe border, we have $b_i \in \{[,],][, I, O\}$. According to formula (1.5), we have:

$$P(v_l, v_{l+1}) = P(S_i \mid v_l v_{l+1}) = \frac{P(S_i) * P(v_l v_{l+1} \mid S_i)}{\sum_{n=1}^{N} (P(S_i) * P(v_l v_{l+1} \mid S_i))}$$

Hence, the phrase border recognition problem convert to the issue of how to insert five possible states, which is (1) [; (2)] ; (3)][; (4) I ; (5) O, into different parts of speech according **to** $P(v_i, v_{i+1})$**.**

Table 1. Border Description Information Table

b_{i-1} / $P(v_{i-1}, v_i)$	[	]	][	I	O
$P(v_{i-1}, v_i) > 0$	b_{i-1} unchanged b_i =I	b_{i-1}=][b_i =I	b_{i-1} unchanged b_i =I	b_{i-1} unchanged b_i =I	b_{i-1}=[b_i =I
$P(v_{i-1}, v_i) = 0$	b_{i-1}=O b_i =O	b_{i-1} unchanged b_i =O	b_{i-1}=] b_i =O $S = S_k$	b_{i-1} unchanged b_i =] $S = S_k$	b_{i-1} unchanged b_i =O

2.3 Improving HMM Model by Taking Advantage of the Relevance Context

However, in model describe above, we have not taken the characteristic of the close relation between the contexts into consideration. Thus, we improve the observation value probability computing:

$$\begin{aligned} b_i(V_k) &= P'(v_l v_{l+1} \mid S_i) \\ &= P(v_{l-1} \mid (v_l v_{l+1}, S_i)) * P(v_l v_{l+1} \mid S_i) * P(v_{l+2} \mid (v_l v_{l+1}, S_i)) \end{aligned} \tag{1.6}$$

In formula 1.6, $P(v_l v_{l+1} \mid S_i)$ is the output probability of the parts of speech strings $v_l v_{l+1}$ that shifting from the current state to state S_i; $P(v_{l-1} \mid (v_l v_{l+1}, S_i))$ is the output probability of previous context that the parts of speech strings $v_l v_{l+1}$ is in the state of S_i; $P(v_{l+2} \mid (v_l v_{l+1}, S_i))$ is the output probability of follow context that parts of speech strings $v_l v_{l+1}$ is in the state of S_i.

3 Ambiguity Elimination

Ambiguity is a common phenomenon that long exists in natural language and it is also an issue that has to face in natural language processing. In this article, if two adjacency parts of speech phrase sequence labeled as $w_i / v_i, w_{i+1} / v_{i+1}, ..., w_j / v_j, (j \geq i+2)$

can makes up a new phrase, then we believe that there exists ambiguity in $w_i / v_i, w_{i+1} / v_{i+1}, ..., w_j / v_j$.

By doing research on Chinese sentence rules, we extract some rule that are useful for ambiguity elimination. Aiming at the special structure word in the Chinese sentence, such as "di" and "de", we construct a basic rules set G. We make some adjustment in the phrase probability that made up by two adjacency term. The original probability multiple k means that the probability of the two adjacency parts of speech making up a phrase is strengthening. The original probability multiple 1/k means that the probability of the two adjacency parts of speech making up a phrase is weakening. By doing experimental test, we find that $k = 10$ has the great discrimination. Part of the rules is shown in Table 2, and the wildcard * means one or more part of speech.

Table 2. Rules Set

Conditions	Separate Result	Example
ap u np, and "u" is the part of speech label of "的"	ap(ap u) np	ap(beautiful/a de/u)bag/n

4 Experimental Results

4.1 Experiment Corpus

Corpus choosing: we use the Corpus of machine translation trees developed by Chinese Academy of Social Sciences.

Testing corpus: we use part of the machine translation tree developed by Chinese Academy of Social Sciences. It has about 1000 sentences, approximate 10,000 words.

Training corpus: we use part of the machine translation tree developed by Chinese Academy of Social Sciences for HMM training. It has about 2,082 sentences, and the average sentence has 8 words.

4.2 Performance Evaluation Indexes (Close Door Testing)

(1) the Precision Rate of phrase recognition=(the number of phrases recognized correctly by HMM/ the number of phrases recognized by HMM)*100%.
The computing result of the Precision Rate is: 8832/10389*100%=85.01%.

(2) the Recall Rate of phrase recognition=(the number of phrases recognized correctly by HMM/ the total number of phrases in the testing corpus)*100%.
The computing result of the Recall Rate is: 8832/10408*100%=84.86%.

(3) Precision Rate and Recall Rate of different kinds of phrase.

Table 3. Testing Results

Testing Program	ap	dp	pp	tp	sp	mp	mcp	np	vp	Sum
The number of phrases recognized correctly by HMM	1012	777	420	282	260	562	90	2728	2701	8832
The number of phrases recognized by HMM	1124	829	441	301	280	586	95	3392	3341	10389
the total number of phrases in the testing corpus	1103	831	454	300	287	597	101	3353	3382	10408
Precision Rate	90.04 %	93.73 %	95.24 %	93.69 %	92.86 %	95.90 %	94.74 %	80.42 %	80.84 %	85.01 %
Recall Rate	91.75 %	93.50 %	92.51 %	94.00 %	90.59 %	94.14 %	89.11 %	81.36 %	79.86 %	84.86 %

4.3 Error Cases and its Analysis

The reason causing the error cases are:

(1) the same parts of speech sequences may have different combinations and the probability of combinations are also different. In our HMM model, we use Maximum Likehood Estimate(MLE) method to do the chunk separation, so in the cases of low probability, HMM can not perform the chunk separation until further information support are provided.

(2) the HMM have only modern Chinese syntax property as oppose to semantics information, so it can often cause error when analyzing sentences that need semantics support.

Table 4. Error cases and its result analysis

Types of error cases	Error cases	Correct analysis
Semantics needed	vp(vp(vp(v<chuan>)u<zhe>)np(ap(np(n<underdress>)u<de>)np(n<children >)))	np(ap(vp(vp(vp(v<chuan>)u<zhe>)np(n<underdress>))u<de>)np(n<children>))

5 Conclusion

In this paper, we develop a shallow parsing model based on HMM. The model can realize phrases separation and type recognition; in the recognition process, the model can eliminate ambiguity by using Chinese sentence rules; finally, we get the model performance indexes by using syntax analyzer to do close-door and open testing. Experimental result show that the model has relatively high precision rate and recall rate.

References

[1] Sun, H., Yu, S.: An overview of the shallow parsing. The Contemporary Linguistics (2000)
[2] Church, K.: A stochastic parts program and noun phrase parser for unrestricted text. In: Proceedings of the Second Conference on Applied Natural Language Processing, pp. 136–143 (1988)
[3] Abney, S.: Parsing by chunks. In: Berwick, R., Abney, S., Tenny, C. (eds.) Principle-Based Parsing. Kluwer Academic Publishers, Dordrecht (1991)
[4] Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. In: Morgan Kaufmann, Waibel, Lee (eds.) Readings in Speech Recognition, 1990, pp. 267–296 (1989)
[5] Su, T., Wu, J.: The improved HMM model based on the spatial correlation. The Computer Engineering And Design, 1023–1026 (2010)

Application of Seismic Sedimentology Method to Beach-Bar Depositional System

Guo Jianqing and Lin Chengyan

Department of Earth Resource and Information, Dong ying College of University of Petroleum, China

Abstract. Seismic Sedimentology is a new and developing boundary subject for study of lithology stratigraphy., petrology, geomorphology and Sedimentary. Which aid is for research the sedimentary structure and history at the equal time frame of three-dimensional sequence stratigraphy with seismic information and technology. It refers three key technologies：90º phase inversion, frequency interpretation and strata slice. In this paper, these methods of seismic sedimentology are applied to beach-bar sediment system identification of WANGJIAGANG oil-field in NIUZHUANG depression of BOHAIWAN basin and getting significant effect. Drawing and conclusion that shallow lake sedimentary clearly since Shahejie four member upper of Paleogene, and with the water deeper ,the main sedimentary face changes beach sand or bar sand microfacies to mud and shallow lake mud rich.

Keywords: Seismic Wavelet, Seismic Sedimentology, 90º Phase Inversion, Frequency Interpretation, Strata Slicing.

1 Introduction

The main part of research area WANGJIAGANG Oilfield is located in the east of NIUZHUANG depression of DONGYING sag in BOHAIWAN basin(Fig.1). The structure in this area is formed in early Eocene. Undergo three periods: chasmic stage、fault stage、depression stage .It is relatively stable in depression period. chasmic stage appear in Yanshan Movement period , Jurassic period - Cretaceous Period. experience from negative reverse to positive reverse. With Yanshan tectonic movement intensify , forming some base faults :Shicun fault, Gaoqing fault, Chennan fault at DONGYING sag. At same time, two sedimental centers appeared, one is located at BOXING sag in west of DONGYING depression, another is located at NIUZHUANG sag in north of DONGYING depression.

2. Research Methods

2.1 90º Phase TransformWireless Network Overview

We can use seismic modeling to discuss seismic phase control on the interpretability of composite waveforms for seismically thin beds. In the frequency domain, a wavelet

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 87–92.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

can be uniquely expressed by amplitude and phase spectra. the phase spectrum displays phase lags of respective signals in the amplitude spectrum(Zeng 2003). Assuming the same amplitude spectrum, a linear-phase change leads to a time shift of a wavelet; a constant-phase shift alters wavelet shape. By modifying the slope of the linearphase shift and constant-phase shift, an infinite number of wavelets can be created from the same amplitude spectrumexamines the constant phase shift of a Ricker wavelet obtained by applying the Hilbert transform.

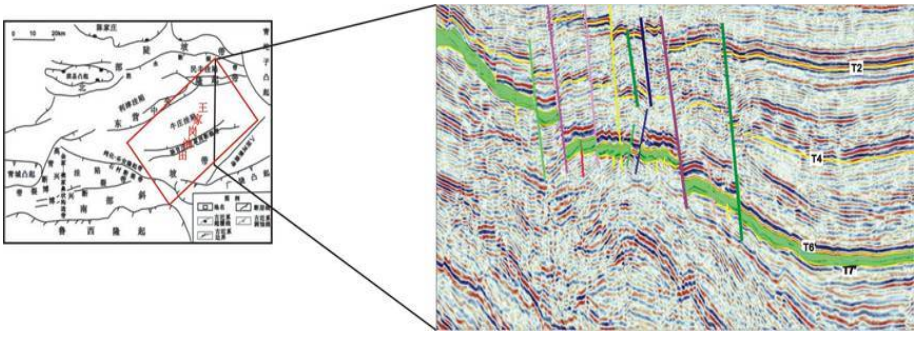


Fig. 1. The station of WANGJIAGANG oilfield and the seimic seciton L1100(green circle represent objective layer)

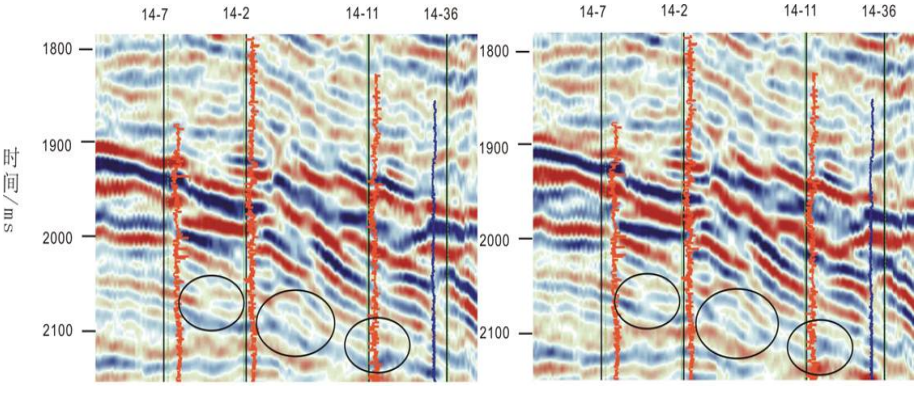


Fig. 2. Wellcross section of upper shahejie four member sediments in wangjigang field. (a) Standard zerophase seismic section. (b) the seismic section after The 90◦-phase inversion (black circle representative change).

Typically sandstone have a low-AI and shale profile have high-AI. Well-log data show AI increases with shaliness low acoustic impedance (AI) sandstones relative to shale or high-porosity zones relative to low-porosity(Zeng, 2009); In a well-log seismic section we use seismic amplitude for geometric and lithologic interpretation, we may choose to deal with dual polarities (trough and peak) at the same time. in the 90◦-phase section, the seismic response to a thin bed is a singlemain trough(Zeng and Jones,2003). The center of the bed lines up with the maximum characteristics, such as

minimum phase, mixed phase, and maximum phase, are caused by nonlinear phase distributions and are not discussed here in detail. A zero-phase wavelet can be converted to a minimumphase or maximum-phase wavelet having the same amplitude spectrum by applying spectral factorization. The interpretive quality of a wavelet is also related to therelative size of its side lobes.

When appling 90°phase rotation to this seismic data. Then, it can be found that the main lobe of strong amplitude peak perfectly coincides with sandstone detected by GR. Sandstone bodies do not have a simple relationship with seismic events. The upper half of a sandstone typically corresponds to a seismic trough (red), whereas the lower half of the sandstone is normally tied to a seismic peak (black), with the midpoint of the sandstone layer roughly at the zero crossing. As a result, seismic polarity and amplitude are neither good indicators of impedance and lithology nor reliable references of lithofacies location and geometry.

2.2 Frequency Interpretation Geological Siginificance

In general, according seimic sedimentology concept, seismic event does not representative geological isochronous interface ,but also not represent lithology boundary(Dong and Zhang, 2006), it is the two respect combination reflection. Different frequency of seismic data, reflection geological information is different. It is dependes on the frequency. So the high frequency seismic event in seismic data reflect equitime sediment boundary(Zhang et al., 2009), low frequency seimic event reflect lithology boundary. And anomalous instantaneous frequency spikes relate to certain bed-thickness

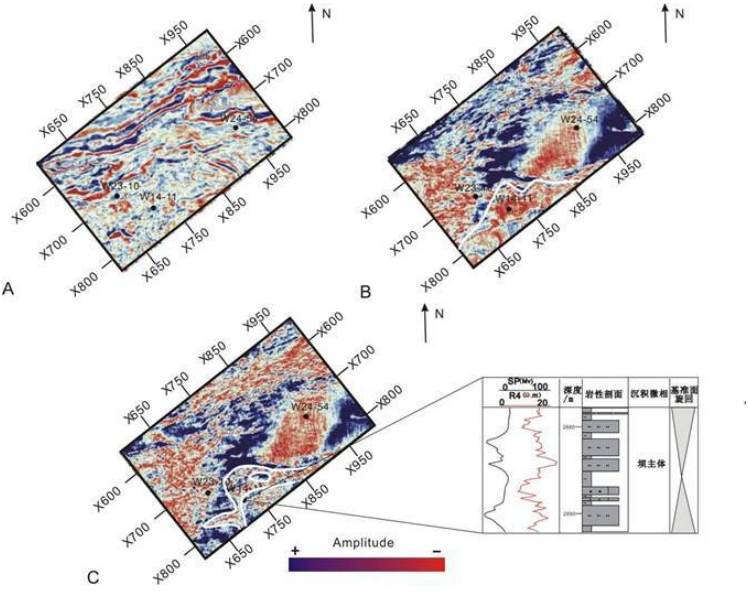


Fig. 3. Comparing the different seismic slicing by using different frequency (A- initial seismic slicing figure;B-20Hz;C-30Hz)

ranges(Zhang and Lin,2010). frequency spikes can be a useful indicator of seismically thin beds that are displayed as amplitude minimal in seismic data, and are otherwise characteristically tricky to recognize(Zhang et al., 2009). These thin beds light be ultrathin reservoirs in a prolific petroleum basin So it has important geological significance. We can use frequency interpretation to proceed seismic data processing(fig.3). and with the frequency increased, the thin bed information can be matched well with litherlity and logging et al(Fig.3). And acording different request to adopt different frequency seismic data volume. At the same time we must pay attention to the anomalous instantaneous frequency spikes, it is possible suggest strategical geological significance.

3 Research on Strata Slice

Slice of 3D seismic data is widely used in seismic interpretation(Zeng and Li,2008), including study of structure, fracture planar distribution and lateral variation of strata and lithology. Traditional seismic slice includes time slice and horizon slice. However,

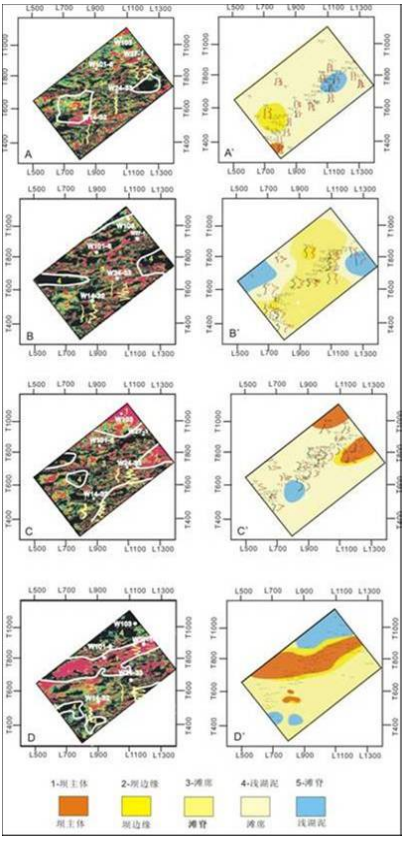


Fig. 4. Comparing the typical stratigraphy strata slicing with the sedimentary microfacies (left-strata slicing figure;right- sedimentary microfacies figure), from D-A means from old to new in the strata

under instruction of seismic sedimentology theory, a new slice technology is proposed(Dong and Zeng, 2008), which is strata slice. This kind of slice is obtained from a series of interfaces proportionally interpolated between two interpreted isochronal interfaces.Horizon slice is seismic amplitude values extracted along interpreted horizon or parallel up or down shifted one(Lin and Zhang ,2008). This method is effective when the upper and lower strata are relative conformable contact. But In most conditions, the interpreted stratum is not parallel to up and down strata. Because the strata thickness is unequal, this kind of slice will result in unique time or diachroneity sediment boundaries identify phenomenon, and its amplitude just reflects lithology boundary and distribution in local range rather than Wavelet analysis shows that the phase of seismic data. and time slicing means base on the precondition that all the strata is horizon including the upper and lower strata ,and obviously it is impossible in real geology condition(Zeng and Loucks, 2007), and will result diachronous phenomenon seriously in indentify stratum. So only stratal slicing is identify stratum according to the real strata sediment condition. Basically can be thought that the boundaries is equitime contact surface.

4 Conclusion

According to the high development of seismic physics technology, beach-bar sedimental systems Strata slice is obtained along the interface that is interpolated between the strata of shahejie four member. according to practical requirement (Fig.4). For different studied areas or different grade sequence boundary surface, it is not definitive how many interpolated horizons needed, which confines flexibility of use of slice browsing tool of available interpretation software. getting data volume to construct isochronous strata is required when take this programe, it must to do first which is flattening aim horizons. After this program, the flattened top and bottom horizons and proportionally interpolated data can be input into Landmark seismic Poststack system, browsing it as time

Acknowledgement. The work was supported by the Project of National Natural Science Foundation(40872094).that is the Application of Seismic Sedimentology method on Fluvial and Delta facies.

References

[1] Zeng, H.L.: How thin is a thin bed? An Alternative Perspective: The Leading Edge 28, 1192–1197 (2009)

[2] Zeng, H.L., Jones, R.: Reconditioning seismic data for improved reservoir characterization, lower Clear Fork and Wichita, Fullerton field, West Texas. In: Hunt, T.J., Lufholm, P.H. (eds.) The Permian Basin: Back to Basics: West Texas Geological, pp. 03–112, 67–78 (2003)

[3] Zeng, H.L.: Significance of seismic phase in interpretation of stratigraphy and sedimentology: Convention. Canadian Society of Petroleum Geologists/Canadian Society of Exploration Geophysicists, Abstract (2003), 266S0131

[4] Dong, C.M., Zhang, X.G., Lin, C.Y.: Concept, method and technology of seismic sedimentology. Acta Sedimentologica Sinica 24(5), 698–703 (2006)
[5] Zhang, Y.N., Zhu, X.M., Liu, C.L.: Seismic sedimentology and its appl ication in south of Central Asia. Petroleum Exploration and Development 36(1), 74279 (2009)
[6] Zhang, X.G., Lin, C.Y.: Seismic Sedimentology and Its Application in Shallow Sea Area, Gentle Slope Belt of Chengning Uplift. Journal of Earth Science 21(4), 471–478 (2010)
[7] Zhang, X.G., Lin, C.Y., Dong, C.M.: From Amplitude to Architecture of Braided Channel Bars: The Sedimentological Interpretation of Stratal Slices in a Shallow-Sea Seismic Survey, Dagang. In: CPS/SEG 2009 Conference, Beijing (2009) ID1150
[8] Zeng, H.L., Loucks, R.G.: Linear Amplitude Patterns in Corpus Christi Bay Frio Subbasin, South Texas: Interpretive Pitfalls or Depositional Features? Geophysics 73(5), A27–A31 (2008)
[9] Li, X., Zeng, H.: Seismic Sedimentology and Depositional Systems in the Upper Cretaceous Olmos Formation, Gold River North Field, Webb County, South Texas. Gulf Coast Association of Geological Societies Transactions 58, 623–634 (2008)
[10] Dong, Y.L., Zeng, H.L.: Study of Seismic Sedimentology in Qinan Sag. Journal of China University of Petroleum (Edition of Nature Science) 32(4), 7–12 (2008) (in Chinese with English Abstract)
[11] Lin, C.Y., Zhang, X.G.: The Concept and Application of Seismic Sedimentology. In: Proceedings Innovative Development Problems in Oil & Gas Industry International Scientific and Practical Conference, Almaty, pp. 62–66 (2008)
[12] Zeng, H.L., Loucks, R.G., Brown, L.F.: Mapping sediment2dispersal patterns and associated systems tracts in fourth and fifth order sequences using seismic sedimentology: Example from Corpus Christi Bay, Texas. AAPG Bulletin 91(7), 98121003 (2007)

Visual Sign System Design of Modern Office Building

Shen Lei

Liaoning University of Technology, China

Abstract. Office buildings are among the most important parts of modern all sorts of organizations. No matter governments, companies, NGO and so on. Because they are the hearts of those organizations, important decisions of the organizations are all come from those office buildings. For different sort of office building, there are different kind of styles, and purposes. However, now days different organizations always located together, for they can save lots of space as well as they can promote mutual communication at the same time. That is to say different office buildings assembled together in one building. But as we mentioned just, different building offices have different features, which means we must find a way to tell the features apart, there must be some kind of visual signs not only to distinguish these organizations, but also to highlight the features of the organizations, so in this paper we try to do some research on "Visual sign design of Modern office building". The visual signs must be in keeping with the features of related organizations. Visual signs have lots of functions, such as the mark of organization images, sales and so on. In this paper we try to do some research about the designing of the signs according to their different functions.

Keywords: Visual, Sign design, Modern office building.

1 Introduction

Office buildings are the basic places of modern activity. Case almost all the important decisions of some organizations are come from their office buildings. Visual signs are very important to office buildings, normally they are used as guides, but at the same time they are on behalf of their organizations' culture. That is to say the styles of signs must agree with their organizations.

However, now there used to be a lot of organizations locate in one building, which means several organization, governments, different companies, NGO, even schools share one building, as they may save space. At this time, their signs must appear in one building, so we must find a way to arrange these signs, not only to act as guides, but also to feature each organization's culture. So in this paper we make some effort to do this. The structure of the paper is as follows: the background and development situation of visual sign design are arranged in the second part, in the third part we do some research about Visual sign design, we discuss visual signs with different styles. Finally a conclusion is given at last. Fig. 1 is a sample of visual signs of some office building

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 93–99.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Fig. 1. Visual signs of some office building

2 Background and Development Situation of Visual Sign Design

Visual signs have lots of functions, such as the mark of organization images, sales and so on.

2.1 The Concept of Visual Sign

From the narrow sense, the concept of signs system design that Call be divided in two respects: First the symbols using to indicate the direction and area; second is the form of expression of symbol carriers in the space. In the broad sense, visual symbols and expressions can be seen as signs system design, which are used to convey the Concept of space. As the rapidly development of modem society, public space, traffic environment and the building environment, function and structure has become increasingly complex, signs system plays a very important role that works in the transportation, stream of people, divide in the area, So the research of it has positive practical significance.

2.2 Classification of Visual Sign

Different organizations have different purposes and culture, so signs vary with organizations. Generally signs can be sorted as office sign, commonweal signs, sales signs and requisite signs and so on. All of these signs are parts of Visual Sign systems, together to realize the system's value. And all these signs will be discussed in this paper. Fig. 2 shows a simple Classification of visual sign.

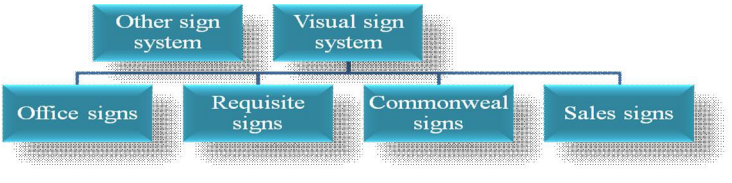


Fig. 2. Simple Classification of visual sign

2.3 The Development of Visual Sign

Visual Sign firstly appeared in europe early at 1940s, and then blossomed both in europe and the USA. It is the requirement of Western industrialized. It is firstly as a guide in freeways. Lead people to resting areas, gas stations, and the directions to the journey's ends. And then with the development of economy, the structures and functions of office buildings have been becoming increasingly diversified. So they need more signs to guide people in some big building, latterly people find that signs also have the functions of representing organizations' cultures and styles, and finally visual sign design become a discipline. Recently more and more attention focused on visual signs in china. And l lot of scholars began to study the related subject. Such as a professor of Central Academy of Art named Yong Xiao begin to setting up such a course in 1997. Also some experts of Chen du suggest there would be a redesign of visual signs in Jin niu District (Chen du) in 2010. But we still have much room to improve. Compared with western countries, we still have much to do. Firstly we still have a shortage of signs from the point of number. Secondly there is a chaos within our sign system. Thirdly English explanations are illegal in signs.However, visual signs are playing an increasingly important role now, lots of research and discussion needed to be done to improve the situation [1].

3 Design of Visual Sign in Modern Office Building

Generally signs can be sorted as office sign, commonweal sign, sales sign and requisite sign and so on. All of these signs are parts of Visual Sign systems, together to realize the system's value.

3.1 Office Visual Signs

Office visual signs are the basic composition of an office. Generally, they need to have some common characteristics, just as seriousness, earnest, conciseness, distinction and so on. They are designed mainly for managers and offices. Basically, there are several rules to design the signs, the signs can be agree with the culture of the organization, or the idea of an organization, and even history, values of the organization. The form of the signs includes signs of departments, bureaus, offices, and breast nameplates. Table 1shows some of famous buildings each with lots of organizations [2].

3.2 Requisite Visual Signs

Requisite visual sign is required by governments in some places, it indicate the requisites of a building or place, such as Emergency exit, Fire-fighting equipment, security identification, traffic system, electricity, water, cable, gas and so on. So the shapes, colors, locations of the sign must be designed by governments. The most important point of this sign is "noticeable". We must make sure that ever one is easy to distinguish the sign, and follow the means. Some researchers have shown that sign color plays a vital role in its legibility. A researcher named Cooper found yellow to be

Table 1. Some of famous buildings each with lots of organizations

Name	Location	Numbers of different organizations	Annotation
JOY CITY	Beijing	More than 55	Shopping mall
Zhongyouer	Beijing	More than 50	Shopping mall
Landmark	Shanghai	More than 70	Shopping mall
SEG Plaza	Shenzhen	More than 100	Digital product mall
Suzhou Innovation Park	Suzhou	More than 1200	Business Incubator
Zhognguancun Business incubato	Beijing	More than 450	Business Incubato
Tuspark	Beijign	Almost 100	Business Incubato

the most effective color with the background of black, light gray. Furthermore, black-on-orange and white-on-green signs have been shown to be detectable at greater distances than black-on-white signs. Jenkins and Cole (1986) found black-on-white signs to provide particularly poor conspicuity. Zwalen and Yu (1991) furthered the understanding of the role color plays in sign detection when they reported their findings that sign color recognition distance was twice that of shape recognition and that the combination of a highly saturated color and specific shape of a sign could double a sign's average recognition distance. The present study followed these principals and the same sign colors were provided within each location. However, to avoid subjects learning what color signs to look for, different color combinations were used across study locations.Specific guidelines ideas: first-level guide sign is region guide sign which are in the entrance of important sections, it can display the current layout of local road map; second-level guide sign is intersection, namely the establishment of guide sign can indicate the direction which are on the general intersection; the third-level guide sign is confirming the destination guide sign in the end [3].

3.3 Commonweal Visual Signs

This kind of sign must be agreed with the culture of environments. One of the main functions of Commonweal visual signs is to make advertisements of commonweal, such as the billboard of environment, water saving, animals protection and so on, such as a billboard said "Accurately measure flow and protects the water supply from possible contamination". Other Commonweal visual

Signs focus on safety cautions, such as a safety cautions "mind gape" in some subway. The building trip distribution, layout of road network and development and utilization of architectural layout, the demand for such guiding information are the basis of the information classification system. With reasonable guide, the elements of information system include office area, recreation area and so on [3].

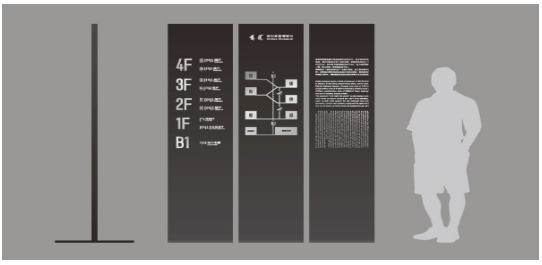


Fig. 3. Characteristics of commonweal signs, office sign, and requisite signs

3.4 Sales Visual Signs

Sales visual signs must make contribution to sales. This kind of sign is designed for company, or organizations with profits. Such as restaurants, barbershops, dress shop, hotels, banks and so on. The most important element of the sign is "sales of brand", we don't need the sign to be agree with the surroundings, but to steal passages spotlights, give them shape memories. Because it is some kind of advertisement, we need to feature the organizations' LOGO [4]. It had been shown by some work, that perpendicular signs are more detectible and legible than are parallel signs, and increasing the size of parallel signs does not improve their visibility. In a survey 30% of the perpendicular signs versus 60% of the parallel signs were missed. But the results were not as extreme. Another survey shows the perpendicular signs were almost never missed while the subjects drove past 30% of the parallel signs, even though the parallel signs were two to three times larger. When the subjects were able to find the signs, the distance at which they could be read was substantially greater for those that were mounted perpendicular to oncoming traffic. The downtown perpendicular signs were read twice as far away as the parallel signs and on the strip development roadways perpendicular mounting increased detection/legibility distance by about 400% over parallel [5].

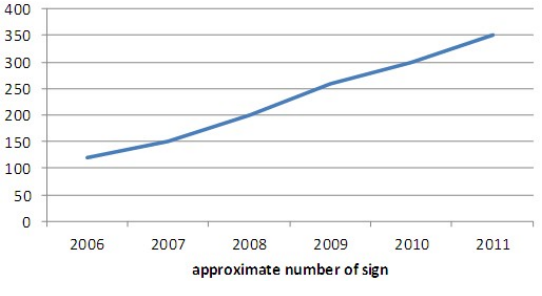


Fig. 4. Approximate number of sign in some business incubator (Beijing)

Signs in the strip development areas had much greater visibility distances and a much higher rate of detection than did the downtown signs. Signs were larger in the strip development condition, and in town there was more congestion, pedestrian

traffic, and visual complexity. However, even between te downtown sites, large location effects were found. While the environmental conditions for the strip locations were relatively uniform, this was not the case in town. Fig. 3 shows characteristics of commonweal signs, office sign, and requisite signs. Fig. 4 shows the approximate number growth trend of sign in some business incubator (Beijing). Fig. 5 is the number of office sign, commonweal sign, sale signs and requisite sign in an office building of a university in Beijing [6].

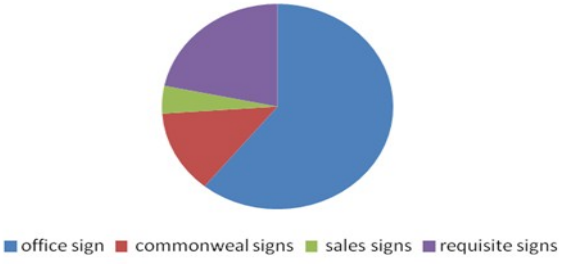


Fig. 5. The number of office sign, commonweal sign, sales sign and requisite sign in an office building of a university in Beijing

4 Conclusion

With the development of business and economic, more and more organizations share one building is an important tendency. Each organization have different property, culture, history, idea and office, so they all need their own signs to represent them, to show their culture and to lead people to their office. Finally, there will be a lot of visual signs; all of the visual signs formed the so called visual sign system. There are several sorts of visual signs with different styles, purposes, culture. In this paper, we made some research about design of visual sign. Generally speaking all the signs can be sorted as office sign, commonweal signs, sales signs and requisite signs and so on. All of these signs are parts of Visual Sign systems, together to realize the system's value. For each kind of sign we all have some discussion about the keystone to design the sign. Finally there still lots of research need to do to further study the subject.

References

[1] Chrysler, S., Stackhouse, S., Tranchida, D., Arthur, E.: Improving street name sign legibility for older drivers. In: Proc. Human Factors and Ergonomics Society 45th Annual Meeting, pp. 1597–1601 (2001)

[2] Forbes, T.W., Fry, J.P., Joyce, R.P., Pain, R.F.: Letter and sign contrast, brightness and size effects on visibility. Highw. Res. Rec. 216, 48–54 (1968)

[3] Wu, R.: Traffic sign's legibility research based on driver's eye movement characteristics. Southeast University Dissertation for a Master's Degree, Nanjing (2008)

[4] Carpenter, R., Fisher, D., Duffy, S., Dutta, A., Upchurch, J., Noyce, D.A.: Overcoming sign design challenges in the central artery tunnel project. In: Proc., ITE 2001 Annual Meeting, Chicago (2001)
[5] Deng, X.D.: Research on selection method of the urban road directional sign information. Journal of Transport Science and Engineering 26(1), 97–102 (2010); GB 5768-2009, Road Traffic Signs and Markings
[6] Mou, Z.-H., Hu, G.M., Wei, D.: Refers to the road sign information graduation system based on the urban road function. Road Traffic and Security 7(1), 15–19 (2007)

Research on Image Intensity Based on Matlab

Wang Yueping[1] and Yi Zhi'an[2]

[1] School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China
[2] School of Sofware, Northeast Petroleum University, Daqing 163318, China

Abstract. In recent years, the application field of image processing is becoming more and more extensive. As a key technology, image intensity is used to improve image visual effect and the clarity of the image. At present, lots of algorithms about this technology have been invention, such as histogram processing, Median filtering, wavelet decomposition and so on, but different algorithms have different advantages and disadvantages. Therefore, different methods should be created in order to realize the technology about image intensity. In this dissertation, the analysis on these advantages and disadvantages will be taken. Take a full analysis on each algorithm is beneficial way to finding a new method for image intensity. Finally, a method combined with many algorithms will be invented, simulated by Matlab and this method will help to improve the effects of the image intensity. Maybe it will also improve the application scope for image intensity.

Keywords: Image Intensity, Histogram Processing, Median filtering.

1 Introduction

Image processing is called computer image processing, it is pointed to transform the simulation image signal to the discrete digital signal using computer, the original image is used for processed, and the output is the improve image or some characteristics[1]. It can improve the practicality of images, in order to require the expected results. In recent years, the application field of image processing is becoming more and more extensive. Image processing has penetrated into civil industry, health care, environmental protection, engineering technology, and other scientific fields, and played more and more important role.

Matlab is an software for numerical calculation, invented by math works. More and more functions and procedures of different interdisciplinary fields are received by this company, therefore matlab can meet some design requirements. Image processing toolbox is included by matlab, it can be applied to digital image processing, such as image restoration and enhanced, image analysis and statistics, two-dimensional transformation, two-dimensional filter design, filter output and so on. In this dissertation, matlab is the main tool for the research and study of image intensity.

Image intensity, a technology about image processing, is used for improving image visual effect and the clarity of the images, emphasizing the characteristic of whole and

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 101–107.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

local, expanding the differences of the characteristics of different objects, in order to output a image that is more suitable for the application occasions[2]. Image intensity is not a nondestructive processing. It is not to increase information for original image, but only to emphasize the useful information, and compress the useless information. It used for the specific problem, and there is no a standard that is made for measuring the quality of image intensity. Fundamentally, A method on image intensity is not fit for all the images. the effect of image intensity is not only based on the algorithm, but also based on the features from image that you want to be enhanced.

There are two realizations for technology of image intensity, and it includes frequency domain analysis and space-time analysis. In space-time Transform you can do operations directly on the image gray level, and it can be classified as point algorithm and neighborhood enhancement algorithm. Point algorithm includes gray level calibration, gray Transform and histogram revisions. The purpose of this technology is making the image imaging even, expanding image dynamic range and improving the contrast. Neighborhood enhancement algorithm is classified as image smooth and sharpen algorithm. Smooth generally is used to eliminate noise, but it can be easy to make the edge fuzzy, and sharpen aims to highlight the edge of the object contour and make target identified easy. Frequency domain Transform is a indirect processing that used to make a analysis on images. Fourier Transform will be used to make the analysis to the frequency domain, and some algorithms including low-pass filter and high-pass filter can be applied. Finally, the inverse Transform on fourier will be used to change the analysis of frequency domain to space-time domain, some other algorithms are also used to improve the accuracy of the results.

In this dissertation, series of algorithm will be used to make analysis on the image with noise, and some important algorithms will be gave a detailed description. All the implementation of algorithms are depended on Matlab.

2 Method Description

In this section, some important algorithms are descried, and these algorithms will be used to make analysis on a image to improve its effects. A digital picture that contains salt & pepper noise is object for study. It is shown in fig.2, and the original image is shown in fig.1.

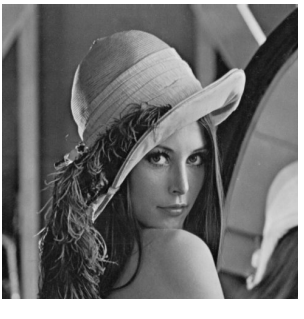


Fig. 1. Original image

Fig. 2. Image with noise

As fig.2 show, the second picture is the picture shown in fig.1 added by salt & pepper noise, that can been realized by using "imnoise" in Matlab.

Now, fig.2 will get a series algorithm Transform, and then we will compare the results with the original image, we can comment the image results. As a image with noise, some filters will be used in order to filter out the noise. Here Median filtering is a good choose.

As Median filtering, pixel desired and pixel value should be sort, then median value will be decided, that value will return to pixel[3]. For example, a image represented by *f(x, y)*, contains a point *(m, n)*, take all the pixel grayscale value in the window recorded as w, order according to side, and take the middle of the value as the grayscale value of the center of the pixels in w. That can be shown in eq.1.

$$g(x, y) = med\{f(m - i, n - j), (i, j \in w)\} \tag{1}$$

If the number of pixels in image *f(x, y)* is odd, the median value is belonged to pixel of point *(x, y)* in the modified image *g(x, y)*, but if the number of pixels in image *f(x, y)* is even, the average of the middle two pixels value is belonged to pixel of point *(x, y)* in the modified image *g(x, y)*. For a Median filtering, to choose a window with right shapes and size is very important, because different shapes and sizes will bring into different effect. The window can be chosen by noise and the details of the image. Now, Median filtering window with linear, phillips, X shape, square, diamond and round is used widely. Function named "medfilt2" in Matlab can realize the Median filtering. The salt & pepper noise can be filtered by this Median filtering, that can shown in fig.3.As it show, the noise in fig.2 is filtered, but is not filtered cleanly, therefore, that also need other algorithms.

Wavelet denoising is one of algorithms for image intensity, that can make the image smoothly[4]. After wavelet Transform, do some processing to remove the noise in image, in order to improve the quality of the image. Firstly, wavelet Transform for image, get the quantities with different size, position and direction. Secondly, change the coefficient, according to different position and direction of some quantities. Make some interested component are magnified and some the component that aren't interested reduced. Thirdly, make inverse Transform to enhance the image.

Wavelet denoising is a problem of filter about signal. It can be treated as a low-pass filter, but its advantage is that some image characteristics are kept[5]. For example, all of the signal can be shown in the eq.2. *s(k)* is pure signal, and *n(k)* is additive random noise.

$$f(k) = s(k) + n(k) \tag{2}$$

But in wavelet Transform the eq.2 can be show as eq.3, $\theta(k)$ is pure signal, and $z(k)$ is the coefficient of noise.

$$\omega(k) = \theta(k) + z(k) \tag{3}$$

In wavelet Transform, a threshold should be set, the wavelet coefficient greater than the threshold is produced by signal, the others is produced by noise[6]. Removing the coefficient produced by noise can achieve the purpose of noise. It is usually three steps for wavelet transform. Firstly, calculate the orthogonal wavelet transform for the signal with noise, choose a suitable wavelet and waeltet decomposition level *j*, move the signal to j, and get the coefficients of the wavelet. Secondly, make the threshold processing to wavelet coefficients, and get the accurate valuations for signal. Thirdly, make inverter wavelet transform to the signal, reconstruct the threshold of wavelet coefficients, and restore the signal, finally get a high quality signal. Wavelet Transform can be realized by function "wavedec2" in Matlab[7]. After wavelet Transform, the fig.3 is shown in fig.4.

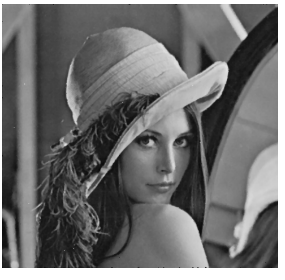


Fig. 3. Image with Median filtering

Fig. 4. Image with wavelet Transform

Wavelet Transform can clean the noise completely, and make the characteristics of original signal peak point retained well, but Gibbs is easy appeared, get the vision of the unnatural signal. Histogram processing is a good choose to make the image more natural.

Digital image histogram is an statistical probability distribution about the grayscales contained in image, such as fig.5.It provides the overview of the grey distribution about image. Histogram enhancement technology is a technology that changes image histogram, change the histogram in order to get a image of high quality. It contains histogram equalization and histogram specification.

The basic idea of the histogram equalization is to make the histogram about original image into histogram with a uniform distribution, in order to improve the contrast degree of image. Histogram equalization technology can only produce a kind of approximate uniform histogram, it doesn't meet all the requirement of image processing. In fact, in order to enhance the gray level range of image, the output of the image histogram is artificial design, and it difficult to describe by mathematical model. At this time, histogram specification is a good choose. It can change the histogram of original image into the histogram with specified shape, and it can be considered an improvement of histogram equalization.

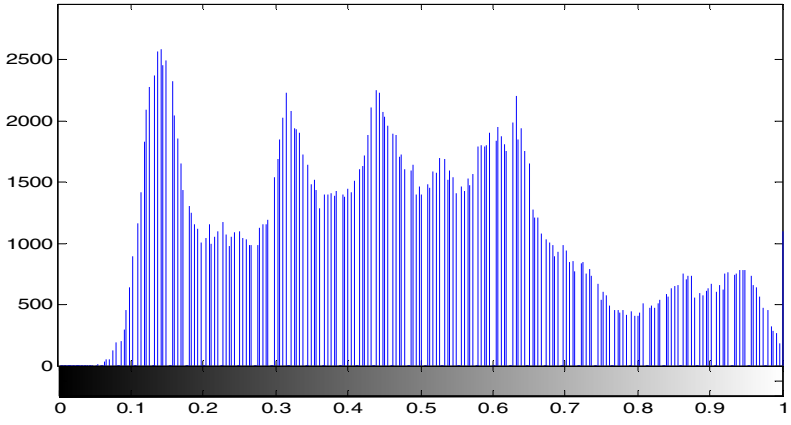


Fig. 5. Histogram after adaptive threshold value enhanced

As digital image is discrete, histogram equalization can not produce a image with balance histogram. But we can get a picture with uniform gray distribution. It can provide people more information, and increase the contrast of the images and the visual image size. After transformed, the gray level can be decreased, maybe some details will be disappeared, but it can usually meet the design requirement. An image with high quality can be got by histogram equalization and histogram specification. Two functions named "imhist" and "histeq" can be used in Matlab to realize histogram equalization and specification.

Fig. 6. Image with histogram processing

Change the image shown in fig.4, using the function in Matlab. The result is shown in fig.6, compared with the images in fig.4, it has a higher quality. The histogram is a good way to show the quality of the image. The image is more equal, and the histogram of the picture is more uniform. Compared with histogram in fig.5, the

histogram in fig.7 is more uniform. It doesn't have peak and widely distributed. Therefore, we can get the conclusion that the image in fig.6 has higher quality than others, and it can also mean that the method of getting image intensity through series of algorithms is better than the method of getting image intensity through any algorithm.

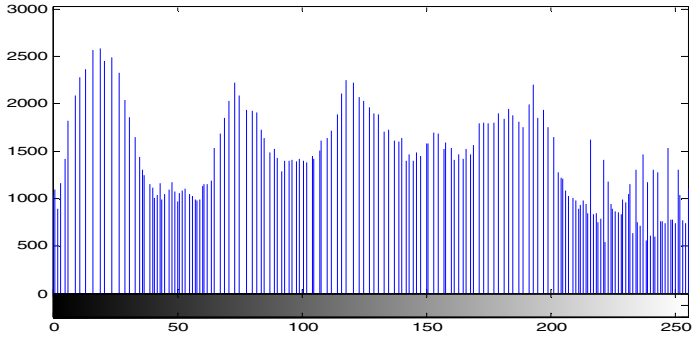


Fig. 7. Histogram with histogram processing

3 Summaries

Conclusion that the image in fig.6 has high quality than others can get through the picture above, It means that the method of getting image intensity through series of algorithms has a better result than the method of getting image intensity through any algorithm.

Digital image processing theory and technology has developed 50 years, it is not limited in aerospace and other minority frontier, developed into an independent science, and has penetrated the many industries and human life. Therefore, in modern society, image processing is becoming more and more important not only in theory but also on practice. In this dissertation, a method is provided, and it can provide a better result of image intensity. As the development of the digital image processing, there will be more methods provided. It can make contributions to the development of image processing.

References

[1] Conzalez, R.C., Woods, R.E.: Digital Image Processing, 2nd edn. Publishing House of Electronics Industry (2002)
[2] Gao, X., Su, W., Gu, H.: Maximizing sum of image intensity square autofocus algorithm. Chinese Journal of Radio Science (2010)
[3] Xin, Y.: A Fast Parallel Median filtering Algorithm Implemented. Microcomputer Applications (2011)
[4] Cui, J.: Application of wavelet Transform in the study of Expression Recognition. Journal of Changchun University of Science and Technology (2011)

[5] Yu, H., Guo, H., Guan, Y.: Application of Wavelet Transform in image processing. Artificial Intelligence and Identification Techniques (2011)
[6] Chen, L., Ding, G., Guolei: Image thresholding based on mutual recognition of histogram. J. Infrared Millim. Waves (2011)
[7] Zhang, J.: Image Processing with Wavelet Transform in Matlab. China Academic Journal Electronic Publishing House (2011)

A Novel Digital Audio Watermarking Algorithm Based on Wavelet Transforming

Liu Yanjun

School of Software, Northeast Petroleum University, Daqing 163318, China

Abstract. With the rapid development of information technology, more and more digital mediums like video, audio can be available online, which brings convenience to peoples' daily life and also results in the copyright violation. How to protect the copyright of digital mediums when transferring them over internet becomes a challenging problem. As a promising way, watermarking technology attracts some attentions. Yet there are still some shortcomings, like medium quality changing and low robustness which hinder its further application. Regarding this problem, a digital audio watermarking algorithm based on wavelet transforming is proposed and the workflow of this algorithm is described in detail. The result of simulation experiment shows that the proposed algorithm has capabilities of robustness, inaudibility and authenticity.

Keywords: digital watermarking, wavelet transforming, audio robustness.

1 Introduction

With the rapid development of the multimedia and the Internet technology, digital files become more and more easy for transmitting and accessing, and this problem can not be solved by traditional code method. Content providers are facing with the great challenge of how to protect the copyright of their digital files and prevent them from distributing without permission. Regarding this problem, the watermarking[1] [2]technique is proposed. It makes possible to embed a identity (such as identification data, serials number, text or image etc.) into multimedia files supporting copyright protection, secret communication, authentication and so on. Therefore, this technique can reach the aim to protect the benefit of authors who have digital article copyright and to protect the development of digital products. Yet the embedding of watermarking will cause the quality loss to the carrier which result in the leakage of watermarking message. To overcome this shortcoming, this paper proposes a kind of audio digital watermarking algorithm based on wavelet analysis is presented, the watermarking information which embeds in low frequency coefficient after wavelet transform. Experimental results show that the proposed watermarking algorithm is robustness and the change of audio quality can not be found by human perceptual system.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 109–115.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Related Works

As the digital watermarking technique has wide applications and great business opportunities, it has become a hot research area in the field of information security. Some of them with representatives can be listed as follows. Cai, Libin [3-5] present an objective audio quality measurement scheme by using semi-fragile digital audio watermarking. The scheme is based on Digital Wavelet Transform (DWT) and quantization, the watermark is a PN-sequence generated from a secret key, the optimization of the quantization step for different audio signals by employing an adaptive control method is also discussed. Chang, Chuan-Yu[6] [7]proposes a DWT-based Counter-Propagation Neural network (CPN) for digital audio watermarking. The db4 filter of the Daubechies wavelet is applied in this paper. The coefficients obtained from 4-level db4 and the corresponding watermark is used for training the CPN. Hu, Jian[8, 9] proposed a digital audio watermarking algorithm, which based on Neural Networks and in the Wavelet Domain and corresponding algorithms of watermark generation, embedding, and extraction. Li, Cheng-Hao presents a new digital watermark approach based on fractal image coding and a method to use the fractal code as a means of embedding a watermark into image is introduced. The achievements listed above partially solve the problem of watermarking embedding in digital audio, but consider less on the quality protecting, no mechanism is employed on the robustness of them.

Regarding this problem, a Novel Digital Audio Watermarking Algorithm Based on Wavelet Transforming is proposed and wavelet transforming model is applied to the design process of watermarking. The simulation experimental results show that the proposed method has capabilities of robustness, inaudibility and authenticity. compared with other traditional algorithms, the scheme has the advantage of extracting watermark without the participation of origin audio signal, and low computing amount.

3 Algorithm Descriptions

Audio watermarking is mainly reflected on the difference between data embedding and extracting scheme is different from earlier methods, there are four categories, which can be listed as follows, the least significant bit method, echo hiding algorithm, spread-spectrum method, the phase encoding method, and time-domain based method. The critical factor which infects the roburstness of audio watermarking lies on the embedding algorithm. Robustness and transparency of digital watermarking are two basic and most important requirements of the digital image watermarking technique. Yet the requirements of robustness and transparency are contradictive to each other. How to balance them is a challenging problem.

3.1 Summary of Wavelet Transforming

Wavelet analysis[10] is a new branch of mathematics, it is the pan functions, Fourier analysis, harmonic analysis, numerical analysis of the perfect crystal; in applications, especially in signal processing, image processing, speech processing and many nonlinear science field, it is considered to be following the Fourier analysis of time-frequency analysis is another effective method. Compared with the Fourier transform of the wavelet transform is a time and frequency domain local transformation which can effectively extract information from the signal by scaling and translation functions such as computing a function or signal of the multi-scale detailed analysis (Multiscale Analysis), Fourier transform to solve many difficult problems can not be solved. The whole process for DWT based audio watermarking can be divided into embedding and extracting.

3.2 Workflow of the Algorithm

(1)The watermarking embedding process
In our approach, we embed watermark by quantizing the DWT coefficients with a specified quantization parameter. The watermark's sensitivity against the modification could be controlled by using different value of. The watermark to be embedded is a pseudo-random binary string that is known only by the author, which can be described like figure1. There are three main steps in embedding process.

1) we choose a mother wavelet function and then compute the L-level discrete wavelet decomposition of the original audio. After the decomposition we obtain L+1 sets of coefficients, which includes 1 set of low-frequency coefficients and L sets of high-frequency coefficients. The value of L is defined by the user.
2) we embed the binary watermark string into the wavelet coefficients. In our approach, the watermark will be embedded into last K-levels of the wavelet coefficients, where K is equal or smaller than L+1. In each level, the coefficients are first divided into blocks, each of which is used to embed one watermark bit. The starting point of the first block should be a sufficiently large number so that the watermark bit will not be embedded to the silent region in the beginning part of audio signal. In each level, we divide the coefficients into 64 blocks. In each block, we choose the first 50 largest coefficients to embed the same watermark bit. For any discrete wavelet transform (DWT), the coefficients are real numbers. The quantization procedure could be performed on these coefficients. In quantization step, every real number is assigned a binary number 0 or 1.
3) we get the watermarked audio by using the corresponding L-level inverse discrete wavelet transform. Before we distribute the watermarked audio into the communication network, we will store a segment of the watermarked audio as the matching template, which will be used in the watermark extraction process.

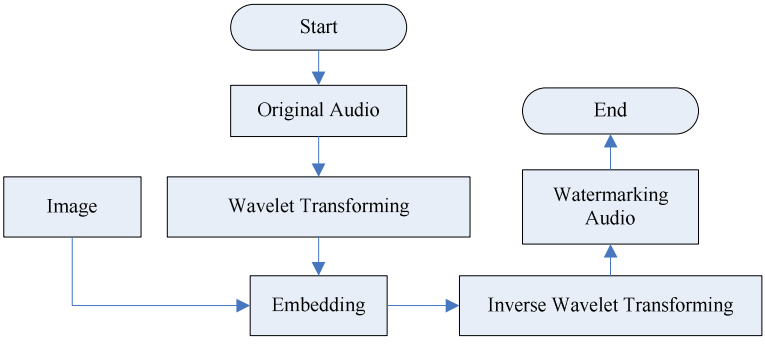


Fig. 1. The workflow of watermarking embedding based on DWT

(2) The watermarking extracting process
The watermarking should be extractable even if common signal processing operations are applied to the original audio signal. In extraction procession, original audio signal is not needed. The detail process can be described like figure 2.

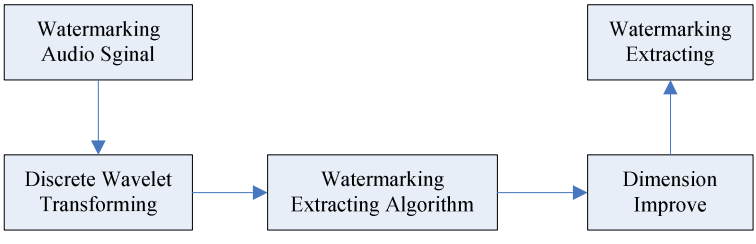


Fig. 2. Watermarking extracting process

Due to some intentional or non-intentional audio processing such as MP3 compression and decompression, the original audio could be shifted and the total number of the sample points could be changed. Therefore, the start point of the watermarked audio must be located before watermark extraction. We employ the classical matched filter to efficiently and fast fulfill the task once the start point is located; the L-level discrete wavelet transform is applied. The mother wavelet function should be the same one as that in embedding procedure. In every level of decomposed coefficients, blocks are divided. The first 50 largest coefficients in each block are selected for extraction. We only perform the extraction on these K levels. As a result, we get K strings of watermark after using the extraction algorithm in every DWT level. Each string corresponds to one level. After the extraction, we can check whether there is a watermark in the audio by using the original watermark string. This is copyright verification. We compare the original watermark with the K extracted strings. If one of the strings is similar to the original watermark, the verification will succeed. Another application is content authentication. The authentication succeeds only if the extracted K watermarks are identical to the original one. Almost any manipulations on audio will fail to pass the authentication.

If the authentication fails, an additional program will be employed to determine the modified audio content.

4 Simulation Experment

4.1 Experiment Design

To validate the effective of the algorithm proposed, the proposed technique has been experimented on a audio clips. that were WAV files sampled at 44.1 kHz, 16bit, 10s length, is used as the original digital audio signal at first, which adopts the 23 X 5 binary image as the watermarking. See Figure 3 (a).

(a) (b)

Fig. 3. Original and eventual image sample

In the process of embedding watermark, the digital audio signal is decomposed third-degree wavelet by selecting wavelet. To study the results of the proposed technique comprehensively. The performance of the proposed scheme is gauged by three different parameters. The visual imperceptibility of the audio watermark is measured by the Peak Signal to Amplitude (DB). The correlation between the original audio clip and the extracted audio clip is monitored. Wavelet decomposition is implemented by Daubechies-4 wavelet base with 3rd level. In order to make the watermarked signal inaudible.

4.2 Result Analyzing

According to the experiment design and parameter setting, the simulation is executed and the result can be described in detail. we embed the watermark into low frequency part of the large magnitude in among the detail coefficients of audio signal. The original audio can be described like figure 4.

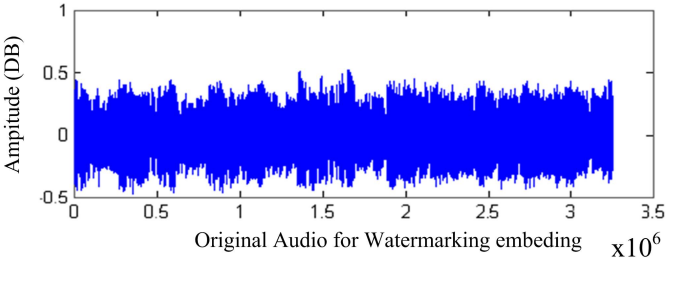


Fig. 4. The original audio for watermark embedding

With the algorithm described in 3.2 we embed the watermarking image into audio, the result can be shown like figure5.

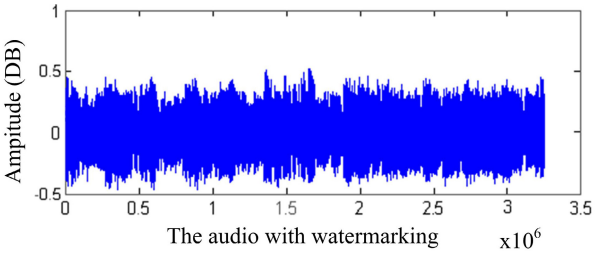


Fig. 5. The audio with watermarking

From the comparison we can find out that the quality of the audio have little change on the Amplitude. The figure3 (b) show that the image watermarking can be extracted without quality loss.

For further testing on the robustness, based on the experiment design in 4.1, the water-marking signal was modeled the Gaussian distribution to the length of 400, the speech signal sampling frequency to 8kHz, and watermark embedding intensity to 0:1. The original audio signal, we can find out that there is almost no difference between the original signal and the watermarking one which can be approved by hearing as well.

5 Conclusion and Future Works

Watermarking technology is a promising way in the audio copyright protecting. Yet there are still some weakness in current watermarking technology for audio. To overcome this shortcoming, the technology of wavelet transforming is applied to watermarking for improving the robustness, inaudibility and authenticity of the watermarking embedding technology and the result of simulation experiment prove the validity of the method. In the future, some efforts will be paid on improving the efficiency of the algorithm proposed.

References

[1] Chung, T.-Y., et al.: Digital watermarking for copyright protection of MPEG2 compressed video. IEEE Transactions on Consumer Electronics 44(3), 895–901 (1998)
[2] Wang, X.-Y., et al.: A new content-based digital audio watermarking algorithm for copyright protection. Association for Computing Machinery, Shanghai (2004)
[3] Cai, L., Zhao, J.: Audio quality measurement by using digital watermarking. Institute of Electrical and Electronics Engineers Inc., Niagara Falls (2004)
[4] Mazurczyk, W., Kotulski, Z.: New VoIP traffic security scheme with digital watermarking. Springer, Gdansk (2006)

[5] Zezula, R., Miurec, J.: Audio digital watermarking algorithm based on SVD in MCLT domain. Inst. of Elec. and Elec. Eng. Computer Society, Cancun (2008)
[6] Chang, C.-Y., Shen, W.-C., Wang, H.-J.: Using counter-propagation neural network for robust digital audio watermarking in DWT domain. Institute of Electrical and Electronics Engineers Inc., Taipei (2007)
[7] Robert, A., Knopp, R.: Detection theory and digital watermarking. SPIE, San Jose (2000)
[8] Hu, J., Qiu, X.-M., He, D.-T.: Digital audio watermarking algorithm based on neural networks. Inst. of Elec. and Elec. Eng. Computer Society, Chengdu (2008)
[9] Triet, T.M., Duc, D.A.: Applying the psychoacoustic audio watermarking technique in Internet Digital Traditional Music Museum in Vietnam. Institute of Electrical and Electronics Engineers Inc., Albuquerque (2004)
[10] Niu, X., Yang, Y.: New algorithm for digital watermarking based on the wavelet transform. Jisuanji Xuebao/Chinese Journal of Computers 23(1), 21–27 (2000)

The Demonstrating and Monitoring Software of Robot Control System

Zou Huan[1], Zhang Xueping[1], Yang Yanxin[1], and Wang Xin[2]

[1] College of Engineering and Technology, Yunnan Agricultural University
Kunming Yunnan 650201, China
[2] College of Water Conservancy Hydropower and Construction Engineering,
Yunnan Agricultural University
Kunming Yunnan 650201, China

Abstract. The robot is one kind of new comprehensive technology product. In the research of industrial robot, the demonstration robot is still the robot which is widely used at present. Therefore this paper introduces the realization and structure of the demonstration and monitoring software of the robot control system. Also this paper makes an entire illustration for the overall technical scheme, discusses the interface layout and the design and implementation of the monitoring software, and realizes the upload and download processing data under different work modes.

Keywords: the robot control system, the demonstration and monitoring software, communication.

1 Introduction

The robot is one kind of new comprehensive technology product, which has fused a lot of subjects [1]. Now the application of industrial robots is very extensive. Now, the international competition is more and more ferocious, the automation development of industrial production is imminent, and industrial robot just can solve such a problem. The appearance of robot, makes the high risk and hard work complemented by people be accomplished instead of robots with higher reliability and safety. The work of robot is accomplished according to strict procedure execution, so its accuracy rate is very high, which can improve the product quality and the technology level and increase the processing speed and manufacturing speed [2]. In the research of industrial robot, the demonstration robot is still the robot which is widely used at present. The research of industrial robot control system has a very important significance. With the current development, the development and application of robot in the industry are large. The robot is one part of a complete industrial automation system [3]. The control systems of all kinds of equipments in the production line are connected effectively via the network or industrial bus, form a comprehensive control system, which has become the development trend of modern production equipment. It will greatly speed up the data and information flow of the whole control system and sharing can be implemented really, so that the efficiency

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 117–123.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

can greatly improved at the process of designing the specifications of the products, system structure and performance [4]. However, there are not only one or two production equipment factories. Since many production equipment factories exist, the equipment in an equipment system may come from different manufacturers. Currently it is hard to form a reasonable automation system, so people put forward the concept of modular and opening and the robot system will develop toward this aspect development [5]. Internationally, many countries also consider the modular and open robot system as the main research direction.

The domestic robot research is late relatively. At present our country takes the robot research and manufacture very seriously. Therefore, the industrial robot has a great significance.

2 Experience Based Approach

The design goal of the robot control system is following: (1) the control degree freedom of the robot control system is six, the movement range and speed of the joints of the robot control system meet the user requirements. (2) the robot control system composes of 6 axis controller, and handheld demonstration and monitoring box, 6 axis exchange drives components etc, and LCD monitors of the demonstration and monitoring box is colorized TFT with size of 5.7 inch. (3) The control system provides demonstration and monitoring, programming, automatic and manual etc. work ways, the coordinate system has a right Angle, joint coordinates, and interpolation means has straight line, joint, circular arc etc. (4) The control system can complete the coordinated control with surrounding auxiliary equipment. (5) The control system provides the necessary security protection.

The performance and function of the robot control system is shown in Table 1.

The software platform used in this system is the Windows CE 6.0 system, this system is a mobile computing platform developed by Microsoft based on an embedded platform, which is a 32-bit system, and designed for the characteristics of the embedded platform. Windows CE system in custom is very convenient, and in accordance with certain steps freely chooses the required system function and mode, which has nothing to do with the processor, inheriting the style and function of Windows interface, can easily be modified and transplanted to apply different application environment.

The function structure of monitoring software is shown in Figure 2, according to their functions, the monitoring software includes: file management, variable management, I/O, parameter setting, communication between upper and lower computers, alarm, graphic simulation, manual module, automatic operation. The detailed descriptions are as following: (1) the file management achieves new, editors, modification, preservation, delete, and other basic documents operation of program files. (2)The variables management achieves the marking editor of integer, long integer, floating point type, byte type, position type variables and the check of the values. (3)I/O completes the operation, the display, identification and the definition of the robot I/O signals. (4)Parameter setting completes the set of operating conditions of demonstrating and monitoring box and the set of all the parameters needed by the robot operations. (5) The communication between upper and lower computers completes the communication function of the demonstrating and monitoring box with

Table 1. The performance and function of the robot control system

The robot parameters		values
Number of axles		six
Min OT		0.01mm
Min action range	S axis	-150° +150°
	L axis	-150° -30°
	U axis	-50° 70°
	R axis	-150° +150°
	B axis	-90° +90°
	T axis	-180° +180°
Max speed	S axis	120°/s
	L axis	120°/s
	U axis	120°/s
	R axis	300°/s
	B axis	300°/s
	T axis	450°/s
Control precision	Dynamo orientation	0.01°
	Extreme oritation	0.01mm

The industrial robot control system is composed of demonstrating and monitoring box, a lower computer ARM and the motion controller. The hardware structure is shown in Figure 1.

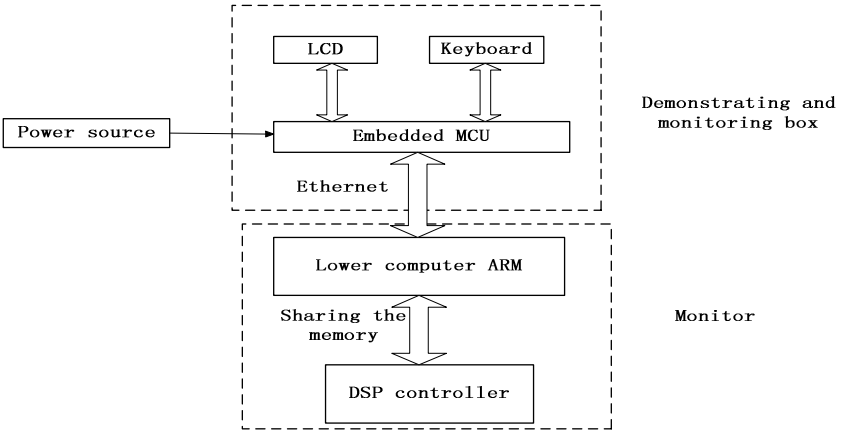


Fig. 1. The hardware structure of the robot control system

the lower computer ARM and downloads programs, parameters, command word and the state information on the robot. (6) Alarm processing completes the detection of all kinds of breakdown in the process of robot run and displays them to the user. (7) Graphic simulation makes the entity simulation function in the process of robot run. (8) Manual module includes two kinds of modes: JOG, a single step, and downloads the order to make robot run in accordance with specified operation. (9) Automatic operation completes the full operation of one program.

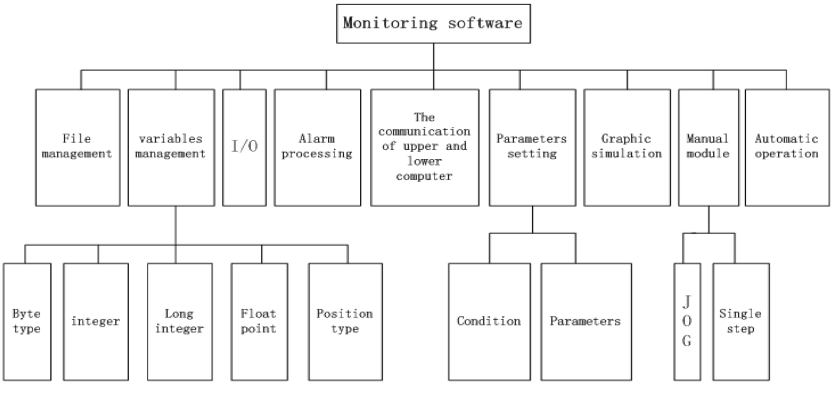


Fig. 2. The function structure of monitoring software

3 The Design and Implement of the DomonstrainG and Monitoring Software

As the man-machine interface, the monitoring interface of demonstrating box should have a reasonable layout. This software uses the interface with a menu, and strives to

make operation humanity and simplicity. According to the interface introduced in this paper, upper interface is the display area for robot operation state.

The monitoring system of demonstrating box is different from other industry monitoring systems, which operates in embedded system. The demonstrating box cannot be operated and edited like common keyboards; therefore we must implement the special instructions input methods of the demonstrating box. This software realizes the editing process of programs using the instruction menu, insert, delete and modify is implemented for one entire line operation instructions, insert instructions are realized by instructions menu, the pop-up of command menu is completed by pressing the command or interpolation key, then select needed instructions. The robot command are divided into mobile instructions, the translation rotating instructions, coordinate transformation program control instructions, input/output instructions, check instructions and PAWS control instruction.

The programs based on MFC under Windows CE can generally be realized using the loaded menu. But this way needs to add a response function for every possible input and the supported instructions parameters, the number will achieve several hundred, make the readability of the program greatly reduce and work efficiency descend, also go against the follow-up software development. We can conveniently read menu content through obtaining the handle of the son menu, but when instruction system needs to be adjusted, the program will need very big change momentum, and prone to BUG, also go against the changes and development of the follow-up process, therefore, the design of the commands menu abandoned the method and adopts the multi-dimensional array storage instructions to realize its functions. In this method three level menu contents are deposited to the array, the position index in the menu and the array index of the instructions and parameters will be corresponded.

The robot operation mode expounded in this paper includes four kinds of way: a single step, JOG, automatic, and back to zero point. The manual mode is designated a joint exercise or translation along a direction, just needs to download instructions, and does not need to download program, back to zero mode transfers back the zero order, single step model only downloads a specified line of the program, automatic mode downloads all programs. Automatic mode will download all procedures to the next place machine, because of downloading all the programs, when the program is too long, because the receiving buffer of Ethernet communication lower computer can't accept too many bytes, this will lead to the loss of the program. In order to solve this problem, when downloading the program, the program will be sectioned and it will be partially downloaded. Before downloading the program, the program can be separated into N sections according to the buffer size, each section with length 2 K, and then download each section to the lower computer according to the order.

The upper computer reports the data with a form of data frame and processes the received data frame. When the lower computer sends data too quickly, several data frames needed to be sent may be connected into a data section to be sent. Therefore in this case the upper computer must segment the received data. Each data frame has an

end identifier "&" to illustrate this data frame having finishing the send. When dividing the data, firstly we find the position of the first end identifier "&", intercept the first paragraph data frame and process it, and then the rest of the data is captured out as a new data section, repeat above actions, until all the data frame is isolated and disposed. The segment of the uploaded data is shown in Figure. 3.

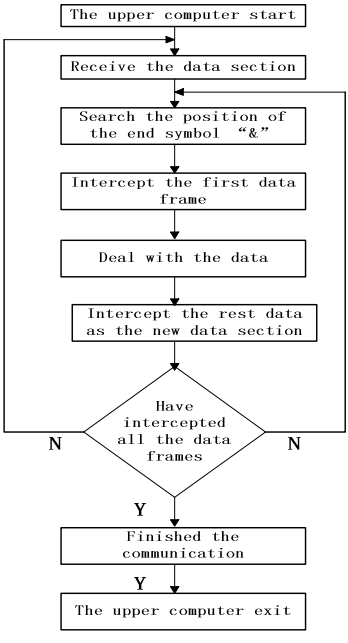


Fig. 3. The segment of the uploaded data

4 Conclusions

This paper introduces the realization and structure of the demonstration and monitoring software of the robot control system. Also this paper makes an entire illustration for the overall technical scheme and the structure of monitoring software of the robot control system, discusses detailedly the interface layout, procedures and instructions editing system design and implementation of the monitoring software, realizes the upload and download processing data under different work modes, as well as the realization method of communication between the graphic simulation process and the main process between. At last this paper detailedly illustrates the implement method and communication agreement between monitoring software and the lower computer.

References

[1] Chen, Y., Shao, J., Zhang, C., et al.: The research progress and prospect of automatic path planning of spraying robots, pp. 149–151 (February 2010)
[2] Gao, L., Cao, M.: The design of demonstration and monitoring box for robots. In: The Academic Conference on Automation in Jiangsu Province, pp. 115–119 (2008)
[3] Liu, S., Chen, Z., Liu, Z.: The application and prospect of virtual reality technology in robot technologies. Information Technology and Informatization 4, 42–44 (2010)
[4] Chen, S.: The design of the demonstration and monitoring box system of robots. Manufacturing Automation 27, 75–78 (2005)
[5] Zhu, W.: The research and development of the high-end demonstration and monitoring box system of robots. The master's degree thesis of Shanghai Jiao tong University (February 2003)

Discussion on Training Mode of Entrepreneurial Talent in Electronic Business

Jin Tao

Tourism College of Zhejiang China, China

Abstract. Entrepreneurial talent training is the inevitable choice of social development. It can turn job seekers into the engine of job creation in the future. Entrepreneurship becomes a hot topic and trend of reform in vocational graduate employment system. College teachers and administrator have to make change and innovation in education system, ideas, methods and content according to the entrepreneurial talent training mode in electronic business.

Keywords: high vocational, entrepreneurial, electronic business.

1 Introduction

Our graduate employment situation in our nation is not optimistic. Governments in all levels pay a lot of attention on undergraduate employment. The premier Wen Jiaobao has said in government work report of the 4th of CCCPC's 13th session, "we should encourage entrepreneurship." The training of vocational students is prepared for obtaining employment. And entrepreneurship is just one kind of employment methods. Vocational entrepreneurship in electronic have advantages of small investment, low cost and is suitable for undergraduates' situation. Meantime, it meets the need of employment situation and alleviates employment pressure. Therefore electronic business has been the option of many vocational high school students and obtains the approval of colleges and society. Lots of institutes have also established e-business incubator to support entrepreneurship. With education industrialization in our country, the number of graduates increased strikingly every year. Electronic business entrepreneurship is fit for vocational students under the limited economic conditions. Simultaneously domestic industry of electronic business is developing rapidly. Large numbers of e-business enterprise cases demonstrate that it is an important way for venture. E-business entrepreneurship has low energy consumption and less social resources which is qualified for environment-friendly level and "emptying the cage, removing the bird" policy. So the electronic business entrepreneurship meets the need of social development.

With fast development of higher education, the number of graduates increases strikingly every year, which worsens the employment situation. College students have to promote enterprise self-competency and turn themselves from job seekers into the engine of job creation and from passive service to active options. They

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 125–131.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

must possess higher quality to face choice. Developing e-business education, fostering enterprise competency and offering more relevant social experience are the embodiment of work- learning integration. Besides it also can broaden the vocational student's employment channels and turn them into the pillar of the society. Stimulating their entrepreneurial passion and improving their business quality surely promote their roles and turn them from job seekers into the engine of job creation.

2 Existing Problems in Training Mode of Electronic Business Entrepreneurial Talent

In 2001, the electronic business major was first set up in some colleges. So far there are more than 300 universities and institutes opened this specialty. They all have their own development emphasis in talent training modes. Some focus on business management operation, some on information technology. Every college gain certain experience and exist some problems during the development of the training mode.

2.1 The Talent Training Goal Is Not Set Accurate

E-business specialty involved a wide range, including various aspects of knowledge such as trading, economy and network techniques. This required a high level to foster students. Only by determining the employment direction for the future talents we can set the benchmark, direction and basis for major construction. Develop professional courses system and establish effective professional education plan can make the plan meet the social need for talents.

2.2 Systematic Insufficiency in E-Business Education Courses Setting

E-business is a rising multi-disciplinary subject in recent 10 years, involving economy, management, trading, internet technology and laws. As for courses system, the integration is not enough. Especially the courses on business and technology lead to a bad theoretical frame to students. Besides there are some drawbacks in the practice course, such as not enough cases to contact and insufficient practice or exercise.

2.3 Shortage of Teachers in This Major

Today most of teacher in this major is transformed from other related professions, some from economy, some from management and others from computer science. These teachers acquire knowledge from e-business textbooks which lack a good theoretical system. The e-business theoretical system is still improving. Lots of teachers have never been working in relative enterprises. They are shortage of actual

operation experience and combat practice. The cases at classes are most from books. There are few activities and discussions during the cases analyses. The enterprise guidance to students is still not enough compared to the real operation. The shortage of teachers influences the quality of e-business talents training directly.

2.4 Shortage of Education Practice

E-business require a strong hand skill, high practice ability and application capability. With practice purpose certain colleges have purchased related e-business simulation software in labs for students. However e-business simulation system is significantly different from real commercial environment. It cannot simulate the integrated real environment, even it is not better than current application. Actual social e-business is complex and varied. Fixing all the commercial action differs from the real social work. Therefore to promote application ability, we should jump out the simulation environment and seek for real business practice.

3 Training Mode of E-Business Entrepreneurial Talents

Each reform has a hard process in a wide range. So is the innovation of e-business entrepreneurial talents training mode. College teachers and administrators have to make some change and innovation in system, ideas, method and content. It also needs to contact with related enterprises, know about the demand conditions and features and do research on constructing e-business knowledge structure system and capability structure system. According to this, we should adjust or set profession course system, complete syllabus, compile textbooks, reform education method continuously, consummate comprehensive practice project, demonstrative venture project and build reasonable talent assessment target. The guidelines of this training mode are as follows:

3.1 Entrepreneurial Talent Theory, Quality and Cognitive Systems

Theory system includes four aspects: professional theory, business theory, management theory and legal theory. Professional theory means specialty basic theory as well as personal theory construction and arts knowledge as well as science knowledge. Business theory is a systematic theory in enterprise operation and marketing. Management theory is a theory on how to manage people, wealth and materials. Legal theory contains economic law, civil law, corporation law and other business regulations.

Quality system is made up of 3 parts: learning quality, creative quality and leading quality. The quality on how to acquire knowledge, reprocess and apply them is called learning quali is ty. Creative quality means the ability to produce innovative ideas and put them into application during the learning and management process. Leading quality is the combination of coordination skill and management capability.

Cognitive system concern 5 components: perception of risk, perception of integrity, perception of responsibility, perception of cooperation and legal cognition. Perception of risk means that talents should face the risks and avoid them with rational option and judgment. Perception of integrity means talent must do everything based on integrity if he wants to gain a good reputation and establish himself in commercial society. Perception of responsibility stands for social sense of duty and the love for society, team and staff. Perception of cooperation means that entrepreneurs must fully understand success to be the result of teamwork. Only full team cooperation and team mutual help lead to a good performance. Legal cognition tell talent that he should be familiar with all kinds of laws, regulations and government support policies on venture with purpose of rational avoiding tax, paying tax, enjoying tax favors and running business within the bounds of the law.

3.2 Entrepreneurial Talent Training Course Structure Design

According to the words above, I consider the main modules of entrepreneurial talent training course system should consist of 3 following aspects:

First is the literacy training module. This module is the basic part of talent entrepreneurial knowledge system. Humanistic literacy is basic factor in talent training. As a humanistic nation, people's humanistic literacy has a huge influence on society. So we should pay a high attention on the training of humanistic literacy during entrepreneurial process. By this kind training, we can improve the students' overall quality, psychology bearing capacity and personal competencies. The aim of this training is to let them master the necessary humanistic knowledge in venture and have an idea about personality and psychological quality an entrepreneur should have. We can apply this theory by offering various special lectures, such as inviting some professional experts to speech. Second we can open some related courses, such as <etiquette practice>, <practical writing>. These measures can improve the students' basic qualities, thinking skill and problem solving skill, strength students' stress tolerance and psychological quality and make students master the practical application skill such as managing applied documents.

Second should be knowledge imparting module. This one is an important part of the system. Learning related knowledge students can understand basic flow, primitive rules, methods and skills. It involves how to establish e-business enterprise and curriculum. We can open up certain courses to put in practice, such as <network marketing>, <cyber retailers>, <international trade and commerce>, <business management>, <HRM>, <business negotiation>, <marketing scheme>, <e-business venture>. By learning this module, students can improve entrepreneurial skills, know how to set up enterprises, lawful operation and master the knowledge about how to adjust marketing environment and demand changes.

Third is enterprise skill training module. It is a higher module in this system. Its purpose is to transform business awareness and knowledge into enterprise skill. There are several methods we could use to achieve our goal. (1) Undergraduates

entrepreneurial contest. Regular venture planning contest, "challenge sup" and other competition activities organized by Communist Youth League can stimulate the development of entrepreneurial practice. Sometimes this also can cultivate students' entrepreneurship and innovative awareness. (2) Communicating lecture on successful entrepreneurial experience. We can invite some graduate entrepreneurs to give lecture to college students, communicate and share success experience and ideas. (3) Open op and operate enterprise to practice. We can found business incubator in the school to provide place for students to run various printing shop, supermarkets, stores and companies. We can take advantage of local government policy to support college students' venture. Then we can promote their leading and managing capabilities.

4 The Method to Stimulate Entrepreneurial Talents Training Mode

4.1 Increase Input and Strengthen the Construction of Vocational College Teachers

Teaching quality and talent training quality are closely related to the quality of teachers. We can take following measures to the construction of professional teachers. First, college can encourage teachers to attend related academic conferences and e-business communication. Second, teachers should take part in some professional training courses organized by colleges and social group. Third, schools should set policy to encourage teachers to assume e-business researches, horizontal subjects and vertical ones. Fourth, we should set the practice regulation on teachers being in enterprises and encourage teachers to stay close to the related enterprises. They should cut into the related enterprises and companies and take part in e-business project development plan. Last should be finished part-time teacher regulations. We should support schools to hire e-business experts to work as practice-guidance teachers or class teachers.

4.2 Make Innovations on Practice Education System

The practice education session in e-business involve 3 aspects: first is business practice in e-business circumstance, such as internet market research, design on online marketing plan, online advertising design and application. Second is simulating all e-business operation practice, such as e-business simulation experiment, such as e-business simulating experiment. Third is e-business system design and development experiment. Then we can enhance practice education through following methods. First, they must build up school-enterprise cooperation mechanism. Enterprise offer job opportunities and position. Schools lead students to attend training in cooperation enterprise which participate in related curriculum. Second, e-business professional laboratory should be established in campus. Third, we could separate one teaching

class into several groups by dormitory or other ways. Let them set up online stores and operate entrepreneurial projects. This can educate students' team spirit and accumulate entrepreneurial practical experience.

4.3 Promote Job-Study Combination

Job-study combination is significant link in vocational school talent training mode. Most colleges cooperate with representative enterprises, setting factory in campus or school in factory. This method can promote students' practice skill and entrepreneurial quality. According our national e-commerce and economic environment development conditions, we can strengthen job-study combination by following means. First, schools should cooperate horizontally with traditional enterprises in e-business. Schools develop and build up e-business application system for traditional enterprise, join the e-business operation project, solve problems in e-business operation and bring in e-business programs. Second, we should encourage students to practice in related firms during holidays or festivals, joining the e-business operation activities. Third, schools and firms should begin order type training, opening order classes and building enterprise type colleges. Both sides form administrative committee to lay down the training plane and foster the e-business talents for the need of firms together. This kind of talents will have a solid foundation in entrepreneurial skill and quality.

4.4 Put Entrepreneurial Quality Education through the Overall Process of Vocational Talents Training

Centered on the goal of fostering e-business talent, we should take full advantage of e-business professional characteristics and superiority of online venture. Starting from completing education system and transforming education mode, the next step is to refine organization structure, build education platform and entrepreneurial quality education mode and finally put quality-oriented education through the overall process of vocational talents training. We should pay attention to developing students' entrepreneurial awareness and enriching their knowledge structure at the freshmen stage; improving their quality by campus practice at the sophomore stage; enhance their practice skill by enterprises post internship. And finally these will cultivate students' innovative entrepreneurial awareness, strong psychological quality and skills.

5 Conclusion

In conclusion, all above means could refine talent entrepreneurial knowledge hierarchy and broaden the students' employment prospects. They will bring positive political and economic effects to colleges and society.

References

1. Liu, F.: Briefly discussion on the question and response of vocational entrepreneurial education. Journal of Wuxi Institute of Commerce (5) (2011)
2. Tan, L., Li, W.: Entrepreneurial talent training:innovation of electronic business talent training mode. Technology Information (27) (2009)
3. Chen, Y., Zhang, X.: Construction of vocational e-business entrepreneurial talent training mode. Vocational and Technical Education (8) (2009)

The Design ECU of Vehicle with the Use of Controller Integrated PICI18F258 SCM

Shujun Hua

School of Mechanical and Electrical Engineering, Hainan University, Haikou 570228 China

Abstract. With the purpose of designing vehicle CAN bus network, we design the network of CAN bus on a certain vehicle model. Vehicle ECU with the functions of collecting sensor information and completing information exchange is developed based on CAN controller integrated PICI18F258 SCM. We use LABWIEW software to conduct system simulation experiment and automobile bench test. The result indicates that this ECU can precisely collect sensor information and achieve reliable communication. This design not only contributes to vehicle data communication, but conveniently develops matched portable diagnostic instrument to read information of vehicle technological status as well. It has a wide application prospect.

Keywords: CAN-BUS, vehicle ECU, LABVIEW software, PIC18F258 SCM.

1 Introduction

Controller LAN belongs to serial communication which is a new type of USB proposed by Bosch Company in SAE conference in 1996. It has been widely used in vehicle data communication between internal measurement and execution units [1]. And it can effectively support real-time distributed control of high safety level. The application range of controller LAN CAN becomes wider and wider, from high speed internet to cheap multiple cabling net. In the automotive electronics industry, the transmission speed can reach 1Mb/s when controller LAN CAN connects to EECU, sensors, antilock system and automatic transmission system. Meantime the controller LAN CAN can be installed in electronic control system such as car light group and automatic transmission system, replacing traditional wiring system.

Based on a new vehicle prototype, this paper designs an ECU system with CAN controller integrated PICI18F258 SCM. For assembly with CAN interface we can read data directly according to protocol. As for assembly without CAN interface we can use PIC18F258 SCM to design vehicle ECU to collect sensor information, use CAN interface to send data to bus in broadcast way and build vehicle network based on CAN bus. This method is benefit for vehicle data communication and developing matched portable diagnostic instrument to read information of vehicle technological status [2].

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 133–138.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 The Formation of Controller LAN CAN

Controller LAN CAN is the only bus protocol at present that can cover most application range in the field of vehicle mobile network. CAN 2.0 specification mainly defines the transmission layer and CAN protocol functions in around layers. But it doesn't set the media linkage unit, resident media and application layer. Therefore users can create data communication directly based on CAN 2.0 specification. However this type of transmission content cannot be modified flexibly which is suitable for fixed telecommunication.

Because of many unique designs, CAN bus possess outstanding reliability, instantaneity and flexibility compared to common telecommunication one. It has following advantages: ① high utilization rate of bus, long data telecommunication distance (over 10km). ② replace expensive vehicle internal wires with cheap twisted pair cables and greatly reduce the wire number. ③ receive or shield messages according to their IDs. ④ reliable error testing and handling mechanism. ⑤ resend the message when it suffers damage. ⑥ node automatically quit the bus in condition of having severe fault. ⑦ The biggest advantages of CAN is that data information in any node does not contain addresses of sending node or receiving node. The information content is marked by former identifier (ID). The ID is exclusive in the whole net. Other nodes in the network will test ID of the message whether its content is related to themselves after receiving the messages. If relevant information, the nodes will receive and deal with it. Or it will be neglected. In this way, different nodes can receive different data which guarantees the instantaneity [3].

Fundamental structure of vehicle information network based on CAN bus is shown in figure 1.

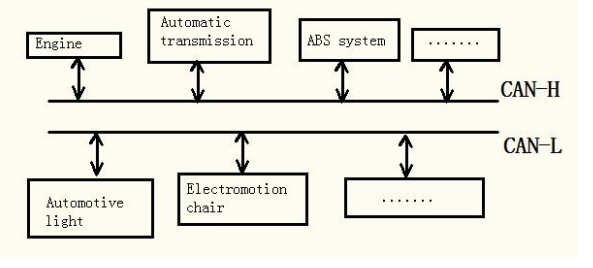


Fig. 1. CAN bus network diagram

2.1 The Features of CAN Micro-Controller Integrated PIC18F258

PIC18F258 is a micro-controller integrating CAN module which contains advanced Reduced Instruction Set framework, enhanced core, 32-level stack, Flash software integrated reservoir, EEROM data storage set, self-programming, ICD and various internal or external interrupt source. It adopts "Harvard" structure which completely separates program and data space. The pins definition is shown in figure 2 [4].

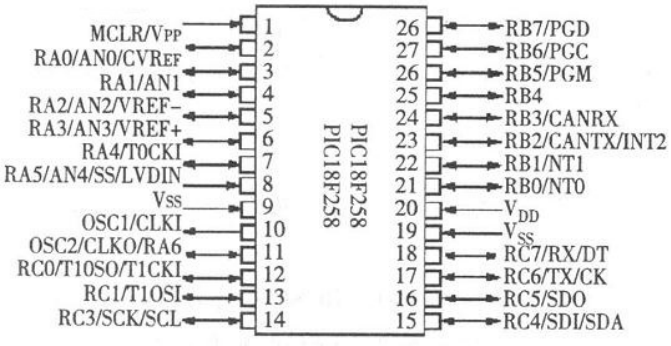


Fig. 2. PIC18F258 microcontroller package diagram

CAN module of PIC18F258 follows the CAN bus protocol and has its own characters. It includes:

1) It supports CAN protocol including CAN 2.0 A and CAN 2.0 B.
2) It supports standard frame, extensive frame, overload frame and error frame.
3) It has 2 receiving buffers with high priority and 3 send buffers with high priority.
4) It has 6 receiving filters, 2 corresponding receiving buffers with high priority and 4 corresponding receiving buffers with low priority.
5) It has 2 shielding filters respectively corresponding to different receiving buffers.

3 Design of Vehicle ECU

CAN nodes is implemented through CAN controller integrated PIC18F258 SCM.

3.1 Hardware Design

We can use timer or counter module to collect relevant pulse signal of sensor, A/D converter of PIC18F258 to acquire relevant voltage signals and CAN transceiver PCA82C250 to fulfill receiving and sending tasks. To improve anti-jamming of the system, high-speed optical coupling chip 6N137 is added into between PIC18F258 SCM and PCA82C250 chip. Hardware circuit principle of CAN telecommunication system is shown in Figure 3 [3].

PIC18F258 SCM is the core of CAN bus interface circuit which is responsible for collection of sensor information, initiation of internal CAN controller and telecommunication tasks such as data sending and receiving. PCA82C250 is the interface between CAN controller and physical bus, providing differential sending and receiving capacity. Then it only needs two data lines including TXD and TXD when conduct data operation, which greatly simplify the hardware circuit and improve the reliability of the system. The pins of CANH and CANL of PCA82C250 respectively connect to CAN bus with a resistor which can act current-limiting to protect PCA82C250 from overflowing impact. Besides one 30pF capacitors are installed between CANH and ground. So is CANL. This can filter the high-frequency interference and electromagnetic radiation. There is one protection diode reverse installed between CAN buses access end and

ground. It can act overvoltage protection with follow current of diode when CAN bus obtains high negative voltage. The 120Ω resistors in both sides of the bus act the function of matching impedance. To improve the system anti-jamming and telecommunication capacity, the design adopts 6N137 high-speed photoelectric coupling circuit to realize electrical isolation among nodes in the bus.

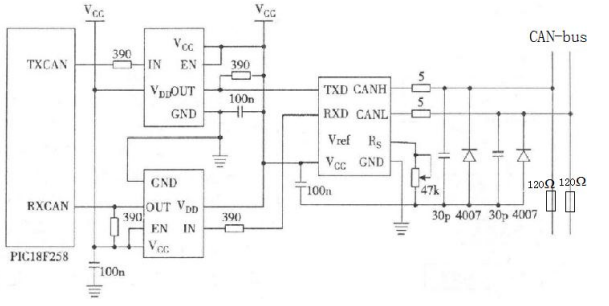


Fig. 3. PIC18F258 communication system circuit diagram

3.2 Software Design

To improve the telecommunication reliability, module structure is adopted in the design of internal software and node. It includes initiation subprogram, data sending and receiving subprograms of CAN node. It uses timer or counter module to collect and shape the sensor pulse signals. By using A/D module integrated in PIC18F258 SCM it can collect some normal sensor simulation signals.

CAN interface communication program mainly involve CAN controller initiation and data sending/ receiving programs. The main program implements data sending and receiving by calling functions. The processes are as shown in figure 4.

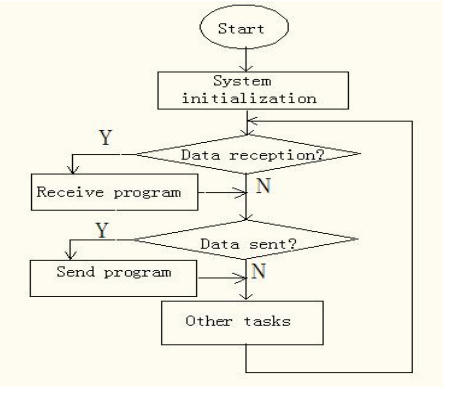


Fig. 4. Main program flow chart

Inside the data receiving program, the data of message object is read by query method. CAN controller initiation mainly includes configuration of CAN module's

working manner, settings of receiving filter, settings of receiving/shielding register, parameters of Baud rate, sending priority setting and settings of interrupt enable register. After finishing initiation settings of CAN controller, it will return to working status and continue normal communication job [6].

4 System Verification

4.1 Simulation Verification

First we choose NI's CAN bus test card. Then the communication with CAN bus can be implemented by calling sub-VI of NI-CAN in LABVIEW program environment. Figure 5 shows the sub-VI provided by NI-CAN.

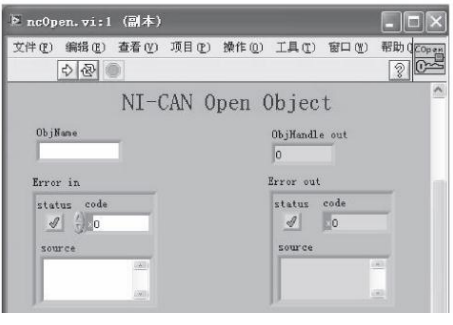


Fig. 5. NI-CAN control panel

In the LABVIEW 8.5 development environment, we can conduct real-time monitoring of the vehicle states (speed, accelerator pedal signals, mileage range, accumulator voltage) and display of failure self-diagnosis of controller units. After CAN bus data being collected and read into sub-VI of NI-CAN, the program will classify these data and conduct analysis. It selects the data frames according to ID values of data frames. CAN bus data collection is divided into two types -- distribution-before and distribution-after. Data will be selected and stored based on the request of test and analysis. Network node simulation is to simulate some node to send messages into the bus in current test port. So ECU function test can be conducted under simulation environment with node simulation in bus. The result indicates that this ECU performs very well.

4.2 Experiment Verification

The experiment uses this vehicle ECU to collect pulse signals (revolving speed) and simulation signals (pressure) of vehicle. Data transmission is conducted through CAN bus. It adopts digital storage oscilloscope (OSC) of vehicle fault diagnosis instrument VAS5051 to inspect the signals and parameters in CAN bus. DSO display under condition of no jamming function is demanded when using DSO to test voltage of CAN bus. During measuring, we should pay attention to the adjustment to time value, voltage value and trigger time of DSO. The test result is shown in figure 6.

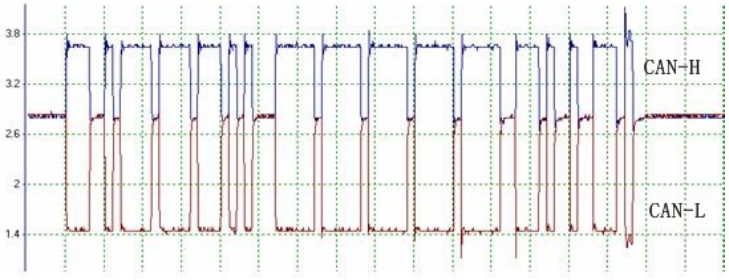


Fig. 6. CAN messages

Experiment adopts CAN 2.0 protocol, using VAS5051 tester to test CAN bus data of some ECU node. According actual test result, we can see that bus did not receive error frame which reflect normal working status. Receiving data indicates that ID codes and data of every ECU node is the same to the predefined ones. The bus receiving and sending is normal.

5 Conclusion

High-performance and high-reliability of CAN bus leads to its wide use in most fancy automobiles. PIC18F258 SCM integrating CAN controller has characters of high integration density, stable performance and strong anti-EMI capability. This paper designs vehicle ECU from both hardware and software aspects. PIC18F258 SCM has multiple interfaces to collect sensor information and integrates CAN interface module which make itself suitable for vehicle CAN bus network. Experiment indicates that this ECU can precisely collect sensor data and conduct reliable data telecommunication on CAN bus. Vehicle ECU based on PIC18F258 SCM designed in this paper has been optimized both in hardware and software which can promote system stability and anti-jamming capacity. So the bit error rate is almost close to zero.

References

[1] Bosch. CAN Specification V2.0 (1991)
[2] Cenag, G., Valenzano, A.: An improved CAN field bus for industrial applications. IEEE Transactions on Industrial Electronics 4(12), 4 (1997)
[3] Vehicle Application Layer standard. J1939-71. SAE Group (December 2004)
[4] Microchip PIC18FXX8 Data Sheet.28/40-Pin High-Performance, Enhanced Flash Microcontrollers with CAN Module
[5] Luofeng: Principle, design and application of vehicle CAN stack system. Electronic Industry Press, Beijing (2010)
[6] Nan, J.: Vehicle SCM and on-board bus technology. Beijing Institute of Technology Press, Beijing (2005)

Encrypt Hardware Device Based on Cryptographic Algorithm

Chen Cong

Department of Computer and Science, Guangxi Wuzhou University, China

Abstract. Information technology has been developing during the past decades, and encryption algorithms are most important to Information technology, in this paper we made some research about some famous international encryption algorithms, such as advanced encryption standard (AES), and Data encryption standard (DES). Smart Cards which are some high-end encryption card are now commonly used all around the world and becoming more and more ubiquitous. We also give a basic discussion of the smart card structure. Encryption algorithms were security enough from the perspective of math, but when realized in hardware, it may possess some electric characteristics, just like electromagnetic or power, which may be used by illegal users to attack the circuit, so the research development of relative resistance technology is important.

Keywords: information security, AES, DES.

1 Introduction

Now great quantities of information transmitted from terminals to terminals, from clouds to individuals, through the Internet and the Internet of things, all the data must be protected by encryption algorithms. In this paper we made some research about the encryption algorithms commonly used, such as AES and 3DES, also one of the hardware realizations of the algorithms-smart card will be discussed, we will give a basic description of the structures, the last part is some problems when the algorithms realized in hardware method. Paper [2] is the research about AES, authors made some research about masking technology. Paper [3] and paper [4] is about DES, the former is the application of DES in network, and the later is the research about the efficient of DES when realized in hardware. Paper [5]

2 Developments of Information Security and Encrypt Hardware Device

Information technology has been developing during the past decades, until now soft ware and IC technology have greatly influenced people's life. Actually they have

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 139–145.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

together connected with almost every device from military to computer games, from science experiments to engineering projects. At the same time great quantities of information transmitted from terminals to terminals, from clouds to individuals, through the internet and the internet of things (IoT), all the data must be protected to make sure of “reality”, “privacy”, “integrity”, ”availability”, and” “undeniable”. “reality” means the source of information is real, users can reach the real information, “privacy” means illegal users cannot reach the information, “integrity” means illegal users should not define and modify the data, “undeniable” means legal users must admit the visit to information, this is particularly important in electronic commerce [1]. Encrypt module is among the most important modules, though soft ware way is in a well established way of technology, which incorporates decades of academic and industrial researches and developments, we have been engaged in low-end encryption chip(hardware), for decades, so it is important to develop high security information security IC technology.

3 Basic Structure of Encrypt Hardware Device

There are lots of kinds of Encrypt hardware devices, all are used in information data encryption. In this paper we are focused in encryption chip. Smart Cards which are some high-end encryption card are now commonly used all around the world and becoming more and more ubiquitous. E-commerce, citizen administration, and others could be, through the Internet, good vehicles to allow service providers to develop new services using the Smart Card as a high-security key element. In a typical smart card based password authentication scheme, users are authenticated with their cards as identification tokens. The smart card takes as input a password from the user, creates a login message from the given password, and sends the message to a remote server, which then checks the validity of the login message before allowing access to any services or resources.

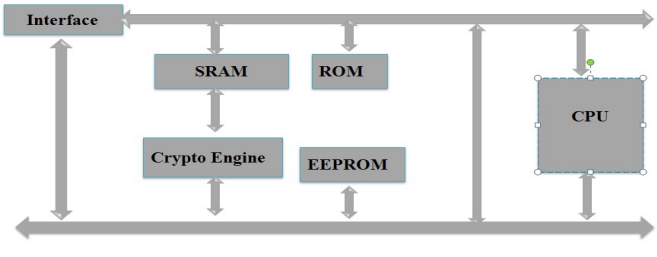


Fig. 1. Basic structure of smart card

Figure 1 shows the basic modules of smart card. Interface is used for communication between card and reader, there are two kind of interface: contact interface, and

contactless interface. Contact interface has a standard of 7816, which define the form, area, protocol and so on of the card, so all the cards have unified contact interface, while the contactless interface use the technology of RFID. CPU was used as a controller of card, all the data will be calculate by the CPU, but some cards have a "coprocessor" to cooperate with CPU. Crypto engine was used to encrypt data, this module was the hardware realized encryption algorithms, the most commonly used algorithms are AES, DES, RSA, and ECC. SRAM act as memory, and EEPROM, ROM were used to store data. All the data in the chip were transmitted in bus.

4 Application of Advanced Encryption Standard Algorithms

Advanced encryption standard was created as a kind of international encryption algorithms. AES is a symmetric-key algorithm in which the sender and receiver of a message share a single, common key, which is used to encrypt and decrypt the message. The data length of a key or message may be chosen to be any of 128, 192, or 256 bits. The AES encryption algorithm has fore main operations: AddRoundKey, SubBytes, ShiftRows, and MixColumns. The operations in decryption are basically the inverse of the operations in encryption [2]. AddRoundKey, ShiftRows, MixColumns are linear operations, and SubBytes is nonlinear. Total AES has 10 rounds each encryption, just repeat the four operations: AddRoundKey, SubBytes, ShiftRows, MixColumns. Encryption key and decryption key are the same one, in encryption procession, key and plaintext were input together to the algorithms, and then ciphertext were the output, which is showed in Figure 2. Although AES has not been commonly used in IC smart card (encryption chip), it is believed security enough for various applications. The operation of subbytes is nonlinear, so in some applications, when the algorithms codes need to be modified for some reasons, the operation must be redefined as linear to make sure the algorithms product the real result.

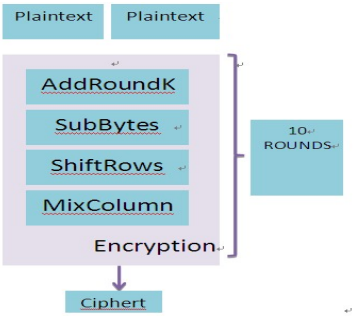


Fig. 2. Procession of AES

5 Data Encryption Algorithm

DES is another symmetric-key algorithm, which means the algorithms use the same key in the procession of encryption and decryption. The procession of decryption just use the reverse key (of encryption) to make the same calculation of encryption.

A. Introduction of DES

The DES (Data Encryption Standard) algorithm is the most widely used encryption algorithm in the world. DES was firstly created by IBM early in the 1970s, and adopted by the USA. But latter EFF proposed a kind of device, which is able to crack DES easily, and EFF gave a detail of the making of that device, thus DES was thought to be not security enough [3].So 3DES was created, actually it was just the triple DES, using the same DES algorithm and 3 different keys to encrypt 3 times. Like that, the procession highly promote the security.3DES was now generally used in almost all sorts of industries [4]. 3DES is especially designed for hardware realization.

B. Details of DES

DES have 64bit key.DES have 16 rounds, each round have several operations. The algorithm divide plaintext to 64bit data, 64 bit data and 64 bit key input and 64bit ciphertext output, after 16 rounds calculation. Details as follsows:

a) Initial permutation: after a initial permutation, 64 bit data was divide to two 32bit data: R0, L0.

*b) First round:*64bit key R0, and L0 were input, and R1, L1 out put. R0 assigned to L1 directly. After a F function R0 and key result a 32 bit data, which was made a calculation of XOR with L0, then output R1

c) 2-16rounds: Act as first round. After 16th round R16, L16 were output.

d) Final permutation: R16, L16 connected together as 64 bit ciphertext.

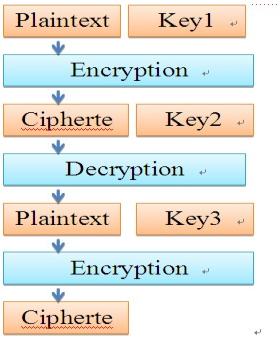


Fig. 3. Procession of 3DES

DES has 64 bit key, but only 56bit was availability, in each round 48 bit was created from 56 bit. Verilog code of F function are as follows:

```
Module F (P, R, K_sub);
Output [1:32]      P;
Input   [1:32]      R;
input   [1:48]      K_sub;
Wire    [1:48] E;
Wire    [1:48] X;
Wire    [1:32] S;
assign E[1:48] = {R[32], R[1], R[2], R[3], R[4], R[5], R[4], R[5],
                  R [6], R [7], R [8], R [9], R [8], R [9], R [10], R [11],
                  R [12], R [13], R [12], R [13], R [14], R [15], R [16],
                  R [17], R [16], R [17], R [18], R [19], R [20], R [21],
                  R [20], R [21], R [22], R [23], R [24], R [25], R [24],
                  R [25], R [26], R [27], R [28], R [29], R [28], R [29],
                  R [30], R [31], R [32], R [1]};
assign X = E ^ K_sub;
sbox1 sbox1( .addr(X[01:06]), .dout(S[01:04]) );
sbox2 sbox2( .addr(X[07:12]), .dout(S[05:08]) );
sbox3 sbox3( .addr(X[13:18]), .dout(S[09:12]) );
sbox4 sbox4( .addr(X[19:24]), .dout(S[13:16]) );
sbox5 sbox5( .addr(X[25:30]), .dout(S[17:20]) );
sbox6 sbox6( .addr(X[31:36]), .dout(S[21:24]) );
sbox7 sbox7( .addr(X[37:42]), .dout(S[25:28]) );
sbox8 sbox8( .addr(X[43:48]), .dout(S[29:32]) );
assign P[1:32] = {S[16], S[7], S[20], S[21], S[29], S[12], S[28],
                  S[17], S[1], S[15], S[23], S[26], S[5], S[18],
                  S[31], S[10], S[2], S[8], S[24], S[14], S[32],
                  S[27], S[3], S[9], S[19], S[13], S[30], S[6],
                  S[22], S[11], S[4], S[25]};
endmodule
```

C. Details of 3DES

3DES was just the triple DES, using the same DES algorithm and 3 different keys to encrypt 3 times. Just use 3 keys to make a calculation of encryption, decryption, and encryption, Figure 3.

6 Attack and Resistance of Encrypt Hardware Device

When encryption algorithms realized in hardware way, especially in the form of IC, it must be protected by special method. Encryption algorithms were security enough from

the perspective of math, but when realized in hardware, it may possess some electric characteristics, just like electromagnetic or power, which may be used by illegal users to attack the circuit [5]. $$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \quad (1)$$ Cov(X, Y) is the covariance. For a attack, electromagnetic and power information must be acquired during the calculation of algorithm of some plaintext, then intermediate data can be calculated by attackers using plaintext and guess key, then calculate the correlation of the intermediate data and acquired electric information, the guess key which is related to the biggest correlation would be the real key. Equation (1) is correlation, during an attack, illegal users will run encrypt some data using the circuit, electric information such as electromagnetic and power of some critic time must be recorded. At the same time guess keys together with data (plaintext) will be input to the algorithms, to output related intermediate data, as the algorithm is known to public. Then correlations of an intermediate data and acquired electric information will be calculated, the guess key which is related to the biggest correlation is the real key. Some of the common resistance solutions include "noisy", "mask", the solution can be implemented in codes or circuit. All of these methods either cut the connection between intermediate data and electric information, or between intermediate data and guess key. Some other way is from the angle of special circuit, such as "asynchronous circuit" and "dual rail".

7 Conclusion

Information security is a more and more important problem now, as various data transmitted through the cloud database and terminals through the internet must be protected. Encryption algorithms are critical in the procession. But the algorithms need to be realized in hardware or software way. So in this paper we made some researches about some commonly used international algorithms, such as AES and DES, we give a detail description about DES and 3DES, so as codes of one of modules in DES (which is called F function), then we point out that Encryption algorithms were security enough from the perspective of math, but when realized in hardware, it may possess some electric characteristics, just like electromagnetic or power, which may be used by illegal users to attack the circuit. Some technologies such as "noisy", "mask", "asynchronous circuit" and "dual rail" are common resistance methods.

References

[1] Messerges, T.S., Dabbish, E.A., Sloan, R.H.: Examining Smart-Card Computers 51(5), 541–552 (2002)

[2] Trichina, E., De Seta, D., Germani, L.: Simplified Adaptive Multiplicative Masking for AES. In: Kaliski Jr., B.S., Koç, Ç.K., Paar, C. (eds.) CHES 2002. LNCS, vol. 2523, pp. 187–197. Springer, Heidelberg (2003)

[3] Eberle, H.: A High-speed DES Implementa-tion for Network Applications. Technical Re-report, Digital Equipment Corporation Systems Research Center (September 1992)
[4] Hoornaert, F., Goubert, J., Desmedt, Y.G.: Efficient Hardware Implementation of the DES. In: Blakely, G.R., Chaum, D. (eds.) CRYPTO 1984. LNCS, vol. 196, pp. 147–173. Springer, Heidelberg (1985)
[5] Agrawal, D., Archambeault, B., Rao, J.R., Rohatgi, P.: The EM Side Channel(s): Attacks and Assessment Methodologies. In: Kaliski Jr., B.S., Koç, Ç.K., Paar, C. (eds.) CHES 2002. LNCS, vol. 2523, pp. 29–45. Springer, Heidelberg (2003), http://www.research.ibm.com.intsec/emf-paper.ps

The Data Mining Technology in the Application of Graduates' Employment

Hu Yan

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
jxnchuyan@126.com

Abstract. The guidance work of graduates' employment plays a very important role in university, college students, and the society. The current employment guidance mode is formed during the distribution system of the past graduates reformation. It has some limitations. Through the association rules application of the computer data mining technology, this paper analyzes the graduates' employment situation in university education. Moreover, it provides references of employment survey technology for the university employment, and the education model improvement.

Keywords: Data mining technology, Graduate, Employment survey.

1 The Overview of Data Mining Technology

With the rapid development of network technology and popularization of computer, various aspects of information have poured into our work and daily life. What way can effectively get the needed meaningful information knowledge from these large amounts of data information is increasingly puzzling us. Moreover, the research work of data mining technology is partly and effectively solved the knowledge scientific management and application. It is the data mining technology and called the knowledge discovery in the database. It refers to how to get the useful information knowledge that people are interested in. Further moreover, the knowledge is implied, unknown and has the potential value. It is include rules, laws, concepts, and models and so on [1].

1.1 The Importance of Data Mining Technology

In recent years, the data mining technology is widely used in the database field. One of the important reasons is the research of association rule mining. It is means how to dig out the useful information from many data. The type of data mining technology can divide into aggregation, time series, association rules, and valuations and so on. Moreover, the whole process of data mining is as the following fig.1 [2].

In recent years, due to the increasingly enrollment, the severity of employment situation is more and more outstanding. Facing the huge, in frequent dynamic change that requires complicated precise management of employment information database, the traditional management model of employment information is getting more and

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 147–151.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

more difficult to use. At the same time, the application of the data mining technology is combined with the modern computer technology and scientific modern management mode. In market research situation of college graduates employment, it is expected to creatively break through the thinking and technology of information management mode for the traditional college students' employment. Moreover, there will be more effective to the management of the increasing employment information database [3]. In short, the application of data mining technology is very important to the system management of university employment information. It cannot only statistic students' employment information in timely and accurately, but also can analyze and excavate the situation of the current employment. It can still evaluations and plans for all the majors, provide decision support for the following students' recruitment, talents training and management. Thereby, it can better combine the education of the students with social requirements.

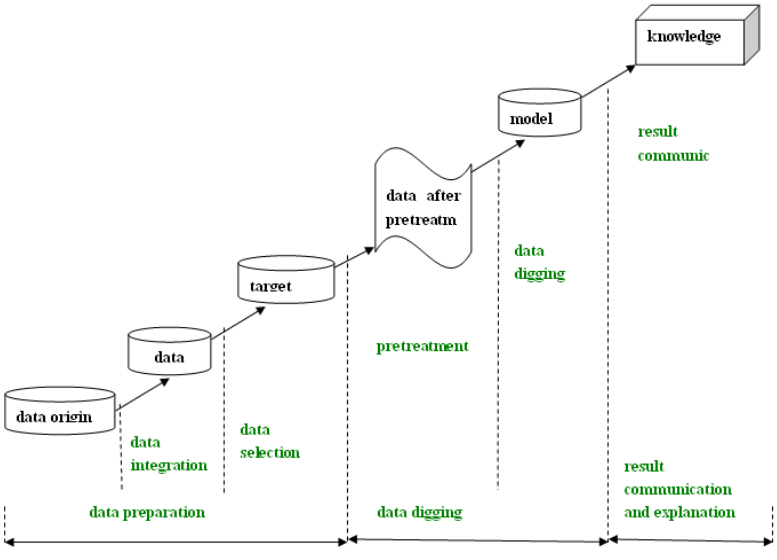


Fig. 1. The whole process schemes of data mining

1.2 The Point Analysis of Data Mining Technology

Such as the analysis of association rules, the purpose is to find out the hidden correlation network of the database. Among them, the correlation is regularity among two or more than two variable dereferences. And there are mainly three association regularity. They are temporal association, and simple correlation. The attribute number in association rules, it can divide into one-dimensional association rules and l association rules. The correlation database is combining with a group of internal related data. The table collection form the correlation database, each table should

have a group of column or field properties (generally store the large tuple). However, the data warehouse is generated by a series of data processing such as embeding, cleaning, transformation, integration, and regular refresh and so on. In actually, the physical structure can be multidimensional data cube.

2 The Application of the Association Rules Analysis in Employment Survey System

In the data mining, the general procedure of correlation rules evaluation is: the determination of problem task solving, choose the right data mining technology (such as statistical methods, machine learning method and database method, etc) and tools. It includes the preparation of the data, and the establishment of the model and the work of knowledge discovery. Moreover, it includes the finally evaluation of the result and assimilation of the knowledge. The analysis is as follows:

2.1 Clear about Question Tasks and Building System

Based on the application of data mining technology in the aspects of some research work, I will divided the complete employment analysis system into two parts: the database treatment (include data query, read, statistics, updating and the simple system interface of the front desk) and the algorithm implementation of background data mining. In there, in order to improve the convenience of database, we should use the Microsoft Access 2003 database platform. The whole system includes data mining algorithm of Microsoft proof reading. The concise change system is with the algorithm that is the algorithm of multidimensional association rules. From the above, we can get a simple system model figure, as the following Fig.2 shows.

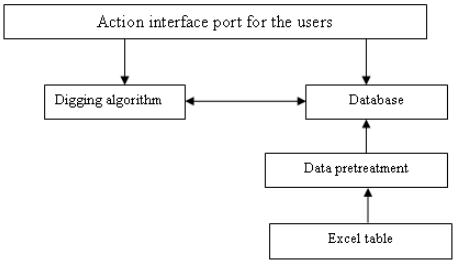


Fig. 2. Simple system model schema

In the implementation process, the main program and function of the system are as follows: pretreatment data, searching dimensional frequent predication, calculating the multidimensional frequent item sets. Moreover, produce strong association rules and correlation analysis.

2.2 The Analysis of Data Pretreatment

This is the important prerequisite for data mining steps. The education system storage the students' information is the self-characteristic. It means, in generally it does not involve multiple database data integration and transformation. The noise is small, data is complete, and time sensitivity is not high with less vacancy value. There has little vacancy value, higher credibility of data value, and the data has the digging value within a few years. Overall, the characteristics of the information on the employment of college graduates are easy to deal with the data source. Among them, to deal with the student employment information data preprocessing work is to generate students' information overall table. Because when the school preserves student information, every student has a unique identification. In generally, the correlation database table is formed with various student id record and integration. Therefore, in order to get a transaction database table with multidimensional attributes, each transaction in a database table article should through the primary key "student id" that can match the link together. However, before the data mining, some relevant original data processing work should be done. Such as in the data preparation, in order to identify the students in the convenient way, we can build "party member" and "gender" as the Boolean index. Moreover, we can use "national", "major", "background", "birth place", and "employing unit property" as the enumeration index. Then we can use the quantity index to do the series of number work with "flexible employment rate", "employment rate" and the "contract ratio". In a word, we should delete the entire unrelated or redundancy attribute in order to reduce the cost. And we can use Null to fill the empty value to avoid the digging data program will take Null for the interested theory. Moreover, we should save some "number" information for the convenient cluster analysis of evaluated students. We can base on the required indexes to coding them. Such as the graduate background is "junior college education, upgrade from junior college student to university student, undergraduate course and graduate student", we can coding them in 1, 2, 3,4. We do it in this way, and this can use on other index coding.

2.3 The Application of Data Mining Method in Data Classification Processing

Considering the great difference between the past and the present employment situation, we suggest during the cluster analysis of graduation data, we can generally choose information in the past five, six years. Because it is more scientific and reasonable, so as we can better ensure the useful data information is representative. In this, we are mainly clustering analyze of the graduates groups. The procedure is roughly about: through evaluation of the undergraduate students' information and the students' id management basement, we build fuzzy similarity matrix. After that is following the major attribute to do the classification. Such as we do the cluster management of the maximal tree, and then based on the detailed condition, we step into the digging evaluation stage. It means we choose one group of typical data to do the digging.

2.4 Start Algorithm and Find Out the Relevant Rules

In this discussion, use the algorithm to find the relevant Apriori rules. Because most association rules are based on support degree and confidence of the framework, so in some frequent predicate, we can produce strong confidence to meet the minimum association rules. Before the mining, because the mining association rules of each attribute domain is in the Boolean type, so it is necessary to enumerate the index and quantitative index value. Such as in the major that is similar to "economics, management, and other", the related index dereference of each student is same with the professional management process. Then it will turn into the Boolean value.

2.5 Compare Evaluation of Digging Results

Through the above multidimensional algorithm for mining association rules, get all the frequent predicate set, and treat it as a series of strong association rules. We can use the concept theory of data mining technology. Make relevant analysis for the strong association rules. For each multidimensional association rules, there needs caring about some of the attribute dimension. Treat it as to the conclusion of the output regulation. Such as through list education background, gender, professional, and a few association rules we can compare the digging results. Then, we can find out the key elements that can influence on several key factors of employment.

3 Conclusion

To sum up, with the growing scale of recruit students in our country college, the graduates increased year by year. That makes the student employment problem is becoming more and more serious. Through the application analysis of multidimensional association rules in the data mining technology, we can effectively solve the graduate employment data and discrete analysis methods and the data management features. Moreover, and from these a large number of employment data extraction or find knowledge, these knowledge not only suitable for our market survey and situation analysis of college graduates employment, it can also provide decision support that assist for the daily teaching management. Moreover, it has really and profound meaning for the students training and employment.

References

[1] Fan, M.: Data mining concept and technology. Mechanical Industry Press, Beijing (2001)
[2] Peng, Y.: Data mining summarization. The Journal of Dehong Teachers College 1(18) (2009)
[3] Tian, G.: The thinking of using information network technology to promote universities employment instruction. Chinese Science and Education (education scientific research) (18) (2011)

Application Research of Virtual Reality Technology in Electronic Technique Teaching

Hu Liang[1] and Bi Xiaoming[2]

[1] Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
[2] Wenzhou Vocational College of Science and Technology, Wenzhou, Zhejiang, 325006
hliang1972@126.com, wzkzybxm@126.com

Abstract. At present, teaching virtual reality as a new media has penetrated to the area of educational technology. It is playing an important role in advancing the method reform of higher education in colleges and universities. Moreover, it improves the teaching quality. The application of virtual reality technology in teaching of electronic technology provides students with a learning platform. Students can practice through hands-on interactions so that their ability of practice, observation, innovation capacities get fully exercises. This will help students acquire vocational skills, enhancing electronic technology "education" efficiency and "learning" efficiency. This paper briefly describes the virtual reality technology and its basic characteristics as well as on the application of virtual reality technology in teaching of electronic technology and the importance of application-specific studies.

Keywords: Virtual reality technology, Electronic technology teaching, Virtual instrument, Virtual electronic lab.

1 Virtual Reality Technology and Its Basic Characteristics

Virtual reality technology is a high technology system that combines with stimulation technology, display technology, sensor technology, network technology and computer graphics technologies and artificial functions. Through visual, auditory sense, tactile effects to users, the users can produce immersive interactive scene simulation. Virtual reality technology use high-resolution display, more sensing interaction, and three-dimensional graphics to generate computer technology that can create out three-dimensional virtual environment. The user, who simply relies on keyboard, and mouse or some entered devices such as wearing dedicated data gloves, helmet, and other special devices, will enter the virtual space, percept, and operate all objects in the virtual environment. Then achieve real-time interactive, gets real feel of object operation that similar with the reality environment. Virtual reality technology emphasized the leading role in the virtual reality world, and has the following three characteristics:

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 153–159.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

1.1 Illusion of Immersion

Key features of virtual reality technology is to make users feel themselves are in part of the virtual reality environment. At the same time, you can get a variety of perceptions, such as visual perception, haptic perception, auditory perception, physical sensations, motion perception, olfactory perception, force sense perception, and so on. All these enable users to produce immersive feel.

1.2 Interactivity

Virtual reality system achieves the natural interactive between human and virtual environments, such as the hand movement, head rotation, people move around, and so on. At the same time with the help of special hardware devices, we can enable the virtual reality system rapidly respond the user input information. Allow users to have the sense of the real human-computer interaction.

1.3 Imagination

Virtual reality helps users break through time and space limitations, immersed in the virtual world to obtain new knowledge, new understanding, thereby increasing user rational cognitive and perceptual awareness, stimulate the generation of new ideas.

2 Application Importance of Virtual Reality Technology in Electronic Technology Teaching

2.1 Beneficial to Encourage Students' Learning Interest

Use virtual reality technology to create virtual environments so that students can learn in a realistic virtual environment. Get rid of the traditional book teaching of dullness and abstract sense, and make up for the limitations of multimedia demonstration teaching to enable students to practice the theoretical knowledge in a virtual environment. Moreover, organically combine knowledge and professional ability to inspire student interest in learning.

2.2 Beneficial to Save Resources for Schools

In traditional electronic experiments, students complete an electronic experiment must use a variety of instruments, and the different electronic experiment need variety of instruments. In this case, to complete electronic specialized courses there need to purchase large amount and type of instrument. Moreover, not only occupied the school significant liquidity, it is also need to give these instruments sufficient space. In addition, the instruments interconnections are complicated with lines bad contact. This brings the resource wastage. For the application of virtual reality

technology virtual instrument platform, we can implement a general-purpose computer with multi-purpose features and a multi-purpose room with laboratory functions. Saves a tremendous amount of equipment funding for schools, and solves the problem of school laboratories limitation, improve efficiently usage of school resources.

2.3 Beneficial to Train the Innovation Ability

The advantage of virtual reality technology in electronic technology teaching is providing openness room environment for students. The student can independently build and manipulate objects in the virtual environment, and clear, intuitive see the results of the operation on the computer screen. For example, creation of the virtual electronic laboratory based on electronics workbench software, gives students all the components that required. Students can design a variety of analog circuits and digital circuits. At the same time, through virtual instrumentation to detect and determine the circuit feasibility, improve the efficiency of experimental teaching. All these are beneficial to students ' autonomous exploration and train their creative ability.

2.4 Beneficial to Improve Experiment Efficiency

Create virtual instrument based on virtual reality technology with computer hardware as the core. Possess storage, display, analysis, transmission, processing, and variety of control functions. For students' electronic experiments, it is much more practical than the traditional instruments. Because virtual instrument interface use the keyboard, mouse, and other input devices to operate away from the failure possibility of mechanical traditional equipment. Determine the virtual experiment of high speed and precision measurement, which can minimizes the measurement error and increase the effectiveness of student electronic experiments.

3 Detailed Application of Virtual Reality Technology during Electronic Technology Teaching

3.1 Self-simulation Laboratory

Application of virtual reality technology can create all kinds of laboratories for students so that students can do the experiments in the virtual laboratory. Circuits, power supplies, and other aspects of simulation experiment are the example. At the same time, allow students assemble an electronic device in the virtual environment. Virtual reality technology can make the most of its courses and training in a virtual lab and operations. Most of the e-skills can be proficient in virtual training workshop exercises. This will help schools to save a lot of money for the expensive experimental training equipment. Simulation experiment of self-design process as shown is Fig.1.

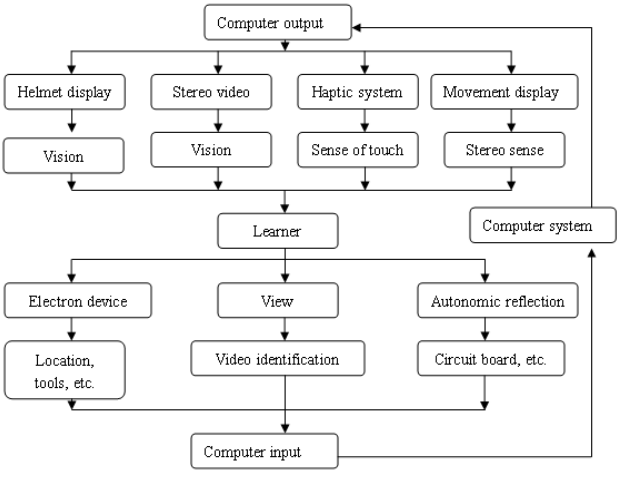


Fig. 1. Application of virtual reality technology in electronic technology teaching

3.2 Virtual Instrumentation

Virtual instrumentation refers to the use of computer systems, hardware and software, and develop common virtual instruments and meters. Use LabVIEW software to develop virtual instrumentation include ammeter, oscilloscopes, voltmeters, signal generator, frequency counter, etc. Students can follow the test requirement and define function of the instrument. At the same time, output the different test results and do the real-time analysis of inspection. For example, use virtual function generator of LabVIEW software and use mouse and keyboard to set up its amplitude, frequency, and wave. The functions are fully in line with the conventional instruments. Virtual voltmeter based on LabVIEW software has more clearly wave than the traditional. Students can adjust the input value and the input waveform arbitrarily in the graphical user interface. That can give students a more intuitive comparison. In addition to covering conventional functions, the virtual oscilloscope based on LabVIEW software has the feature of display, storage, recording, print waveform, and other functions. Through the virtual storage and "playback" function of oscilloscope, the curve can be

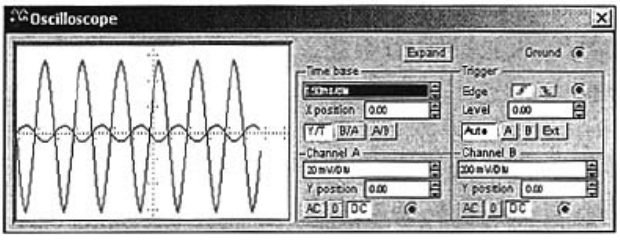


Fig. 2. Virtual Oscilloscope

displayed on the screen, and playback speed can be freely adjusted. Students can observe the waveform changes clearly and intuitively. Moreover, evaluate experimental results in real-time in order to improve classroom experiment teaching efficiency. Show in Fig.2.

Based on electronic bench (Electronics Workbench) technology to build the virtual oscilloscope, adds Panel extensions except all the real oscilloscope features. Virtual oscilloscope can use more efficient. Fig.3 is the detail:

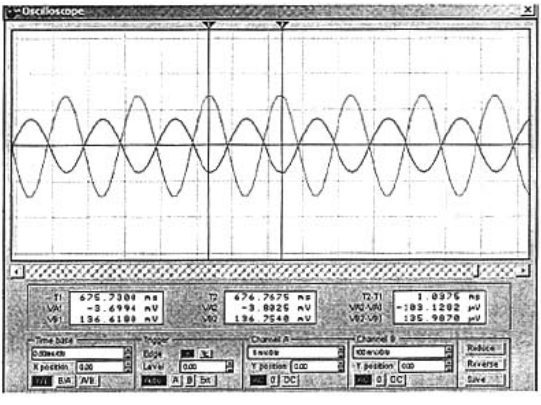


Fig. 3. Virtual Oscilloscope of panel expanding

3.3 Virtual Electronic Laboratory

At present, some colleges and universities already have virtual electronic experiment during the teaching. There mostly use virtual electronic experiment software of Electronics Workbench. Workbench creates and test based on computers. Moreover, test on the computer platform with real-world laboratory experimental experience.

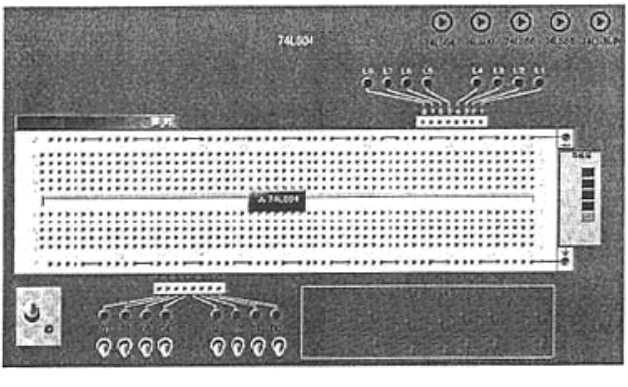


Fig. 4. Simulation breadboard of digital circuit

In the virtual reality environment that created by Workbench software, we can design and tested a variety of electronic circuit such as analog circuit, high frequency circuit, electrical engineering, digital circuit and so on. For example, the plug experiment of digital circuit breadboard can plug wiring and test the function to 74LS series integrated chip in the virtual environment. Fig.4 is the detail:

3.4 Analogue Simulation System of Skill Training

Use virtual reality technology to design and create electronic skills and training simulation system, electrician skills and training simulation systems. Skills training simulation system of multimedia simulation use sophisticated technologies and apply it to the teaching of electronic technology. This can fully reflect the teaching philosophy of theoretical knowledge and practical vocational education. The oscilloscope, for example, the simulation system of skill training emphasizes students' interactive, circumstances and realism. During the training, students can feel the reality in the simulation environment. This is same with the real oscilloscope operation. It is mainly because the procedure in simulation system of oscilloscope panel, the display window control buttons and knobs of the basic measurement procedures are lifelike in front the students. It embodies the basic feature of immersion in virtual reality. The skill in training simulation system of oscilloscope has more sense of reality than the oscilloscope virtual instrument.

4 Conclusion

Overall, the virtual reality technology has the positive impact to improve the modern education technology, ease school pressure and improve teaching activities. The application of virtual reality technology in electronic technology teaching is the important ways to reform the mode of traditional electronic experiments. It makes up for the inadequacy of the traditional electronic experiment teaching and optimizes electronic experimental teaching. It will help to develop students ' practical skills, observation skills and innovation ability then to improve students ' professional skills.

Acknowledgement. This article obtains the support of Vocational colleges research of "double qualified" teachers team construction from the planning issue of 12th Five Year Plan in Jiangxi science of education(No. 11YB114). This paper is one of the research results.

References

[1] Huang, Z.: Application practice of modern education technology in the electronic technology courses. Culture and Education (24) (2006)
[2] Guo, L.: The application of virtual reality technology during the communication experimental learning. Dalian Maritime University (2007)

[3] Dai, H.: The application of virtual reality technology during the medium vocational electronic education (36) (2008)
[4] Zhang, C., Shan, C., Wang, X.: The application of virtual reality technology during the circuit education. In: China Twentieth Academic Article Circuits and Systems Conference and 2007 Electronic Information Symposium of Hong Kong, Macao and Mainland (2007)
[5] Li, H.: Analysis of virtual reality technology and the application research in college teaching. Education of Computer (6) (2009)
[6] Long, H.: Application research of virtual reality technology during the electronic vocational education. China Educational Technology & Equipment (27) (2011)

The Platform Design of Practice Teaching in the Electromechanical Major That Based on the Virtual Reality Technique

Hu Liang, Yang Wen, and Huang Chengzhong

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
hliang1972@126.com, ncyangwen@126.com, nchuangchengzhong@126.com

Abstract. This article starts form the virtual reality technique and simply describes its characteristic. It researches the technology principle and theory model of virtual lab that based on the virtual reality. Moreover, this article provides the frame design and application of practice education platform in the electromechanical major, which is combined with the subsystem of virtual reality technique, model subsystem and virtual instrument subsystem and based on the virtual reality technique. The result shows through this system can effectively increase the practice teaching effect of electromechanical major.

Keywords: virtual reality technique, electromechanical, practice system.

1 Important Significant of Professional Training Teaching for Mechanical and Electrical Integration

The core-training target of higher vocational school provides production, establishment, and management of comprehensive technical talents with high quality and ability that can satisfy the enterprises requirement. This target decides the basic characteristic of aiming at training education for developing the professional technical talents. Moreover, this is also determining the important position of training teaching during the personnel training process. Training teaching has the important significance for the increasing of practical ability, manipulative ability, innovation ability, responsibility sense, and even the professional dedication. Training teaching experiences also establish the development basement for the students' further career. Mechanical and electrical integration is one of the significant majors in the higher vocational school. In the school, teachers train the professional theory and comprehensive professional ability of mechanical and electrical integration technology. The school is mainly work at training the comprehensive talents with the first level technology quality and ability of equipment application, management, development, marketing and technical service.

The core-training target of higher vocational school provides production, establishment, and management of comprehensive technical talents with high quality and ability that can satisfy the enterprises requirement. This target decides the basic

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 161–165.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

characteristic of aiming at training education for developing the professional technical talents. Moreover, this is also determining the important position of training teaching during the personnel training process. Training teaching has the important significance for the increasing of practical ability, manipulative ability, innovation ability, responsibility sense, and even the professional dedication. Training teaching experiences also establish the development basement for the students' further career. Mechanical and electrical integration is one of the significant majors in the higher vocational school. In the school, teachers train the professional theory and comprehensive professional ability of mechanical and electrical integration technology. The school is mainly work at training the comprehensive talents with the first level technology quality and ability of equipment application, management, development, marketing and technical service.

The mechanical and electrical integration includes hydraumatic, sensor, network, PLC, communication and pneumatic. They interpenetrate with each other. The technologies are using widely in industry, energy, transportation, national defense and other fields. It is the important combination in the modern industry technology filed. Numerical control machine, flexible manufacturing system of MPS, circuit board packaging, robot system, and other supermatic equipments are using widely in our Yangtze River Delta and industry developing areas. In this way, grasps the theory knowledge and professional ability of mechanical and electrical integration is under the urgent need. Moreover, it is also imperative to become the talents with the ability of operation, maintenance, test, management even marketing and develop the equipment.

The higher vocational school needs to pay attention to the training teaching in order to suit for the normal development of high and new technical industry, the region economy, and satisfy the technical talents' requirement. Moreover, training teaching is the important method for training students the comprehensive professional ability. Research the mechanical and electrical integration as the development of training teaching system, establish and perfect the suitable system for the high quality talents will has great significance for the society requirements of professional ability talents.

2 Establishment of Training Teaching System for the Mechanical and Electrical Integration

Higher vocational education has obviously professional education that can train the students' professional ability. The major of mechanical and electrical integration has particularly prominent requirements of training students' comprehensive abilities. In this condition, the course system of this professional major should targets with training the comprehensive profession quality and ability. Theory teaching and training teaching is the big support to achieve this target. During the teaching system establishment, there needs following the basic rules of objects development, and complies the basic principle from the easier to harder, from the shallower to the deeper and from the single to the comprehensive. In order to reach the steadily application of teaching target, we need to ensure the theory teaching content, skill certification teaching and training teaching contents are coordinate with each other

and combine together. Based on this principle, the writer divides the content system into basic ability training, professional ability training, and comprehensive training. The structure is the echelonment.

2.1 Basic Skill of Training Teaching

Basic ability training is the most basic stage in the training teaching system of mechanical and electrical integration. Related with basic theory teaching, this phase is aiming at metalworking practice, experiment of electrician and electron, CAD design and mapping training, turning lathe training and so on. The training teaching target of this stage is practicing the students to familiar with basic experiment equipments. Moreover, train them to possess the basic understanding of tool software and accumulate the basic usage experiences. Through the experiments and practice, they can understand the CAD and other graph software protraction, basic data measurement, collect and grasp the basic experiment methods. Students can be familiar with the primary metalworking of pliers, milling, turning, digging, and soldering. From the beginning, the school pays energy to train the practice ability, innovation ability, and adaptability for the students.

2.2 Professional Ability of Training Teaching

Professional ability of training teaching is training students the abilities of technology knowledge application, practical operation, actual question settlement, and other key parts. In this stage, students should be familiar with professional ability and obtain the related professional qualifications in table 1. It is the target of training teaching. In the detailed training, the school should direction with training the comprehensive application ability, and aiming at the specific major direction that students collected. Moreover, training teaching is the details that follow the process of "basic ability training→simulation operation training→actual operation training→expanding training→ identification of profession vocational qualification". Through the acquaintance and punning of professional ability, it is laying a solid foundation for the important comprehensive ability training.

2.3 Comprehensive Ability of Training Teaching

Comprehensive ability training teaching is the final target of basic ability training and professional ability training. Moreover, it is the key part of achieving comprehensive professional talents' training of quality and ability. Comprehensive ability training teaching includes production training, undergraduate training, and position practice. This training teaching target is making students to form the direct acknowledge of working environment, further practice and students adaptability improvement and working ability in the social environment and working position. It can short the employment adaptable time. Train the professional quality and team cooperation for the students is the main target of comprehensive ability training. The training contents in this stage is basing on applicability of the national professional classification and professional qualification. The training requirement and assessment criteria needs to link up with the related professional qualification standard. Strengthen the training system availability, pertinency, and continuity.

Follow the national and social qualification and requirements to train the students' comprehensive professional ability and quality. Combine the teaching, learning and practicing in the higher vocational teaching. Otherwise, objectively promote the double certificates of professional qualification and academic certificate, ensure the students can smoothly graduate, obtain employment, and the enterprise can find the required talents.

3 Professional Training Teaching Management and Protection Measurements for Mechanical and Electrical Integration

The training teaching of mechanical and electrical integration major not only needs the basic teaching system frame, but also needs each necessary condition that manages and determines the system effective operation.

3.1 Form the Scientific Training Plan and Exam Approach that Correspond with Actual Work

In each stage training teaching, training plan formulation needs to center with practical and scientific. The reasonable training plan that only corresponds with the practical professional background can really reach the target of training the professional talents that suit for the social requirements. During the plan formulating, there needs fully consider about the teaching facilities, students qualities, period arrangement, and various courses' articulation. Training teaching should first ensure the correspondence of practical work. The evaluation standard should use industry standard and regulation. Only the practical exam standard can reflect students' progress and shortage. For the detailed examine methods, we should use the combination of specification evaluation and process evaluation. The specification evaluation is aiming at experimental training. Process evaluation is aiming at practical training. The examine method will under perfection and optimization during the practice.

3.2 Training Teaching Mode of Creating the Work Environment That Close to the Actual Environment

The training teaching mode that closed to the real working environment will help students for improving adaptation in the working environment. Just like using the ability competition, we can group the students, arrange the training tasks, and limit the time. Because the time and work load limitation, each team has the competition relation and form the working competition mode in naturally. This competition mode formulates the actual working process. Moreover, it provides the teamwork for each group and the coordination is suit for the practical working environment. This kind of training teaching can practice the professional ability of mechanical and electrical integration, and make students to realize the significance of teamwork.

3.3 Build Training Basement That Has Advanced Equipments and Correspond with the Present Condition

In order to reach the target of training students' the comprehensive professional ability, there needs to support the really or the training environment that close to real. For the environment, the school should establish a training basement that has advanced technical equipments. Moreover, it needs to correspond with the present industry condition. For example, the parts sorting station, pneumatic machinery and so on. The equipments need to as true as possible that can determine the students can study and practice under the real work environment. Complete equipments and "real" training basement is the basic support to train professional technical talents.

3.4 Strengthen the Teacher Group

The qualified teacher is the necessary condition for training teaching. Based on the requirement of theory education and training teaching, the teacher of mechanical and electrical integration major should have the related abundant theory knowledge, practical experiences, and professional quality. They can provide the knowledge to the students, and need to do the samples and guide students during the production, training sites. Moreover, they need to familiar with the industry regulation, occupation standard and job requirements. The major of mechanical and electrical integration is the comprehensive major, so there need to have the mechanical teacher and electrical teacher.

Acknowledgment. This article obtains the support of Vocational colleges research of "double qualified" teachers team construction from the planning issue of 12th Five Year Plan in Jiangxi science of education(No. 11YB114). This paper is one of the research results.

References

[1] Zhao, X.: Mechatronics system. Higher Education Press (2004)
[2] Wang, H., Zhu, Z.: The thinking of mechatronic training teaching. Science and Technical Information (2) (2006)
[3] Cheng, J., Dong, L., Liang, J.: Electrical and electronic engineering revolution and practice in the training teaching for the mechanical and electrical integration in higher vocational school. Science and Technology of West China (13) (2010)
[4] Li, X., Zhou, W., Huang, W.: Teaching mode practice and exploration of ccomprehensive training for the mechanical and electrical integration in higher vocational school. Shananxi Education (02) (2008)
[5] Geng, B.: Discussion of training teaching revolution in higher vocational school. Wen Jiao Zi Liao (15) (2007)

The Exploration Model of New-Type Software and the Application in the System Exploration of Teacher Management

Cao Yu

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
jxnccaoyu@126.com

Abstract. During the software process, the primary work is selecting the software exploration model and software process that suitable for the practical condition. This article describes and summarizes the exploration model of new-type software. Moreover, it discusses the relations, characteristics, and application field among each model. Adapt the model that based on spiral software of spiral model and research the teacher management system for the colleges and universities.

Keywords: Software exploration model, component element, spiral model, teacher management system.

1 Introduction

With the development of education business after reform and opening up, the strategy of taking the “double-qualified” inter-disciplinary talent training as the major in building teachers scale. It largely promoted the further development of the education business. In other words, the significance of building “double-qualified” teachers scale culture not only requires teacher advance with the time and overall innovation, but also requires modern teachers to have a strict self-restraint requirements. That means, in addition to teach responsibility, comprehensive professional knowledge and skills, and skilled practical experience, the higher inter-disciplinary professional quality is also needed. This can comprehensively enhance the national teacher industry standard, and promote the development of education business. Therefore, in the current situation, to strengthen the building of "double-qualified" teachers scale training has great guiding significance and practical value to the improvement of the educational quality in higher vocational colleges of China.

2 Significance of Building "double-qualified" Teachers Scale Training

Largely build and develop inter-disciplinary talent is the major task of higher vocational education. Higher vocational college is an important part in the national education system. The comprehensive development and construction of higher

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 167–172.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

vocational colleges is conducive to the strategy of rejuvenating the country through science and education. Especially in the stage of full access to the well-off society in our country, the higher vocational colleges have the arduous historical task by themselves. Firstly, in the situation of technology and economic coordinately comprehensive development, the need of personnel training improves unceasingly. There is under the urgently requirement of updating the concepts in the national education business for the further improve teaching and learning. Moreover, the teaching content innovation, the teaching system and teaching method improvement is also need. Secondly, in the knowledge economy age, knowledge and technology are primary production capability. This means trend of global competition is becoming increasingly apparent, especially the market competition increases strongly. Thirdly, in the "Tenth Five-Year" period until now, our country has already profoundly cleared the system of personnel training mode in higher vocational colleges. It means it is enable to coordinate the comprehensive development of social progress. In other words, higher vocational college will be the main part of the national education business. Train the inter-disciplinary talent and increase the training efforts will be the focal point of improving and promoting education business development.

Comprehensive and coordinated development of higher vocational colleges must take "double-qualified" development direction as the foothold. The key of education development is the academic staff. From the very beginning, the scale construction of China's vocational education teachers was using the non-modular and non-professional teaching body. For example, at that time, many vocational colleges would deploy academic staff from middle school, while academic staff in higher vocational colleges are mainly came from regular institutions. This un-regularization and un-vocational teachers switched to this job without training, and constrained the forming of the higher teachers' teaching quality. It can be seen that the professional capabilities of teachers in our higher vocational college is quite limited, which chiefly appeared in lack of working and teaching experiences. For this, based on this situation, in 1990s, the former state education commission proposed the idea of "double-qualified" teachers. This is in order to enhance the practice and professional capabilities of the academic staff, in order to gradually and coordinately promote the adverse current situation in professional competence of teachers, teaching body quality, and teaching body scale limitation.

On the one hand, the development of "double-qualified" teaching body is the key to improve the higher vocational education. Moreover, from the global point of view, the development of vocational education is the trend of the times. Promote the domestic education business attaching great importance to building "double-qualified" teaching body scale should be the inevitable choice. Especially along with decades of reform and opening up, the traditional industries is under the continued improving. The high-tech productivity developed even rapidly and many products replaced increasingly frequent. At the same time, a lot of economic structural adjustment pace continued to accelerate, which caused the service standards of the tertiary industry reached an unprecedented increase. To this end, as an important part of national education business, the higher vocational teachers are bound to accompany the current situation. Moreover, they need to enhance and make a change, while as much as possible to train or create a group of "double-qualified" teachers with strong management capabilities, service level, theoretical knowledge, and

practical ability. It is worth noting that the "double-qualified" teachers are not only with a variety of qualifications of academic staff, but also refer to outstanding practical ability, high-quality inter-disciplinary talents.

On the other hand, the "double-qualified" teachers are the key to achieve vocational and technical education of personnel training objectives. "Double-qualified" teachers should both have good ethics cultivation, teaching abilities, and good industry professional attitude. Moreover, they need knowledge, skills, and practical ability. Besides, hold a "double certificates "is very important. Teach by these professional teachers, the higher vocational college students can not only learn a lot of theoretical knowledge, but also successfully complete the pre-job vocational training during the school time. They will have the ability to work independently and engage in some kind of professional job that in order to adapt to the requirement of practitioners in social talent market in the quality of "applicable "," efficiency", and benefit".

3 Strategies to Build "double-qualified" Teaching Body in Higher Vocational Colleges

Build a “double-qualified” teaching body which can adapt to employment-oriented higher vocational education, and strengthen skill and practice teaching requirements is the key of training high-quality applied talents the community needs. Moreover, it is also the main direction for establishing the teaching body of higher vocational education.

Strengthen the implementation of the community supporting efforts to improve the operability of "double-qualified" policies . Increase the support power, improve, and perfect the evaluation criteria of the "double-qualified" teachers. It can be said that the defining standards of "double-qualified"academic staff in the education sector has many similarities and differences sayings. Overall, some are strong on theory, but poor in practice. Some are strong in practice, but out of line between theory and practice. Therefore, to make "double-qualified" as model of academic staff training policy guidelines, it has to be able to establish "dual qualified" teachers adjudicators organization, and try to promulgate standards review mechanism and feasibility of the approach. Specifically, a number of pilot projects could be developed at the beginning, and implemented progressively. In addition, we know that all sectors of the industry standards have similarities and differences on the nature. Therefore, the set of "double-qualified" teacher evaluation criteria must try to go with the practice initiatives to implement and gradually refine to improve, in order to improve the operability of its implementation. In addition, the construction of the "dual qualified" teaching body is also with high and low points by multiplying levels that due to the actual division. For example, the primary "double-qualified" teachers, mid-level "double-qualified" teachers, and advanced "double-qualified" teachers and other advanced titles. In this way, it will be able to fully promote and supervise the initiative of the academic staff to improve the self-study and self-development, rather than completing the task of teaching while teaching.

Gradually improve the treatment of "double-qualified" academic staff, and provide financial support. Construction with adequate funding, excellent personnel treatment,

and excellent office environment are key premises of "double-qualified" teaching team. In other words, the "double-qualified" teachers have more than one undergraduate or higher education record, and other professional qualification certificates. It is essential for education of personnel training. Especially under the current age of the knowledge-based economy, training cycle of this inter-disciplinary talent is relatively long and difficult. For example, in Germany and other western countries, the general academic staff can get this qualification standard after age of 30. Therefore, based on the role the "double-qualified"teachers play as the mainstay of the national cause of education, government agencies should strengthen their policies support and strengthen support efforts. Meanwhile, the government can also organize beneficial activities such as teachers' forum, seminars, special academic skills display platform. This is enabling to comprehensively improve the overall quality of the “double-qualified” teachers talent, enhance the quality level, doing its best to ensure that a variety of policy support initiatives to reflect the far-sighted, innovative, and practicality. This is making the true sense to provide a broad and favorable exchange platform for “double-qualified”.

Establish an improving and perfect mechanism of “double-qualified” teacher training . Strengthen the enforcement of “double-qualified” talent training as the base construction. For now, the enforcement of construction of training base of academic staff which is similar to this “double-qualified” talent is relatively large, and seems taking shape in the current. Once the "China Youth Daily" reported Tianjin had invested nearly 1.5 billion to the community in construction funds in the last five years, while there appeared a lot of higher level vocational education and training organizations. At the same time, from this education and training institutions there would be more than 100,000 graduates at all levels each year, while trainees were up to 500,000 people. Hence, it indicated that our country really did a lot of useful attempts in varying degrees. Although, the implementation of these initiatives already has a no small gap with “double-qualified” teacher training, through the positive and useful attempt and over time, it must be a large extent of improving the operability of combination of theory and practice in “double-qualified” talent training. In addition, many educators believe that if the ministry of education acts an official capacity and organize relevant units or institutions with necessary qualifications to unite, and to create the plan of “double-qualified” talent, it would further expand formation of “double-qualified” teaching scale. Namely, through formation of national standard high level higher vocational teachers college or other organizations, it will expand the scale of “double-qualified” talent cultivation, and thus benefit the country and people, revitalize national education.

Run “double-qualified” network university. Carrying out specific training to academic staff with online university is one of the favorable measures to cultivate “double-qualified” teachers in the era of information industry technology for modern vocational colleges. To this end, the follow-up and implementation of such initiatives should have a plan, and be put into action by levels and phased. In particular, it could be completed open education institutions leading by government departments in macro (principles and policies, educational content, eligibility criteria), and distinctive in a specific teaching (at any time for admission, free choice of curricula, to voluntarily participate in the qualification), for teachers to learn professional technology as “skills”, “practical”, “lifelong” and “scientific” to create conditions.

Plans and regular training.Combine the demand of this professional course in our college and make a training program of "double qualified"teacher's course. Employ training experts to coach these future "double qualified" talent on psychology and pedagogy. At the same time, activities or training issues which involve with professional skills and actual operation could combine lectures and extracurricular academic study to give Guidance suited to local conditions. And then, the teachers will understand the specific approach in practice, to raise the academic standards of combination of theory and practice. For example, Culture methods of training academic staff in college will through the professional teaching experiment to study and practice. In this way, teachers will try to figure out the theory and coupled with the operation and practice of the experimental course is able to continuously enrich the teaching experience. In addition, in order to do this work better, we will break the previous system of the teaching and practice and implement the combination of theory and practice to achieve complementary advantages. Moreover, the long-term practice teaching staff could select a number of other professional courses to study, and then based on the theoretical study constantly enrich their own practical experience, to ensure the quality and functions of "double qualified "teacher completely play.

4 Conclusion

In short, the vocational college in the stage of training personnel should pay great attention on teachers' "lifelong learning", "collective learning" and "full education". At the same time, transfer the basis of theory into the creativity in the process of practice. For this purpose, the vocational college should actively do the things as follows: First, maintain continuous learning, based on learning professional knowledge, strive to combine theory and practice, aimed at improving the professionalism. Second, in the process of collective learning, in addition to the educational administration staffs need to arouse learning, decision-makers, managers and primary staff also need to value learning. Finally, if we want to transfer learning into creativity not only need to learn plenty of knowledge and get more information, but also need to learn how to use the modern intelligence, improve the efficiency of learning and teaching, and finally make the process of learning into the process of working and self-improving, while the "double-qualified" teachers will gradually growing in learning and working.

Acknowledgement. This article obtains the support of Vocational colleges research of "double qualified" teachers team construction from the planning issue of 12th Five Year Plan in Jiangxi science of education(No. 11YB114). This paper is one of the research results.

References

[1] Sun, Z.: The exploration of higher vocational colleges' "double-qualified" teachers training mechanism. China Higher Education Evaluation (01) (2009)

[2] Ye, M.: The exploration of higher vocational colleges' "double-qualified" teachers training mechanism. Continuing Education (10) (2009)

[3] Zhan, Y.: The exploration and analysis of higher vocational colleges' "double qualified"teachers training mechanism. China Electric Power Education (10) (2010)
[4] Xu, R.: The strategy research of making higher vocational college "double type" teachers. Education in Jiangsu (18) (2010)
[5] Sun, Z., Ye, M.: The exploration of higher vocational colleges' "double-qualified" teachers training mechanism. Vocational & Technical Education Forum (15) (2009)
[6] Yang, Y., Wan, Q.: Several thoughts on higher vocational colleges' "double-qualified" teachers training. Career Horizon (10) (2008)
[7] Xi, J.: Influencing factors and effective ways of higher vocational colleges' "double-qualified"teachers training. Journal of Zhongzhou University

Education Software Design of Electronic Technique That Based on Virtual Reality Technology

Hu Liang[1] and Bi Xiaoming[2]

[1] Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
[2] Wenzhou Vocational College of Science and Technology, Wenzhou, Zhejiang, 325006
hliang1972@126.com, wzkzybxm@126.com

Abstract. Virtual reality is called artificial environment. It is the high technological imitate system that assisted by the computer technology. It combines computer graphic technique, computer simulation technique, artificial intelligent, sensor technology, display technology and network multiprocessing technology. This article brief introduces the key technique of virtual reality and virtual lab. Moreover, combines the education characteristic of electronic technique, provide the education software design that based on virtual reality technology.

Keywords: virtual reality technique, virtual lab, electronic technology.

1 Introduction

Electrical and electronic technology is one of the professional technology basis courses in the vocational school. It has the important comprehensive, guide, and practical function for the computer or automatic control major students. However, most students feel boring about learning theory and they are not good at manufacture. Therefore, during the electrical and electronic course teaching, the teachers should innovate the education method, change the traditional way that can adjust the science development for the students development.

2 Teaching Questions in the Electrical and Electronic Course Education

First, the students. Vocational school is weakening in the student basement. One hand is the knowledge that learned in the high school has not consolidated that they do not know how to learn the deeper knowledge. The students learning method and ability influence this. On the other hand, maybe it is for the learning attitude of students. Most of them think they cannot enter the colleges for the lack of ability. Therefore, they slight over the course in the vocational school. This kind of negative attitude influences the learning of electrical and electronic technology.

Second, teachers. Some teachers think the knowledge structure is bad in the vocational school. It is hard effectively combine the practice and theory. Some

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 173–178.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

students thinking the learning effect is not good, they should responsible for themselves. These two kinds of teaching attitude is directly influence the teaching method research and operation. Therefore, in the electrical and electronic technology, teacher attitude influence the learning effect of learning effect. Otherwise, the professional quality differences among the teacher have the various acceptances. A part of teachers cannot innovate the new theory and method. It will lead the course teaching only rest on the traditional and dated teaching method that cannot adjust for the education development.

In the last, the course question. The course of electrical and electronic technology is mainly in the theory teaching and assists with practical teaching. It has the great theoretical and practice. Therefore, the teacher should pay attention to the traditional theory and student practice operation. However, influenced by the traditional teaching method, the present electrical and electronic education is centered with knowledge and science system. Moreover, integrate the training course for the learning. Both of the teachings are separated with single training and equipment. Moreover, the laboratory is simple and teacher always guide the students. The students should only follow the steps. This kind of course is boring that students have no interest to learn and will influence the learning effect. This article list some teaching content and class arrangement of one higher vocational school.

Table 1. Teaching content arrangement of electrician and electronic technology in vocational school

Teaching contents	Course arrangement	Theory teaching hour	Course training hour
Circuit and evaluation	32	18	14
Single-phase sinusoidal circuits	22	14	8
Three-phase sinusoidal circuits	26	16	10
Magnetic circuit and transformer	24	12	12
Asynchronous motor application	26	16	10
Safety power supply and power utilization	6	6	0
	136	82	54

3 The Strategy of Improving the Teaching Effectiveness about Electrician and Electronic Technology in Vocational School

With the education revolution and economy development, the market is under the increasing requirement of electrician and electronic talents. On the same, there have

higher talent quality requirement. The manager in vocational school has known that trend. They do the revolution that based on the new education trend and student employment. In there, education method is the most important one. The following table is the revolution contents that aiming at electrician and electronic technology in one vocational school.

Table 2. Revolution content of teaching method

Mark number	The revolution content of education method
1	Take fully advantages of model, teaching aid to assist the education and help them to understand the electrical equipment structure.
2	Use wall chart, multimedia to express the inside and outside of electrical equipment, help them to comprehend the abstract electromagnetism.
3	Train the teachers of network technology, explore the boutique course on the network, and achieve the teaching quality improvement.
4	Open the lab and training room to improve their operational capacity.
5	Organize the college and university ability competition and technology match to train the students' innovation ability.
6	Effectively organize the students to visit the factory and do the post practice.

From these, the writher provides the following teaching strategies. Hope these methods can be the reference for improving the teaching effectiveness of electrician and electronic technology in vocational school.

Improve the principle of course teaching effectiveness of electrician and electronic technology. In the first, based on the present course arrangement disadvantages, the teacher should follow the principle of "hour match the content" while arranging the class hour. The course of electrician and electronic technology is the basic course for computer, electrician, automatic control and other majors. At present, the book market has different textbook editions of electrician and electronic technology. These materials have some different contents and the particular emphasis during the programming is various too. Therefore, the teacher should choose the book with target and choice. Pay attention to the practice and express the center function of textbook during the education.

In the second, the technology of electrician and electronic is the combination course of theory and practice. During the teaching process, the teacher should based on the teaching content, reasonable arrange the class hour of theory teaching and practice teaching that cannot independent the practical training. The arrangement

should suitable combine the theory and practice. This is the second principle to improve the effectiveness of electrician and electronic technology. The teacher should under the learning condition, school experiment condition to arrange the theory teaching and practice teaching. This method can let students to apply the theory knowledge into practice and through practice to test the theory knowledge. Only this kind of education can effectively help the students to internalize the knowledge of electrician and electronic technology to establish the firm foundation of learning other knowledge.

In the last, the society and economy is under the rapid development. The knowledge and technology update is in the same step with the period development. The vocational school is the education institution for training the practical talents, so the teaching method and mode development should express the new technology and new knowledge. This is the third principle for improving the course teaching effectiveness of electrician and electronic technology. With the period development, the technology and knowledge is under the increasing and the teaching content in the textbook is changing at the same time. The knowledge structure is more and more closely with the practical work. Therefore, in order to improve the theory basis and practice ability of students, the teachers should ceaseless increase the self-quality, teach the new technology and knowledge that combine with practice life and work. This can increase the students' learning interest, effectively open their view, and improve the learning ability.

Establish the harmonious relationship between teacher and student is the basement to improve the teaching effectiveness. During the students' learning, teachers are not only the teacher and they are the seniority and friends. Therefore, the harmonious relationship between teacher and student is very important. The teachers should positive research the new technology and knowledge of electrician and electronic technology, ceaseless improve the self-knowledge structure and conquer the students with high quality and professional knowledge. Otherwise, the teacher needs effectively communicate with students. This can understand their thinking, requirement then to transform their learning attitude. The teaching effect on this basement can yield twice the result with half the effort.

Combine the traditional course characteristics of electrician and electronic technology to adjust the teaching mode. The traditional education mode of electrician and electronic technology is centered with theory knowledge. There has not much attention on the students' practical ability training and lead the perfect knowledge system with bad operational skill. This kind of student cannot suite for the market development. Therefore, the teacher should combine the course characteristics of electrician and electronic technology, adjust the teaching mode, and strengthen the practice part. Moreover, train the operational ability, the skill of solving questions and innovation capacity. In addition, the class hour arrangement should closely combine the theory teaching and practice teaching for the uniformed, interinfiltration and both foundation. On the other hand, separate the training course into two parts. They are basic part and ability improvement part. The basic part is aiming at the used equipment and tools. Students can understand the deeply training content then correctly use the equipment. Ability improvement part is under the teachers' organization and guides for the students to do the independent research. Learn the knowledge by them, apply the learned knowledge, and test some propositions. This

kind of teaching mode can effectively increase the innovation ability and research capacity.

4 Use the Advanced Education Technology, Innovate New Method

With the economy and technology development, network information technology has got the widely application. Multimedia technology is the represent of network information technology in the class. It has the directly expression of sound, video, image, and article combination. It can increase the learning interest, save the blackboard writhing. Moreover, the complicated circuit, electronic equipment, and electromagnet response diagram, students can understand them by the multimedia. It will help students effectively understand the contents. Multimedia courseware can use flash to teach. Teacher show the complicated knowledge that can help students to learn more easily. For example, the asynchronous binary system can calculate the dynamic process. Students can understand that directly and more clearly. At the same time, use multimedia can express students learning initiative. Teacher can use the heuristic, understanding research and interactive method to encourage student show their own opinion. They should train the comprehensive ability through discussion and find new knowledge. Other wise, EDA is the new kink of simulating software. Through the modern evaluation we can establish the EDA platform, bring the new theory and technology. Then, do the experiment for find the new theory and technology. Through the theory, stimulate and practice, assist students to learn more knowledge. Students on the better basement will obtain the better training course. On the other hand, students can know, it is very hot. Then , train the better practice, improve the electronic ability.

In order to improve the education effectiveness, there have many strategies. The teacher can under the students' learning condition and ability to create the teaching method with effective evaluation. Moreover, this can increase the learning initiative and train the comprehensive practical ability of students.

5 Summary

During the course teaching of electrician and electronic technology in vocational school, the teacher should express the independent function of guide and organization. Fully mobilize the students' learning initiative that they can establish the good learning confidence of electrician and electronic theory. Then they can do the practical course operation under the theory. After this, the students can learn the theory knowledge, operate the theory, based on the teacher's guide to deeply understand the knowledge. Moreover, train the practical ability, cooperation ability, innovation skill and increase the teaching effect and learning effect. This kind of course teaching is the effective teaching that most teacher and students should learn.

References

[1] Hu, X., He, J.: The course practical education revolution research of in higher vocational colleges. The Journal of Yangtze University (Natural Science edition) (04) (2006)
[2] Wang, B.: Teaching mode revolution research of Electronical and Electronic Technology. The Journal of Shandong Institute of Commerce and Technology (03) (2005)
[3] Chen, X.: Higher vocational school the course design of electrician and electronic technology. Enterprisers (Theory edition) (01) (2009)
[4] Li, K.: Higher vocational school the course exploration revolution of electrician and electronic technology. The Journal of Hubei Radio and TV University (04) (2008)
[5] Huo, Q.: Teaches indefinitely method—brief discussion of teaching design in the electrician and electronic class. New Course Research (middle month edition) (03) (2010)
[6] Jin, M.: Course practice and exploration of electrician and electronic technology in higher vocational school. Vocation Education Research (02) (2010)
[7] Ding, Z.: Education quality discussion and research about electrician and electronic technology in higher vocational school. China Modern Education Equipment (06) (2009)
[8] Chen, X.: Brief discussion of course revolution about electrician and electronic technology in higher vocational education. Intelligence (06) (2009)
[9] Li, J.: Course revolution design of electrician and electronic technology in higher vocation school. The Science Education Article Collects (Half month edition) (04) (2009)

Study on Innovation Education Technology Based on New Computer Technology

Cao Yu

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
jxnccaoyu@126.com

Abstract. This article does the brief introduction about the modern new computer technology, especially the network new technique and virtual reality technique, integral imaging technique, panoramic video technology and other computer multimedia technologies. Moreover, this article evaluates the traditional education technology, points out the present tension development trend that the traditional education technique cannot adapt to the social requirement. This article also looks into the innovation education technology that based on the new computer technology in colleges and universities.

Keywords: Virtual reality technique, panoramic video technology, computer network, education technique.

1 Introduction

The modern education technology is using the advanced education thinking, mode, method, and science technology for the education. Apply the modern education technology during the higher vocational education can improve the teaching effectiveness, optimize the teaching quality, increase the competition ability and reach the independent sustainable development. In the social development process, the higher vocational school is mainly training the higher technology talents. The teaching period is usually pay attention to the practical ability training. Compare with the traditional education mode, in the higher vocational school we operate the education activity with combination of theory and practice. This can let students do the practice under the theory guide and deepen the knowledge during the practice. This article will start from the following parts and aim at the questions and strategies during the daily education to do the evaluation.

2 Characteristic of Higher Vocational Education

Different from the common higher school, the higher vocational school is aiming at training practical ability during the education. The theory representation is not in the advantage. This kind of education mode will lead the undergraduates will place the

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 179–184.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

huge advantages during the employment. This is satisfying the employers practice ability requirement. Students can join into the new work environment very fast. Aims at the characteristics of higher vocational education, here include several aspects. Firstly, aim at the different position requirements, the school education should combine with the period development during the talent training. Secondly, the daily education should train the practical ability from the basement. There will have some chances for the students to join into the relative industries in order to improve their practical ability. In the last, during the education process of higher vocational school, aim at the position characteristic to do the evaluation, we should add the professional requirement into the entire education activity. This can let students based on the detailed employment requirement to perfect the independent question for the further development.

3 Education Technology Application in Higher Vocational School

The modern education technology has six parts. First, education information transform and the technology; Second, education information restores and research technology; Third, education information production and management technology; Fourth, education information express and copy technology; Fifth, modeling and simulating of education test information; Sixth is the experimental process improvement and optimized technology. The application advantage of modern education technology is corresponding with the training target in higher vocational school. We can use the modern education technology to promote the new-type teaching environmental establishment in higher vocational school. The education technology application in the higher vocational school should have the following characteristics.

Application Requirement. During the daily development of higher vocational school, there pays attention to the practical ability training for training the practical talents. The course setting should combine the theory and practice. Students can do the practical operation under the theory guide, and deepen the theory understanding during the process of practical operation. This needs during the education in higher vocational school, we should incorporate with the practical development trend, pay attention to the industrial structure adjustment and talents requirements. Only this can train the new-type practical talents, promote the school improvement and follow the social development.

Various Requirements. The course-setting period in the higher vocational school, there have several requirements except the students' practical abilities. These requirements not only requires students the application ability, but also need the increase of comprehensive skill. This article will compare part majors between Jilin Science and Technology School and one regular college in He'nan. The detailed evaluation is in the following:

Through the compare we can find out the daily education in higher vocational school has higher requirement than the regular college not only in technology requirement but also in the practical ability.

Financial Resources Requirement. Based on the requirement of practical ability in higher vocational school, the school development needs amount of lab or other important classrooms. This process provides the relative requirement about the school financial resources. In other words, during the practical ability training in higher vocational school, there need to equip the financial basement except the enterprise assistant. Only this can entirely train the students' practical ability and satisfy the employment requirement from students.

Subject Requirement. The education platform establishment should express the teachers' guide with students' subject places. The platforms include resource manufacture platform, resource management platform, resource utilization platform, and resource communication platform. There should equip resource catalogue browser, resource search, foreground service management, system backstage management, resource statistic, and independent knowledge manager. This can explore the function of independent learning and research learning. Audio frequency has auto-timing broadcast, cycling broadcast, single channel in the sequence broadcast with many program. Moreover, it can support the video source and form various. It can broadcast every sound and imaging equipment signals on time. The users on Internet or local area network can do the site teaching and broadcast many different classrooms at the same time.

Teaching Requirement. The students' sources in our higher vocational school have not perfect performance in the college entrance examination. These students has the weak basement knowledge, exist the learning initiative and weak understanding. During the teaching process, the teacher should combine the students' practical condition with pertinency. In the entire education process, the teaching includes the following aspects. First, teacher should fully combine the advantages of education technologies and integrate the practical learning condition with pertinency education. Second, the teachers should pay attention to train the innovation sense and practical ability during the teaching to establish the firm basement for the future. In the last, the education technology should effectively join into the various subjects' education in order to provide the new teaching environment for the students.

4 Existed Questions of Present Education Technology in the Higher Vocational School

With the recent years' development, the higher vocational schools get better performance. The education technology is influenced by the economy development with various aspects to appear the low education quality and weak students' quality. Aiming at the present existed question, the detailed evaluation is in the following.

The Software and Hardware Establishment Is Out of Step. The software and hardware establishment is out of step is the ordinary question during the education technology development. With the network popularity, education informationization promotes the education revolution and increase the society development. In the higher vocation school's development, the hardware equipment establishment obtains the better performance. However, the software is very weak that cannot catch up with the hardware establishment. This condition directly influences the education technology application and development in higher vocational school.

Pay Attention to the Effectiveness and Despise the Effect. This process includes the following aspects. First, pay attention to the technology and despise the theory. Many teachers not pay attention to the theory, they doubt and negated the theory and pay attention to the technology. In the second, pay attention to the manufacture and despise the application. When people make the education software, they think use the education technology can promote the education. They are not pay attention to the design, feedback and evaluation. This leads the lack of advanced and perfect education software. In the last, pay attention to the reword and despise the application. In the daily education process, much teachers use the education technology only for the independent title.

The Big Differences between Theory Research and Practical Requirement. Based on the practical development condition in higher vocational school, much teachers pay attention to the theory and despise the practice. This kind of teaching mode will influences the education quality and affects the students' further learning. On the other hand, some teachers are on the contrary. They pay attention to the practice and despise the theory. This kind of method can increase the students' practice ability greatly. However, the theory understanding is in the blurry condition. From here we can know, there still exist great differences between theory research and practical requirement in the higher vocational school.

5 The Plan Present and Improvement of New-Type Education Technology

The plan present and improvement of net-type education technology can increase the teaching quality with school activities perfection. It can establish the firm foundation for students' further employment. Aiming at the plan, here the writer provides some aspects.

Change the Thinking. First, aiming at the application of modern education technology, teachers should abandon the old thinking and take fully advantage of modern technology and teaching contents. This can make the abundant the teaching content and increase the teaching quality. In the second, the school manager should organize the teacher training to promote teachers' comprehensive ability and create the new teaching mode, transform the old education opinion. In the last, aiming at the education technology application, the higher vocational school should based on the

independent condition, correspond with the society development to follow the period steps.

The Improvement Strategy of Education Technology Plan. The education technology application should the effectiveness combination between theory and practice. Aiming at the improvement strategy, there has the following two parts.

Teacher Training. Teacher as the implementer of the entire education technology plan is not only relating with the application effect and close to students' further. Strengthen teachers' technology training can improve the technology understanding. Moreover, teachers can use the correct technology in the further teaching. They can increase the teaching quality while promote the development of higher vocational school.

Learning Support. In the education technology application, the basic target is improving the students' learning ability and supports them during the learning. This kind of support is corresponding with the present network development trend. Aiming at students' learning, teaching organization and other teaching evaluation, combine the theory and practice can train students' practical operation ability while pay attention to the theory knowledge deepen. Only this can train the perfect comprehensive talents and put them into the further learning and work.

6 Summary

In the summary, with the rapid development of the society, the technology plan support and application in higher vocational school can increase the teaching quality and greatly improve the students' practice ability and innovation skill that establish the firm foundation for the future. Therefore, here we need the higher vocational school can combine the society development trend with school condition, provide the scientific and effective teaching technology plane and perfect it. This is the best way to train the new-type talents that can promote the society development.

Acknowledgement. This article obtains the support of Vocational colleges research of "double qualified" teachers team construction from the planning issue of 12th Five Year Plan in Jiangxi science of education(No. 11YB114). This paper is one of the research results.

References

[1] Song, X.: Modern education technology in the education development. The Journal of He'nan Education School (Science and Technology Edition) (01) (2006)
[2] Jiao, L.: The compare between higher vocational education and regular college education. Education and Vocation (21) (2005)
[3] Yang, Y.: Education theory and teaching method revolution that based on the modern education technology platform. Qing Education (11) (2006)

[4] Xing, G., Wei, D.: Based view of education technology——the new research of education technology resource. China Audio Visual Education (01) (2005)
[5] Chen, Y.: Discussion of operation theory in the new period of higher vocational school. China Higher Education Research (06) (2004)
[6] Ji, X.: Modern education technology application and student main body. The Journal of Zhejiang Wanli University (01) (2003)
[7] Li, J.: Create the lively information environment for students and teachers—life environmental view of education information. Education Research (02) (2002)
[8] Deng, R., Liu, M., Chen, S.: Research evaluation about application strategy of modern education technology in higher vocational school. Professional Circle (16) (2007)

Research on Multimedia Database

Hu Liang, Huang Manyuan, and Huang Chengzhong

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
hliang1972@126.com, 331629826@qq.com, nchuangchengzhong@126.com

Abstract. Multimedia database is the basic subject as well as the key problem in multimedia field. This article discusses the key technique and evaluates the structure that based on the present multimedia database. Moreover, it combines the complexity of ideological and political education, and provides the multimedia database character and application foreground that aiming at ideological and political education.

Keywords: Multimedia, Database, Ideological and political education.

1 Introduction

With the rapid popularity of computer and network, they push the society development and perfect the living mode. The progress of ideological and political education is the prospect of social development. It is centered with satisfying the social requirement and focus on promoting the talent comprehensive development. During the ideological and political development, there not only regard with the talent value, moral look and the correct view of society, but also concentrate with the society steady and national blossom. The multimedia education is under the traditional mode to join the computer network and through the advantages to make up the insufficient of traditional education. This article will do the following evaluation that aim at the questions and development prospect of ideological and political education development in the multimedia education.

2 Ideological and Political Education Characteristic in the Multimedia Education Mode

The multimedia education in the ideological and political education can through the network advantage to expand the education content and arrange except than impart the theory knowledge. Moreover, it can abundant the student ideological and political cognition while deepen the ideological and political understanding. Aiming at the character of ideological and political education in the multimedia education mode, the detailed analysis is in the following:

Education Subject Highlighting. In the traditional ideological and political education, teacher is the organizer and sponsor during the teaching activity. The

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 185–189.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

teacher places the highlighting and students can only accept the knowledge during the learning. This method will lead the learning weariness mind. The usage of multimedia education mode breaks the "force indoctrination" and the teacher transforms to the knowledge "manager" rather than the initiator. Through the multimedia, teacher can expand the original contents with the society development. This can provide newest and nearest references for the student learning. This learning mode can extrude the students' subject and encourage their positivity of ideological and political learning. Students can make a commitment to the further study.

Education Object Importance. The target of network ideological and political education is the students in the network society. In the first place, compare with the traditional education, the students initiative are more than the passivity. In the network education, they research the information rather than passive accept. They can obtain the various contents in the different levels. In the second, the network can express the subject function, through the information control to influence other students to reach the effect that traditional education cannot achieve.

Education Striking Effect. The multimedia education operation during the ideological and political education has the striking education effect that except the above two advantages. Aiming at the education effect evaluation, we do the analysis through table 1. From the basis to evaluate multimedia education mode, the teacher separately statistics the regular class and the experimental class that limited in one semester. During the statistic, we do the evaluation that aiming at the various parts.

Through the one semester experiment we can find the experimental class that use the multimedia education has better performance that the regular class that not use the multimedia education. From this, we use the multimedia education can improve the ideological and political performance. Moreover, this method can deepen their ideological understanding and establish the correct thinking and value.

3 The Positive Meaning of Multimedia Education in the Ideological and Political Education

The teachers in the ideological and political education, they use multimedia education mode. When they bring the new teaching mode, they encourage the students' learning initiative. This can let the students can learn better in the further ideological and political study. Aiming at the positive meaning, the detailed evaluation is the following.

Reach the Teaching Resources Sharing. During the ideological and political education, the multimedia education can provide the large amount of materials for the teachers by the personal advantages. The abundant contents to prepare the lessons are very useful. The teachers can download the required materials and combine the teaching contents to teach with the pertinency.

Increase the Teaching Actual Effect and Expand the Students View. The actual effect is the start and terminal of network education as well as the standard to measure

the class quality. Jiang Zemin said, "We need to pay attention and fully take the advantage of network technology. Improve the teaching actual effect of ideological and political education, enlarge the coverage, and increase the consequence." The network education includes voice, image, video, and animation. This can fully mobilize the students' participation, increase the selectivity, and express teach students in accordance with their aptitude. This can transform the abstract, hard, and boring theories into detailed, vivid, and visualized that can easily understand and accept. Moreover, it brings the new energy and vitality to the traditional education that increase the interesting and effectiveness of ideological and political teaching.

Train the Student Autonomy and Research. Based on the theory and realizing of ideological and political education, if the teachers want deepen the students understanding form the basis during the education process, they need to combine the society development trend and seek truth from facts in order to prove the students thinking. The application of multimedia education can through the computer advantages, combine the contents of ideological and political education, aiming at the social condition to improve the students' autonomy and research ability while deepen their understanding. At the same time, the multimedia openness leads students can publish their own opinion on the webpage while searching for the required materials. They can discuss, analyze and improve the own performance. Meanwhile, this method can encourage the learning interest to reach the study target.

4 The Core Center of Multimedia Education Mode in the Ideological and Political Education

In the ideological and political education, except the correct thinking guide, the multimedia education needs the education department support and the effective coordination of human power, material resources, and financial resources. Only this can effectively implement the multimedia education in the ideological and political education. Aiming at the core center of multimedia education mode in the ideological and political education, the detailed evaluation is under the following.

The Ideological and Political Educator Training in Multimedia Education. The multimedia educator in the ideological and political education needs the experienced computer operation ability and the educational capacity. This is the only way to take fully advantages of multimedia in the ideological and political education. The educator training includes the following. First, the manager of each college should train the educator on time to insure the increasing of teaching capacity. This can confirm the perfect combination off theory and practice during the teaching process. Second, during the step training, there should pay attention to the computer operation. The multimedia education requires the teacher has the relative computer operation ability. Some elder teachers cannot experienced operate the computer with the result that cannot use the multimedia education. This shows strengthen the computer operation training is the basic safeguard for the multimedia education development. Thirdly, the educator of ideological and political education needs the aggressive pioneering spirit during the multimedia education. in the last, aiming at the

multimedia usage, we should combine the social development and integrate the practical learning ability and knowledge receptivity. Only this can reach the teaching target.

Network System Establishment of Ideological and Political Education. Build the perfect network system of ideological and political education can improve the entire college teaching and promote the multimedia education development. However, during the establishment process, there exist some aspects: firstly, the network information system should include comprehensiveness and objectivity. Secondly, the network information should be positive and obverse without offset information that standing in the advancement. In the last, the information should keep the main body of educator. Publish the correct information on time and accept battle with the error message to play the effect of "protection".

5 Ideological and Political Development Potential in the Multimedia Education

The multimedia education operation in the ideological and political education can promote the teaching revolution. Moreover, it can deepen the students thinking, expand their knowledge and abundant the accumulation of knowledge. Compare with the traditional education, no matter in the teaching target or teaching mode, the multimedia education places in the superiority. In addition, during the multimedia education operation, we can express the theory of ideological and political education with animation, video, or other forms. This can let students deepen their understanding through the independent sense. From this we can see, the multimedia education provide a fair communication and popular platform for the students. Aiming at the development potential of multimedia education in the ideological and political education, the relative workers conclude the following evaluation through the compare of traditional education and multimedia education.

Table 1. The compare of traditional education and multimedia education

Subject	Traditional education	Multimedia education
Resource	Fasten, rigid, resource limit	Network resource sharing
Mode	Fasten	Flexible
Main body	Teacher	Student
Arrange	Narrow	Abundant
Significant	Performance is the center	Students' autonomy and innovation is the center

In the table 1, through the compare of traditional education and multimedia education we can find out no matter in the resources or the education subject, the multimedia education has the definitely advantages. This teaching mode can greatly

train the student innovation sense that they can deepen the ideological and political content while provide their own opinion. Therefore, we can infer that during the ideological and political education, the multimedia education has the vast potential for future development. The relative education departments and college managers combine the practical condition and popular the multimedia education with pertinency. They support the development with human power, material resources, and financial resources. This can increase the teaching quality while establish the stable basement for the student growth.

6 Summary

In conclusion, the rapid development of 21th century technology, the multimedia education promotion in the ideological and political education pushes the education revolution and improves the teaching quality. There needs the education workers can combine the social development trend and network advantages. During the ideological and political education, use the multimedia education pertinency. While deepen the students' thinking, it can help them to establish the correct value that settle the basement for the further growth.

References

[1] Jiling, Z.: Network education mode of information technology research and practice that based on constructivism learning theory. Examination Week (54) (2010)
[2] Hu, L.: The malpractice and strategy of over dependence on the multimedia in college teachiing. In: The Memoir of the Second Nationwide Agriculture and Forestry Colleges Educational Science Postgraduate Academic Forum (2010)
[3] Zhang, S., Ji, W.: Abuse education of completely multimedia education. Higher Education in Chemical Engineering (01) (2005)
[4] Fu, C., Zhuang, C., Wei, Y.: Application research and analysis of modern education technology in higher education–Hebei normal university of science and technology is the example. The Journal of Hebei Normal University of Science and Technology, Social Sciences (04) (2008)
[5] Xia, X.: Advantages and disadvantages analysis in multimedia education. Decision and Information (Financial observation) (09) (2008)
[6] Wang, R., Qu, R., Li, H.: Use multimedia education can increase the teaching quality. The Journal of Jiamusi Education Institute (01) (2003)
[7] Wu, Y.: The thinking of multimedia education quality in our colleges and universities. Xuelilun (31) (2009)
[8] Zheng, S.: Multimedia technology application for increasing the teaching quality. Journal of Liaoning Agricultural Vocation-Technical College (03) (2009)

The Experimental Teaching Platform of Mechanotronics Major That Based on Java EE Technology

Hu Liang and Huang Chengzhong

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
hliang1972@126.com, nchuangchengzhong@126.com

Abstract. This article will take the advantage of Java EE technology to reach the education platform thinking, system combination that based on Web. Moreover, the article will describe the advantage and disadvantage of this system during the practical application. At the same time, introduces how to use the open software of Java EE technology such as Hibernate, Spring, Proxool, Struts, and Tomcat into the experimental teaching platform of Web network. This article does the brief introduction of how to use the Java EE to explore the experimental teaching platform of mechanotronics major.

Keywords: Java EE, Web, experimental teaching platform.

1 Important Significant of Professional Training Teaching for Mechanical and Electrical Integration

The core-training target of higher vocational school provides production, establishment, and management of comprehensive technical talents with high quality and ability that can satisfy the enterprises requirement. This target decides the basic characteristic of aiming at training education for developing the professional technical talents. Moreover, this is also determining the important position of training teaching during the personnel training process. Training teaching has the important significance for the increasing of practical ability, manipulative ability, innovation ability, responsibility sense, and even the professional dedication. Training teaching experiences also establish the development basement for the students' further career. Mechanical and electrical integration is one of the significant majors in the higher vocational school. In the school, teachers train the professional theory and comprehensive professional ability of mechanical and electrical integration technology. The school is mainly work at training the comprehensive talents with the first level technology quality and ability of equipment application, management, development, marketing and technical service.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 191–196.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

The core-training target of higher vocational school provides production, establishment, and management of comprehensive technical talents with high quality and ability that can satisfy the enterprises requirement. This target decides the basic characteristic of aiming at training education for developing the professional technical talents. Moreover, this is also determining the important position of training teaching during the personnel training process. Training teaching has the important significance for the increasing of practical ability, manipulative ability, innovation ability, responsibility sense, and even the professional dedication. Training teaching experiences also establish the development basement for the students' further career. Mechanical and electrical integration is one of the significant majors in the higher vocational school. In the school, teachers train the professional theory and comprehensive professional ability of mechanical and electrical integration technology. The school is mainly work at training the comprehensive talents with the first level technology quality and ability of equipment application, management, development, marketing and technical service.

The mechanical and electrical integration includes hydraumatic, sensor, network, PLC, communication and pneumatic. They interpenetrate with each other. The technologies are using widely in industry, energy, transportation, national defense and other fields. It is the important combination in the modern industry technology filed. Numerical control machine, flexible manufacturing system of MPS, circuit board packaging, robot system, and other supermatic equipments are using widely in our Yangtze River Delta and industry developing areas. In this way, grasps the theory knowledge and professional ability of mechanical and electrical integration is under the urgent need. Moreover, it is also imperative to become the talents with the ability of operation, maintenance, test, management even marketing and develop the equipment.

The higher vocational school needs to pay attention to the training teaching in order to suit for the normal development of high and new technical industry, the region economy, and satisfy the technical talents' requirement. Moreover, training teaching is the important method for training students the comprehensive professional ability. Research the mechanical and electrical integration as the development of training teaching system, establish and perfect the suitable system for the high quality talents will has great significance for the society requirements of professional ability talents.

2 Establishment of Training Teaching System for the Mechanical and Electrical Integration

Higher vocational education has obviously professional education that can train the students' professional ability. The major of mechanical and electrical integration has particularly prominent requirements of training students' comprehensive abilities. In this condition, the course system of this professional major should targets with

training the comprehensive profession quality and ability. Theory teaching and training teaching is the big support to achieve this target. During the teaching system establishment, there needs following the basic rules of objects development, and complies the basic principle from the easier to harder, from the shallower to the deeper and from the single to the comprehensive. In order to reach the steadily application of teaching target, we need to ensure the theory teaching content, skill certification teaching and training teaching contents are coordinate with each other and combine together. Based on this principle, the writer divides the content system into basic ability training, professional ability training, and comprehensive training. The structure is the echelonment.

2.1 Basic Skill of Training Teaching

Basic ability training is the most basic stage in the training teaching system of mechanical and electrical integration. Related with basic theory teaching, this phase is aiming at metalworking practice, experiment of electrician and electron, CAD design and mapping training, turning lathe training and so on. The training teaching target of this stage is practicing the students to familiar with basic experiment equipments. Moreover, train them to possess the basic understanding of tool software and accumulate the basic usage experiences. Through the experiments and practice, they can understand the CAD and other graph software protraction, basic data measurement, collect and grasp the basic experiment methods. Students can be familiar with the primary metalworking of pliers, milling, turning, digging, and soldering. From the beginning, the school pays energy to train the practice ability, innovation ability, and adaptability for the students.

2.2 Professional Ability of Training Teaching

Professional ability of training teaching is training students the abilities of technology knowledge application, practical operation, actual question settlement, and other key parts. In this stage, students should be familiar with professional ability and obtain the related professional qualifications in table 1. It is the target of training teaching. In the detailed training, the school should direction with training the comprehensive application ability, and aiming at the specific major direction that students collected. Moreover, training teaching is the details that follow the process of “basic ability training→simulation operation training→actual operation training→expanding training→ identification of profession vocational qualification”. Through the acquaintance and punning of professional ability, it is laying a solid foundation for the important comprehensive ability training.

2.3 Comprehensive Ability of Training Teaching

Comprehensive ability training teaching is the final target of basic ability training and professional ability training. Moreover, it is the key part of achieving comprehensive

professional talents' training of quality and ability. Comprehensive ability training teaching includes production training, undergraduate training, and position practice. This training teaching target is making students to form the direct acknowledge of working environment, further practice and students adaptability improvement and working ability in the social environment and working position. It can short the employment adaptable time. Train the professional quality and team cooperation for the students is the main target of comprehensive ability training. The training contents in this stage is basing on applicability of the national professional classification and professional qualification. The training requirement and assessment criteria needs to link up with the related professional qualification standard. Strengthen the training system availability, pertinency, and continuity. Follow the national and social qualification and requirements to train the students' comprehensive professional ability and quality. Combine the teaching, learning and practicing in the higher vocational teaching. Otherwise, objectively promote the double certificates of professional qualification and academic certificate, ensure the students can smoothly graduate, obtain employment, and the enterprise can find the required talents.

3 Professional Training Teaching Management and Protection Measurements for Mechanical and Electrical Integration

The training teaching of mechanical and electrical integration major not only needs the basic teaching system frame, but also needs each necessary condition that manages and determines the system effective operation.

3.1 Form the Scientific Training Plan and Exam Approach That Correspond with Actual Work

In each stage training teaching, training plan formulation needs to center with practical and scientific. The reasonable training plan that only corresponds with the practical professional background can really reach the target of training the professional talents that suit for the social requirements. During the plan formulating, there needs fully consider about the teaching facilities, students qualities, period arrangement, and various courses' articulation. Training teaching should first ensure the correspondence of practical work. The evaluation standard should use industry standard and regulation. Only the practical exam standard can reflect students' progress and shortage. For the detailed examine methods, we should use the combination of specification evaluation and process evaluation. The specification evaluation is aiming at experimental training. Process evaluation is aiming at practical training. The examine method will under perfection and optimization during the practice.

3.2 Training Teaching Mode of Creating the Work Environment That Close to the Actual Environment

The training teaching mode that closed to the real working environment will help students for improving adaptation in the working environment. Just like using the ability competition, we can group the students, arrange the training tasks, and limit the time. Because the time and work load limitation, each team has the competition relation and form the working competition mode in naturally. This competition mode formulates the actual working process. Moreover, it provides the teamwork for each group and the coordination is suit for the practical working environment. This kind of training teaching can practice the professional ability of mechanical and electrical integration, and make students to realize the significance of teamwork.

3.3 Build Training Basement That Has Advanced Equipments and Correspond with the Present Condition

In order to reach the target of training students' the comprehensive professional ability, there needs to support the really or the training environment that close to real. For the environment, the school should establish a training basement that has advanced technical equipments. Moreover, it needs to correspond with the present industry condition. For example, the parts sorting station, pneumatic machinery and so on. The equipments need to as true as possible that can determine the students can study and practice under the real work environment. Complete equipments and "real" training basement is the basic support to train professional technical talents.

3.4 Strengthen the Teacher Group

The qualified teacher is the necessary condition for training teaching. Based on the requirement of theory education and training teaching, the teacher of mechanical and electrical integration major should have the related abundant theory knowledge, practical experiences, and professional quality. They can provide the knowledge to the students, and need to do the samples and guide students during the production, training sites. Moreover, they need to familiar with the industry regulation, occupation standard and job requirements. The major of mechanical and electrical integration is the comprehensive major, so there need to have the mechanical teacher and electrical teacher.

Acknowledgment. This article obtains the support of Vocational colleges research of "double qualified" teachers team construction from the planning issue of 12th Five Year Plan in Jiangxi science of education(No. 11YB114). This paper is one of the research results.

References

[1] Zhao, X.: Mechatronics system. Higher Education Press (2004)
[2] Wang, H., Zhu, Z.: The thinking of mechatronic training teaching. Science and Technical Information (2) (2006)
[3] Cheng, J., Dong, L., Liang, J.: Electrical and electronic engineering revolution and practice in the training teaching for the mechanical and electrical integration in higher vocational school. Science and Technology of West China (13) (2010)
[4] Li, X., Zhou, W., Huang, W.: Teaching mode practice and exploration of comprehensive training for the mechanical and electrical integration in higher vocational school. Shananxi Education (02) (2008)
[5] Geng, B.: Discussion of training teaching revolution in higher vocational school. Wen Jiao Zi Liao (15) (2007)

The Analysis of Grid Database Technology in the Computer Application System

Juan Chen

Department of Information Engineering, Hunan Mechanical and Electrical Polytechnic, Changsha City, Hunan Province

Abstract. Through the introduction and analysis of the concept of the grid, grid database technology, discuss the status of the grid database technology, the content and direction of the research, and discuss problems that the technology faces in computer applications.

Keywords: grid, grid database technology, applications.

1 Introduction

A lot of convenience is brought to people's lives and learning with the internet advent. However, the various servers become Resource Island due to the lack of effective contact among them, and lead to a lot of useful information resources which is archived by storage devices can not be effectively sharing and exchange. Grid is a new network computing platform following the World Wide Web, and its purpose is to provide users with a infrastructure that people can comprehensive share resources. Data as an important Grid resource occupies an important position in the grid computing, and how conveniently access to the grid data is to become an important topic in grid research field. By far, with grid technology development it varies from the original computational grid to today's storage grid, data grid, information grid, etc. and the grid database technology is produced through combination of grid technology and database.

2 Introduction of Grid Database Technology

Grid use of the Internet, all kinds of resources which are widely distributed geographically,

For example computing resources, storage resources , software resources and data resources together as a logical whole make up a virtual supercomputer, provide users with integrated information application services, in order to share Internet resource comprehensive. Therefore the sharing of resources is the essential characteristics of the grid. The core problem that is resource sharing and collaborative working in a dynamic virtual organization is going to be addressed.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 197–201.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2.1 Grid Architecture

The grid architecture that is the basis for building a grid system is a framework which constructs grid. It defines the sharing relationship about the negotiation, establishment, management, use between users and resources the basic mechanisms. Five levels hourglass structure and open grid services architecture are generally admitted, five levels hourglass structure is introduced in this paper.

Five levels hourglass structure treat the "agreement" as hourglass architecture's heart, and stress the position of agreement in the grid resource sharing and interoperability. The simple structure and clear hierarchy are its main features, and it focuses on the qualitative description rather than a specific protocol definition and is helpful to the overall understanding of the computational grid. Another important feature of five levels hourglass is hourglass shape. Due to the different number of each part of the agreement, the short supply number of the core protocol, So the core protocol form a bottleneck in the protocol hierarchy and the resource layer and link layer together form a bottleneck part of the core. It is similar to the traditional TCP / IP network protocol and various components are distributed in five different levels according to the functions, shown in Figure 1. From the outside to inside: application layer, convergence layer and resource layer, link layer, structural layer.

The basic functions of the grid structural layer is control of local resources and to provide an interface to access these resources; the basic function of the link layer is to achieve mutual communication, and It defines the core communication and authentication protocols as network transaction processing for the grid; the resource layer function is to achieve the sharing of individual resources; protocols and services (including the API / SDK) describes the commonalities of the resources, and it does not involve the specific characteristics of resources; the application layer exists in the virtual organization environment, and is constructed according to the services defined in any one level.

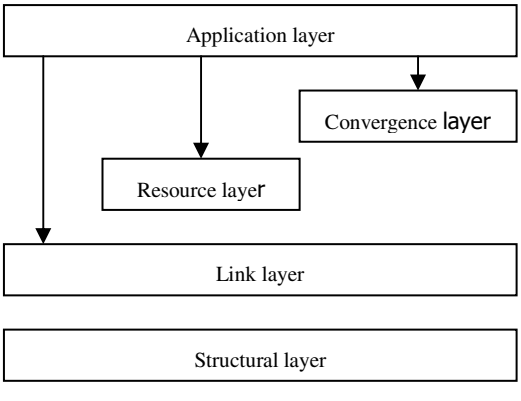


Fig. 1. The hierarchical graph of five-story hourglass structure

The Open Grid Services Architecture (OGSA) is the most important grid architecture following the five-story hourglass structure, and it is a kind of structure whose core is service, and is called the next generation of grid structure.

2.2 The Status of the Grid Database Technology

The grid database gridded the existing database. The purpose is to achieve efficient data management functions and provide support for wide area data resource sharing. It is based on Open Grid Services Architecture Institute which provides mesh database service. Let Grid users or grid service via a mesh database service access to variety of heterogeneous databases of grid.

In the field of grid database technology research, the research scope and scale of the U.S. and Europe is relatively large, in a leading position. On the basis of facing computing grid Initially to study and realize the function of data management the Globus system has become the most famous grid data management system development platform.

China has also done a lot of research work to promote the application of the data grid. A solution is proposed for multiple applications encountered problem of mass data storage, management, processing and joint services.

3 In-Depth Study of Grid Database

3.1 The Content of Grid Database Technology Research

Grid Database research mainly includes three aspects of grid database management systems, grid database integration and support for the new grid.

The grid database management system is an important resource of the grid. Two steps can be divided to build the grid database management system: the first step is to provide a middleware so database management system packaged as a grid service to grid access to the grid database; the second is expansion of existing database management system, and it is directly provided by the grid to achieve a distributed database and grid services.

The grid database integration is to use the information in two or more grid database, and use this information to build a large database. The virtual database is a federal database and it is only a federal model, and all users are unable to perceive the fact of the existence of multiple independent databases. The virtual database is an ideal target too much attention to detail and custom integration, incremental integration of living between the two. Incremental integration, developers need to complete every detail of the integrated and advanced data access and integration of components can be automatically completed some post- integration steps in support of the new grid data stream processing , information retrieval and scientific data analysis grid.

3.2 The Development of Grid Database

Now with grid technology increasingly widely used in the database grid database will be a rapid and steady development. Comprehensive analysis of research in the grid database at home and abroad, the trends of grid database development in the following areas:

1. Focus on database access and integration standards. Although the issues involved in grid database integration is not fully resolved, currently existing grid database access and the integrated draft standard and this draft is available to let the actual software system achieve the goal.
2. More research on grid database management system. The access and integration of grid database using the method of middleware, expect future database will support the grid more directly.
3. Continue to explore the grid the new demands. Scientific research in the grid have a lot of problems about data management to be solved, Some of these problems can not be solved just simply the data to mount the database and they are able to promote the development of database technology.
4. The new database technology will be carried out under grid environment. If you put these studies into the grid environment , you will receive an unexpected effect.

3.3 The Problems of Grid Database Faced By

In the process of the development of grid database advancing forward, the following question is worth to discuss.

3.3.1 Focus on the Safety of the Grid Database

The grid database implementation on data resources which are distributed, heterogeneous, autonomous, mass and other characteristics in the grid a unified access and integration, variety of data resources sharing and collaboration within the scope of the WAN on the Internet. These will bring the security issues of data sharing. Therefore pay particular attention to issues of data storage and user authentication in process of data transfer, authorization and access control and audit and data integrity to develop the safety standards of the grid database. In addition, there is the need to make sure to take adequate measures to deal with the impact of worms and viruses.

3.3.2 Increase Research of Grid Database Performance Monitoring

With the grid database technology shifted from the fields of scientific research gradually closer to new applications of people's lives, along with traditional database monitoring system can not well adapted to the grid database performance monitoring. Therefore, there is the need to enhance research of the grid database performance monitoring system that is useful for the implementation and expansion of new applications of grid database.

3.3.3 To Carry Out the Research of Grid Database in the Field of Artificial Intelligence

Data mining built on a grid database platform combined with grid computing ideas and technical advantages is able to efficient processing, analysis and mining of wide-area distribution of vast amounts of data, with a view to their application to expert systems, artificial intelligence.

4 Conclusions

With grid technology increasingly widely used in the database , the grid technology provide a wider space for the development of the database, Grid database technology which combines grid technology and database technology has good prospects for development to solve data access and integration under the Internet environment. Grid as an important emerging technology there are a lot of controversial content, and how grid database rapidly and steadily develop remains to be the people to further study.

References

[1] Zhao, H., Tan, G.: The improvements of Grid database and applications in the Intelligent Transportation Systems (2006)
[2] Hu, G.: Datebase study based on grid. Science Technology and Engineering 6(12), 1644–1647 (2006)
[3] Liu, Y., Wu, Y.: GridDatabase Access and Integration. Fujian Computer 7, 34–35 (2007)
[4] Ling, Z., Kang, W., Xin, F.: The requirements and solutions of grid database services. Computer Science 33(12), 75–77 (2006)
[5] Antonioletti, M., Atkinson, M., Baxter, R., Borley, A., ChueHong, N.P., Collins, B., Hardman, N., Hume, A., Knox, A., Jackson, M., et al
[6] The Design and Implementation of Grid Database Services in OGSA-DAI, http://aspen.ucs.indiana.edu/CCPEwebresource/c815watson/c815OGSA-DAI-6.pdf
[7] Xia, Y., Zhao, H., Zhang, L., Li, J.: Analysis and application of grid database technology. Chongqing Institute of Technology (Natural Science) 21(12), 98–101 (2007)

Study of Smart Warehouse Management System Based on the IOT

Wen Ding

Hunan Mechanical & Electrical Polytechnic,
410151 Changsha, China

Abstract. With the development of enterprises and the constant demands of the product diversity, traditional warehouse management model cannot meet that, due to its heavy workload and low efficiency. This paper presents a new type of intelligent warehouse management system - Smart Warehouse Management System based on the IOT, and expounded the principles and structure of it. This system has great advantages compared with the traditional mode, and we expect good prospects for its development.

Keywords: internet of things, warehouse management system, RFID.

1 Introduction

Warehouse information management plays a decisive role in goods supply chain management, and only with a smoothly proceeding of purchasing, inventory control and shipping in the warehouse management, can we effectively reduce the cost of the enterprise and improve service quality and competitiveness. With the development of economic and warehouse management technology, inventory management operates faster and faster, at the same time, warehouse operations and management control has become more and more complicated. It is hard for manual operations to accomplish a huge amount of data due to its low efficiency and time-consuming, what's more, tremendous losses will be caused if there are errors. Therefore, the use of intelligent warehouse management tools has become an urgent need for the current businesses.

2 Internet of Things (IOT) and Its Characteristics

2.1 The Concept of Internet of Things

Internet of Things is a network concept, which connected any items to the Internet, realizing to the exchange of information and communications, to achieve intelligent identification, location tracking, monitoring and management according to the agreed protocol through radio frequency identification (RFID), infrared sensors, global positioning systems, laser scanners, and other sensing devices. The role of IOT is to achieve all management personnel and cargo on the network connecting, so that makes it more convenient to manage and control warehouse information.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 203–207.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2.2 The Composition of IOT

From the technology architecture, IOT can be divided into the sensor network, transmission network and application network. The sensor network achieve to Identification of the "objects" through two-dimensional code, RFID, sensor-based, just as a sensory nerve endings, which can identify objects and collect information sources. Transmission network has a role of data transmission and computing through the Internet, radio and television networks, communication networks and cloud computing platforms, just as human neural network and the brain, which transport and process the sensor network information. Application network is a Input and output control terminal, and its role is to achieve the needs of the users (individuals, organizations) to reach the intelligent application of the IOT.

2.3 Characteristics of IOT

Firstly, the IOT get large amounts of information through using of advanced sensor technology. All sensors, respectively, are independent sources of information, and deliverer the captured information to the transmission system in different formats. The advantages of the sensor is being able to collect information in real time to upload, and can be subject to the control of the network to gather information according to a certain sampling frequency.

Secondly, based on the Internet, IOT can achieve a real-time arrival of the object information and management operations. Internet technology is still the core technology of the IOT, and achieving a real-time arrival of the object information and management operations through an integration of Internet with radio and television networks and the communications networks. The collected information can be massive to describe since the sensors are varieties. The IOT adopts new technologies to adapt to a variety of heterogeneous networks and protocols to ensure data integrity, a wide range of coverage and timely delivery.

Thirdly, on the basis of the sensor and network interconnection, the IOT also implements intelligent processing of data and the implementation of intelligent control on objects. The IOT takes advantage of cloud computing, pattern recognition and other emerging smart technologies to achieve the expansion of the function after the analysis of, processing and handling sensor information through the network upload back to warehouse management application mode feedback in accordance with the needs of different users.

3 The Application of Internet of Things (IOT) to Intelligent Warehouse Information Management System

3.1 Warehouse Inventory Management

Warehouse inventory management refers to reflection of a variety of storage and the flow of materials in a timely manner and provision of the basis for production management and cost accounting according to management of warehouse, cargo and other account and type and data of in / out library into / storehouse. Provide inventory

occupancy of funds, materials backlog, shortages / super storage situation, ABC classification such as different statistical analysis for managers and decision-makers through the inventory analysis. Through the tracking of batches, to achieve specifically approve and quality tracking smoothly.

3.2 Demand Analysis of Inventory Management

With the continuous development and expansion of enterprises and the continuous expansion of the data of the goods in the warehouse, traditional inventory management model appears stretched, and necessary requirements are following:

(1) Ensuring that the amount of goods are stored and improving the response speed In Inventory Management are basic requirements to meet customers' demand, ensure the normal operation of the enterprise and improve the economic efficiency of enterprises.
(2) Standardization and simplification of the cumbersome business processes such as in/out, inventory and transport is also the requirements of modern enterprise inventory management. Standardization and automation of business processes can improve the responsiveness and efficiency of inventory management, and scientific integrity of the management process and a smooth organizational structure can ensure no redundancy and efficient operation in inventory management and logistics business.
(3) With the improvement of treatment of staff, in order to reduce the cost of doing business, modern inventory management also requires improving automation of warehouse systems and work efficiency;
(4) Information visualization of goods subject to more and more attention as an important part of the inventory management. Visualization of the cargo information allows the staff on the entire network of chain an intuitive grasp of the activity status information of the items in the business and logistics process, to help completing the tasks and real-time fault-dimensional. Visual clarity of the inventory information will become a major trend of the inventory management

3.3 Design of Intelligent Warehouse Management System Based on the Internet of Things (IOT)

Intelligent warehouse management program based on the Internet of Things (IOT), in terms of hardware, consists of handheld electronic label, barcode readers and fixed readers; in terms of software, consists of a host management system and shelf electronic tags. The overall structure of the system is shown in Figure 1.

Handheld electronic label bar code reader is recording equipment in the operation of in/out of storage in warehouse management system, and it is also the core equipment of the whole system. Staffs use handheld reader to scan the bar code of goods, and then write the storage information to the RFID tags to complete the inventory of goods.

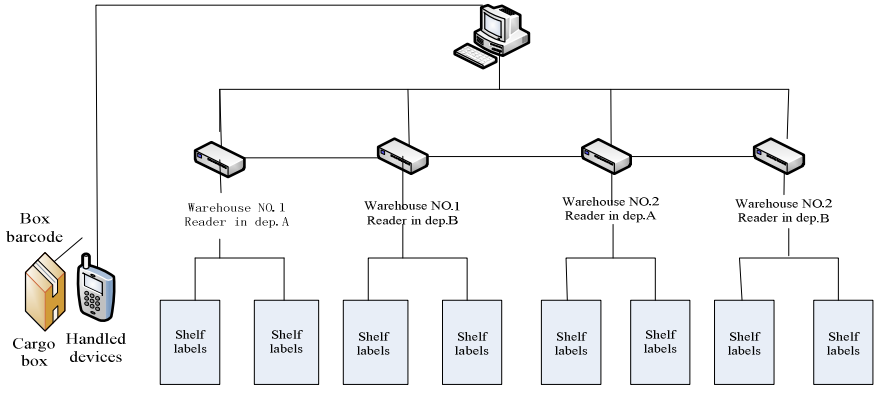


Fig. 1. An overall system structure

The role of fixed reader is connecting the management host system to read real-time inventory information and Update the database. For the large warehousing company, managers can use multiple fixed readers to divide warehouse into several partitions, and each fixed reader is responsible for a partition, so you can have electronic tags on each shelf level and well-organized monitoring of the region.

Host management system consists of information management module, the storage operation management module, handheld device management module, and the inventory information query module, inventory alarming module, equipment and label detection module. Each module works together to help the intelligent warehouse information management system to achieve the all-weather multi-faceted regulation of the entire warehouse.

In warehouse, each region has a number of shelves, and Electronic tags on the shelf are used to record the inventory of goods of this shelf; when goods in the warehouse changes, the staff use handheld readers to scan the goods on the shelves, and new information is available on the region's fixed-reader.

Common bar code is parallel line pattern arranged by black bars and white bars which are reflective varying greatly. The bar code can express a group of independent information, including the marked items producer, manufacturer, product name, production date, book words, the message start and end location, category, date, etc. Thus, it has been widely used in many areas, such as the circulation of commodities, library management, postal management, banking system, etc.

In Inventory Management, the bar code of the goods is custom information of the goods within the system, and it records the cargo storage area number, storage time, the serial number etc.

3.4 The Advantages of IOT in Warehouse Management

(1) IOT can help to quickly tally

Using IOT, just read the bar code of the goods or goods shelves to get the goods information of warehouse number, location and previous logistics process information

to quickly complete the tally process. Compared with the traditional tally, IOT reduces the operator's manual labor and improves the quality of speed.

(2) IOT can simplify the process of inventory of goods

Using the Internet of Things (IOT) technology, we can achieve to read data from remote multiple goods and sent it to the system host. This can greatly simplify the operations to complete the inventory of goods. Compared with the traditional inventory of goods, IOT eliminates manual inspection error rate, reducing the workload and decreasing the time of take stock, to increase the number of inventory.

(3) IOT can enhance the level of warehouse automation management

It is able to automate check the goods out of storage non-contacting with IOT. Compared with traditional management, it is not required to open the boxes to get the cargo information, so significantly improved storage speed. At the same time, barcode detailed cargo information to help businesses strengthen the meticulous management.

4 Conclusions

Intelligent warehouse information management system based on the IOT get a lot of information of different goods through the use of advanced sensor technology, and on the basis of Internet and cloud computing technologies, achieves intelligent processing and control of the goods in in/out of storage and cargo handling process, improving the efficiency of warehouse management and reduce the error rate for enterprises. Also reduces costs and the workload of staff. More and more enterprises will adopt it to get greater assistance for the socio-economic and rapid development of enterprises.

References

[1] Kim, Y.B., Moon, M., Keunhyukyeom: A framework for rapid development of RFID applications. Department of computer engineering, pp. 609–735. Pusan National University, Pusan (2008)

[2] Sabbaghi, A., Vaidyanathan, G.: Effectiveness and efficiency of RFID Technology in supply chain management strategic values and challenges. Journal of Theoretical and Applied Electronic Commerce Research 3(2)(8), 71–81 (2008)

[3] Liu, M.R., Zhang, Q.L., Ni, L.M., Tseng, M.M.: An RFID-Based Distributed Control System for Mass Customization Manufacturing. In: Cao, J., Yang, L.T., Guo, M., Lau, F. (eds.) ISPA 2004. LNCS, vol. 3358, pp. 1039–1049. Springer, Heidelberg (2004)

[4] Liang, G.: Development and application of radio frequency identification technology and the Internet of Things. Jinka Program (10) (2004)

[5] Meng, X.-N.: RFID applied to supply chain management, p. 12. Beijing Jiaotong University (2006)

[6] Xiao, Y.-Y.: RFID applied to retailer logistics management, p. 16. Southwest Jiaotong University (2008)

[7] Feng, L.: RFID applied to supply chain logistics management, p. 19. Southwest Jiaotong University (2009)

[8] Lou, W.-I., Kong, F.-S., Lou, W.-Z.: Data warehouse and knowledge discovery. Computer Engineering and Applications (10) (2000)

Design and Implementation of Visual C + + Based Digital Image Processing System

Hu Liang

Jiangxi Science and Technology Normal University, Nanchang, Jiangxi, 330013, P.R. China
hliang1972@126.com

Abstract. Visual C + + is object-oriented visual programming language. This article discusses the use of Visual C + + development tools to quickly build a digital image processing system, and the related issues on the input, display and processing of the digital image. Practice has proved that the system has real time, friendly interface and simple operation, suitable for building experimental platform of digital image processing for the continuous development of digital image processing technology and universities.

Keywords: Digital image processing, image enhancement, vc.

1 Introduction

Digital image processing, is to handle image by use of computers and other digital technology, in order to achieve improved image quality and target recognition as an intended purpose, including image transform, image coding, image enhancement, image restoration, image segmentation, with the computer hardware and software technology. Digital image processing has played an important role in an academic research and engineering applications. As an object-oriented visual programming, VC + + is portable and interface development from the bottom to the user-oriented software applications and other software development.

Based on this consideration, this paper presents a VC in the windows platform with a digital image processing system, including image input, display and processing. Using object-oriented technology, a variety of image processing algorithms encapsulate the implementation process, with running speed, portability and simple operation features.

2 System Components and Design

2.1 The Overall Framework

System is designed in accordance with the Windows interface. The overall framework includes basic image input, image processing, image output. Image processing is a key part of the system, including orthogonal transformation, point transformation, image enhancement, image restoration, image coding and image segmentation. The system interface is simple and easy to operate.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 209–214.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2.2 The Image Input and Display

The form of the input image file is BMP, each of which includes the bitmap file header, bitmap information header, the palette information and bitmap data. In the programming, the function ReadDIBFile (file) is to achieve graphical input, and SetPixel (i, j, RGB (Red, Green, Blue)) functions for image display, which is the RGB color index values of an image. When you open File, the value of the program is to read, and then find the value of the final color and images displayed on the screen dot.

2.3 Image Processing

Image processing is the main content of the system, including the following:

2.4 Orthogonal Transformation

Orthogonal transformation module consists of discrete Fourier transform module and the discrete cosine transform module. Discrete Fourier transform is often used in image enhancement and image compression, etc., which a complex arithmetic operation is leaving slower. Discrete cosine transform are based on real orthogonal transformation. The transformation matrix contains only cosine components can be eliminated Fourier transform Gibbs effect shortcomings. Pros and cons of one-dimensional Fourier transform formula[1] as follows:

$$F(U)=\sum_{n=0}^{N-1} f(x)e^{-j\frac{2\pi}{N}ux}(u=0,1,2\cdots N-1)$$

$$f(x)=\frac{1}{N}\sum F(U)e^{j\frac{2\pi}{N}ux}(x=0,1,2\cdots N-1)$$

The frequency domain is a complex function, whose absolute value represents the spectrum or amplitude spectrum. Pros and cons of one-dimensional cosine transform formula is as follows:

$$F(U)=C(u)\sqrt{\frac{2}{N}}\sum_{n=0}^{N-1} f(x)\cos\frac{(2x+1)u\pi}{2N}(u=0,1,2\cdots N-1)$$

$$f(x)=\sqrt{\frac{2}{N}}\sum_{n=0}^{N-1} C(u)F(u)\cos\frac{(2x+1)u\pi}{2N}(x=0,1,2\cdots N-1)$$

$$\text{in}\quad C(u)=\begin{cases}\frac{1}{\sqrt{2}} & u=0\\ 1 & other\end{cases}$$

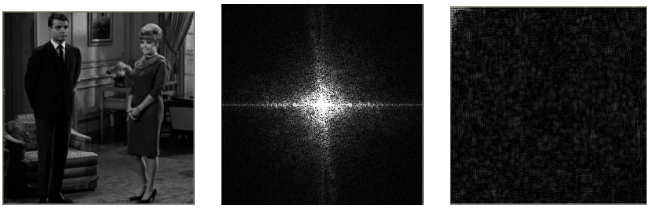


Fig. 1. The results of orthogonal transformation image: (a) couple Figure; (b) Fourier transform; (c) cosine transform

To apply two orthogonal transformations above in the digital image for image processing, two-dimensional orthogonal transformation should be put forward. Two-dimensional orthogonal transform operation is divided into horizontal and vertical directions on the one-dimensional orthogonal transform operation. Image processing results shown in Figure 1.

2.5 Point Operations

Histogram display and histogram equalization: Histogram shows the distribution characteristics of the gray image. Histogram equalization is to transform the original plans for the uniformly distributed histogram form, which can increase the dynamic range of pixel gray value. The results shown in Figure 2.

Fig. 2. Image after equalization

Gray-scale transformation

A new image $g = T(f)$ is changed the contrast of the image by transforming image f to T, thereby improving the human visual effects. The system curve is given by the dialog box, moving its control points to get the desired transformation.

2.6 Image Enhancement

Once the quality of the image is affected by noise, the image enhancement methods can be degraded only interesting part of the image to improve prominent image features.

Image smoothing: Image smoothing is achieved by low-pass filter. System uses a low-pass spatial filtering technology, by which low-pass convolution distinguishes between the different positions of pixels of the output pixel value by defining areas of the template. System settings through the dialog box allows users to smooth templates. custom template can define the coefficients of the template, height, width, the center element location, and the values of the template elements. Figure 3:

The value of the filter: This is a non-linear signal processing methods, in which moving window by an odd point, the value of the center of the window is replaced with the value of each point in the window .

Image sharpening: Processing is to enable the sharpening blurred images . Commonly, there are the gradient and Laplacian sharpening methods.

Gradient sharpening:

After gradient sharpening, image function $f(x, y)$ is changed into the image $g(x, y)$ as follows:

$$g(x,y) = \begin{cases} G[f(x,y)] & G[f(x,y)] \geq T \\ f(x,y) \end{cases}$$

Among them, T is a non-negative threshold set by the dialog $G[f(x,y)]$ box. is a gradient vector at the point (x, y) in a function of the image $f(x,y)$, defined as:

$$G[f(x,y)] = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \\ \dfrac{\partial f}{\partial y} \end{bmatrix}$$

Laplacian sharpening:

Laplace operator can be converted to template operation as Laplace operator template. The template is commonly:

$$H = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8* & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Compared with the gradient sharpening, the template is weak in noise reduction, thus it is necessary to smooth the image first if sharpening the Laplace transform template.

Pseudo-color Enhancement

By Pseudo-color enhancement, the original gray image areas with different gray values is assigned different colors to more clearly distinguish them.

2.7 Image Restoration

To improve the degraded image restoration is one of the image degraded mainly for specific reasons, in which compensates for lower quality factors so that the improved image is as much as possible approximation of the original image. The implementation method is to use some prior knowledge of degradation, the establishment of a mathematical model of degradation, and then reverse the deduction based on the model calculations in order to restore the original landscape images.

Maximum a Posteriori (MAP) Estimation

This is a posteriori estimation based on Markov fields and Gibbs then field theory, in which the image is regarded as a non-stationary random field, and the image model is then represented as a smooth process in which unstable mean value is fluctuate zero mean and then Diego on behalf of the conditions of model calculation.

Inverse Filtering Method
Based on signal processing in the inverse filtering method for image restoration.

The Maximum and Minimum Filter
Minimum filtering method is to use the maximum extraction recovery of target images.

Image Coding
The purpose of coding for data compression is to eliminate redundant data. System uses a run length encoding method, a continuous string with the same value with the string length, and a representative value instead. Encoding process is to follow the coding standard JPEG image to divide the image into 8 * 8 pieces which is quantized after DCT transformation. It recovers with quantization table and calculates the mean square error with the original image.

$$H = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8* & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Image Segmentation
Image segmentation is a disjoint process by which the image is divided into several small area which is a common property of some sense of the connected set of pixels. System uses the gray threshold segmentation scheme, and its expression is

$$g(x, y) = \begin{cases} 0 & f(x, y) < T \\ 255 & f(x, y) \geq T \end{cases}$$

Is a non-negative threshold. In order to obtain a suitable threshold , the system conducts a connected component analysis, and pseudo-color coded according to tag before carrying out image segmentation. The results shown in Figure 3

Fig. 3. Segmentation: (a) the image connected component labeling. (b) Segmentation results after

3 Conclusion

The developed system is to process image with good quality by a simple operation. Because the system is with multi functions, it just demonstrates some of the more representative examples for common image processing work to achieve,. The system

adopts the object-oriented software design methods, making the algorithm of each module algorithm library as the independent application environment. So the software system is easy to maintain and extend. In this system, there is further development in more complex and more specific image processing methods.

References

[1] He, D.: Digital Image Processing, pp. 54–158. Xidian University Press, Xi'an (2003)
[2] He, B., Ma, T., Wang, Y. (eds.): Visual C + + Digital Image Processing. People's Posts and Telecommunications Press, Beijing (2002)
[3] Song, K., Ruining, Liu, W.: The Collection of Visual C + + Development Technology. People's Posts and Telecommunications Press, Beijing (2007)

A New Multi-chaos Based Image Encryption Algorithm

Nan Lin, Xiaofeng Guo, Ping Xu, and Yuqin Wang

Zhengzhou Institute of Information Science and Technology Zhengzhou, China
linnan_2011@126.com

Abstract. A new multi-chaos based image encryption algorithm is proposed in this paper. Four chaotic mappings are involved in the encryption algorithm. The renew function of CML mapping is determined by the status of Chebyshev mapping. The encryption of pixel value and the permutation of pixel position are obtained with CML and Chebyshev iteration. It is from analysis and experimental results that the encryption algorithm possesses higher security.

Keywords: Multi-Chaos, Chebyshev Mapping, Cml, Image Encryption.

1 Introduction

With the rapid development of Internet and multimedia data processing, protection of multimedia data against illegal copying and distribution has become extremely important. To meet this requirement, many new encryption algorithms have been proposed. The chaos-based cryptographic algorithm have suggested some new ways to develop efficient multimedia encryption schemes[1-10], which have been motivated by the chaotic properties(pseudorandom property, non-periodicity and topological transitivity and extreme sensitivity to initial conditions and parameters).

In [3] and [4], the positions of image pixels are permutated by an image total shuffling matrix, and the pixel values of the permuted image are encrypted by a hyper-chaotic system. The security weakness of the algorithm in [3] is analyzed in [5,6], and the algorithm is broken by two attacks proposed in [5,6].

In this paper, a new image encryption algorithm is proposed to solve the problem of the algorithm in [3]. In the new image encryption algorithm, four chaotic systems are applied. The partial renew function of CML is chosen according to the state of Chebyshev mapping. Then the encryption of the positions and values of the plain image are performed based on CML and Chebyshev iterations. Analysis and experimental results show the security of the new image encryption algorithm.

2 Chaotic Systems

In the new encryption algorithm, four chaotic systems are applied, which are Chebyshev mapping, Sin mapping, Cubic mapping and 2D coupled map lattice (CML). The functions of them are as follows:

Chebyshev mapping:

$$x_{n+1} = \cos(a \cos^{-1} x_n), \tag{1}$$

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 215–221.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Sin mapping:

$$x_{n+1} = b\sin(\pi x_n), \tag{2}$$

Cubic mapping:

$$x_{n+1} = \lambda x_n(1-x_n^2), \tag{3}$$

CML:

$$y_{n+1}^{i,j} = (1-\varepsilon) f(y_n^{i,j}) + \frac{1}{2}\varepsilon[f(y_n^{i+1,j}) + f(y_n^{i,j+1})] \tag{4}$$

where 2≤a, -1≤xn≤1, b=0.99, 0<xn<1, λ=2.59, 0<xn<1. $0 < \varepsilon < 1$, $0 < y_n^{i,j} < 1$, $0 \le x \le 1$, $f(x)$ is the partial renew function of CML. The periodic boundary conditions of CML are $y_n^{i+H,j} = y_n^{i,j}$, $y_n^{i,j+W} = y_n^{i,j+1}$.

In this algorithm, the partial renew function $f(x)$ of CML is determined by Chebyshev mapping, Sin mapping and Cubic mapping. Sin mapping or Cubic mapping is chosen according to the state of Chebyshev mapping as the partial renew function $f(x)$, the detailed chosen method is shown in the following formula:

$$f(x_n) = \begin{cases} b\sin(\pi x_n) & if\ x_{chebyshev} < 0.5 \\ \lambda x_n(1-x_n^2) & if\ x_{chebyshev} \ge 0.5 \end{cases}, \tag{5}$$

Where $x_{chebyshev}$ is the state of Chebyshev mapping.

3 Multi-chaos Based Image Encryption Algorithm

Assume that a plian image is $\boldsymbol{P} = (p_{i,j})_{N\times N}$, where $p_{i,j} \in \mathcal{P}$, $\mathcal{P} = \{0,1,2,\cdots,255\}$, $\mathcal{N} = \{1,2,\cdots,N\}$. Let $\boldsymbol{A} = (a_{i,j})_{N\times N}$ denote the DCT coefficient matrix of plaintext after pre-coding (for example block partitioning), DCT transformation and quantization. Two positive integer M_1, M_2 ($M_1, M_2 \le 1024$) are chosen as the generation parameters of the keystream, which are also part of the secret key.

Before encryption, the DCT coefficient matrix is divided into $k-1$ blocks of size $T\times T$, then $\boldsymbol{A} = \{A^t\}_{t=1}^{k-1}$, where $A^t = (a_{i,j}^t)_{T\times T}$, $t = 0,1,2\cdots,k-1$, $i, j = 0,1,2\cdots,T-1$. For the t th block $A^t = (a_{i,j}^t)_{T\times T}$, $t = 0,1,2\cdots,k-1$, the encryption procedure are detailed as follows:

1. Pre-iterate Chebyshev mapping by initial value x_0 for n_0 times, and denote the new state of Chebyshev mapping by x_0;
2. Interate Chebyshev mapping by x_0, and denote the state by x_t; then determine the partial renew function $f(x)$ of CML with x_t and Eq.错误！未找到引用源。;
3. Iterate CML with the row initial vector $IV = (x_{0,1}, x_{0,2}, \cdots, x_{0,T})$ and $f(x)$, and denote the obtained $T\times T$ state matrix by $R = (r_{i,j})_{T\times T}$;

4. Generate the matrix $S=(s_{i,j})_{T\times T}$ which is the keystream of this algorithm by quantization $R=(r_{i,j})_{T\times T}$ with the following equation:

$$s_{i,j}=\begin{cases}\left\lfloor r_{i,j}\times 2^{M_1}\right\rfloor \bmod M_1 & i=j=0\\ \left\lfloor r_{i,j}\times 2^{M_2}\right\rfloor \bmod M_2 & else\end{cases}. \tag{6}$$

5. Encrypt $A^t=(a^t_{i,j})_{T\times T}$ by the matrix $S=(s_{i,j})_{T\times T}$ with

$$c^t_{i,j}=\begin{cases}a^t_{i,j}\oplus s_{i,j} & i=j=0\\ sign(a^t_{i,j})[\lfloor a^t_{i,j}\rfloor\oplus s_{i,j}] & else\end{cases}, \tag{7}$$

and the cipher block $C^t=(c^t_{i,j})_{T\times T}$ of $A^t=(a^t_{i,j})_{T\times T}$ is obtained, where $sign(a^t_{i,j})=\begin{cases}1 & a^t_{i,j}>0\\ -1 & a^t_{i,j}\le 0\end{cases}$.

6. If $t<k-1$, then let $x_0=x_t$, go to 1, carry out 2 to 5; otherwise, continue;
7. Let $X=\{x_t\}_{t=0}^{k-1}$, rearrange sequence $\{x_0,x_1,\cdots x_{k-1}\}$ from small to large, and sequence $\{x_{r_0},x_{r_1},\cdots x_{r_{k-1}}\}$ is obtained. Obviously, there is a permutation function σ: $\sigma(i)=r_i$, which is shown as follows:

$$\begin{pmatrix}0 & 1 & 2 & \cdots & k-1\\ \sigma(0)=r_0 & \sigma(1) & \sigma(2) & \cdots & \sigma(k-1)\end{pmatrix} \tag{8}$$

8. Permute $\boldsymbol{C}=\{C^t\}_{t=1}^{k-1}$ by block with permutation function σ, and obtain the cipher image $\boldsymbol{D}=\{D^t\}_{t=1}^{k-1}$, $D^t=C^{\sigma(t)}$, that is

$$\begin{pmatrix}C^0 & C^1 & C^2 & \cdots & C^{k-1}\\ C^{\sigma(0)}=C^{r_0} & C^{\sigma(1)} & C^{\sigma(2)} & \cdots & C^{\sigma(k-1)}\end{pmatrix} \tag{9}$$

4 Decryption Algorithm

The decryption procedure is the inverse procedure of the encryption algorithm. In the decryption procedure, the sequence $X=\{x_t\}_{t=0}^{k-1}$ is first generated with Chebyshev mapping, and the inverse permutation function σ^{-1} is obtained, then the ciphertext is first inversely permuted by the inverse permutation function σ^{-1}; thirdly, the data after inverse permutation is divided into $k-1$ blocks of size $T\times T$, and finally the plaintext is obtained by decrypting the data block by block with matrix $S=(s_{i,j})_{T\times T}$ which is generated from CML.

5 Key Space Analysis

In this algorithm, the secret key is $Key=\{x_0,M_1,M_2,IV\}$, where x_0 is the initial value of Chebyshev mapping, $IV=(x_{0,1},x_{0,2},\cdots,x_{0,T})$ is the row initial value of CML. The key space of the algorithm is $K=K(x_0)\times K(IV)\times K(M_1)\times K(M_2)$. If double-precision

floating-point number of 64bit of IEEE is applied in the encryption algorithm, since x_0 and $IV = (x_{0,1}, x_{0,2}, \cdots, x_{0,T})$ can not equal to 0 and 1, then $K(x_0) = 2^{52} - 2$, $K(IV) = (2^{52} - 2)^T$. If the block size $T \times T = 8 \times 8$, and let $K(M_1) = K(M_2) = 1$ (that is M_1, M_2 are two secret constant) then the key space of this algorithm is $K = K(x_0) \times K(IV) \times K(M_1) \times K(M_2) = (2^{52} - 2) \times (2^{52} - 2)^8 \approx 2^{468}$. So the keyspace of this algorithm is considered to be large enough for practical applications.

6 Experiments

In this section the encryption and decryption experiments of a 256×256 pixel image "Lena.bmp" are performed under Matlab6.5 on a PC with 2GB RAM. In the experiments, the block size is $T \times T = 8 \times 8$, the initial value of Chebyshev mapping $x_0 = 0.123456$, the row initial value of CML is $IV = (0.152, 0.1735, 0.1653, 0.343, 0.2536, 0.635, 0.472, 0.2397)$, and the parameter $M_1 = M_2 = 1024$.

(a) Plain image (b) Cipher image

Fig. 1. Plain image and cipher image

6.1 Key Sensitivity Test

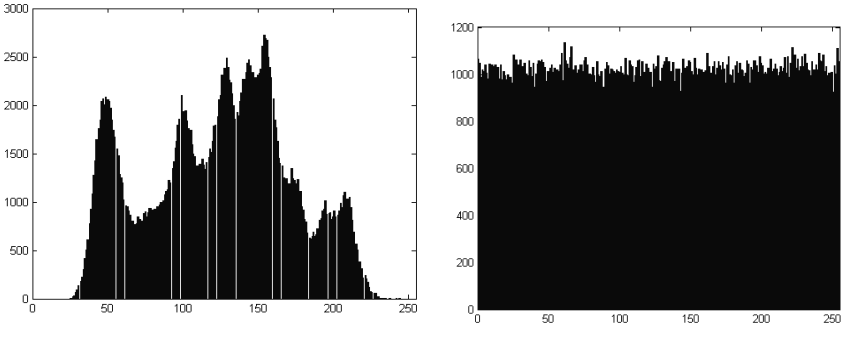
In this subsection, the key sensitivity of the algorithm is test. Plain image "Lena.bmp" is encrypted by using the test key $x_0 = 0.123456$, $M_1 = M_2 = 1024$, $IV = (0.152, 0.1735, 0.1653, 0.343, 0.2536, 0.635, 0.472, 0.2397)$, and decrypted with all the parameters unchanged except for the initial value of Chebyshev mapping $x_0 = 0.123457$. The experimental results are shown in Figure 2, where Figure 2(a) is the decrypted image with the tiny changed key, and Figure 2(b) is the decrypted image with the correct key. We can find that a slight change of the key will generate a completely different decryption result and can't get the correct plain image, that is, the encryption scheme is sensitive to the secret key.

(a) decryption results with wrong key (b) decryption with correct key

Fig. 2. Key sensitivity text results

6.2 Image Statistic Characteristic

The histograms of the plain image and the cipher image are evaluated in this subsection. The histogram of image "Lena.bmp" is shown in Figure 3(a), and that of the cipher image is shown in Figure 3(b). According to the experimental results, the histogram of the cipher image is significantly different from that of the plain image and is fairly uniform.



(a) Histogram of plain image (b) Histogram of cipher image

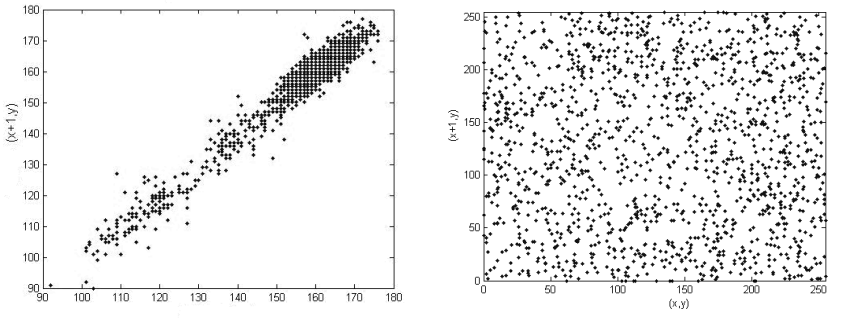
Fig. 3. Histogram of images before and after encryption

6.3 Analysis of Correlation of Two Adjacent Pixels

In this section, some simulations are carried out to test the correlation between two vertically adjacent pixels, two horizontally adjacent pixels and two diagonally adjacent pixels respectively. Firstly, 4096 pairs of two adjacent pixels are chosen randomly from the plain image and the cipher image, and the correlation distribution of two vertically adjacent pixels in the plain image and that in the cipher image is tested. The results are shown in Figure 4. Then the correlation coefficient of each pair is calculated by Eq (10). And the calculation results of correlation coefficients are shown in Table 1. We can find that the correlation of two adjacent pixels has decreased obviously, that is, the close correlation property between pixels in plain image has been removed.

$$r_{xy} = \mathrm{cov}(x, y)/(\sqrt{D(x)}\sqrt{D(y)}), \tag{9}$$

where $\mathrm{cov}(x, y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - E(x))(y_i - E(y))$, $E(x) = \frac{1}{N}\sum_{i=1}^{N} x_i$, $D(x) = \frac{1}{N}\sum_{i=1}^{N}(x_i - E(x_i))^2$.



(a) Correlation in the plain image (b) Correlation in the cipher image

Fig. 4. Correlations of two vertically adjacent pixels in the image before and after encryption

Table 1. Correlation coefficients of two adjacent pixels in images before and after encryption

	Horizontal	Vertical	Diagonal
Plain image	0.988378	0.988778	0.983117
Cipher image	0.035188	0.046506	0.060353

7 Conclusions

A new multi-chaos based image encryption algorithm is proposed in this paper. Four chaotic mappings are involved in the encryption algorithm. The renew function of CML mapping is determined by the status of Chebyshev mapping. The encryption of pixel value and the permutation of pixel position are obtained with CML and Chebyshev iteration. It is from analysis and experimental results that the encryption algorithm possesses higher security.

References

[1] Mao, Y.B., Chen, G.: Chaos-based image encryption. In: Handbook of Computational Geometry for Pattern Recognition, Computer Vision, Neurocomputing and Robotics. Springer, New York (2003) (in press)

[2] Li, S., Chen, G., Zheng, X.: Chaos-based encryption for digital images and videos. In: Furht, B., Kirovski, D. (eds.) Multimedia Security Handbook, ch. 4, pp. 133–167. CRC Press (2004)

[3] Gao, T., Chen, Z.: A new image encryption algorithm based on hyper-chaos. Phys. Lett. A 372, 394–400 (2008)

[4] Gao, T., Chen, Z.: Image encryption based on a new total shuffling algorithm. Chaos, Solitons and Fractals 38, 213–220 (2008)

[5] Rhouma, R., Belghith, S.: Cryptanalysis of a new image encryption algorithm based on hyper-chaos. Phys. Lett. A (2008) (in press), doi:10.1016/j.physleta.2008.07.057

[6] Ge, X., Liu, F., Lu, B., Yang, C.: Improvement of Rhouma's Attacks on Gao Algorithm. Physics Letters A 374, 1362–1367 (2010)
[7] Pareek, N.K., Patidar, V., Sud, K.K.: Image encryption using chaotic logistic map. Image and Vision Computing 24(9), 926–934 (2006)
[8] Behnia, S., Akhshani, A., Mahmodi, H.: A novel algorithm for image encryption based on mixture of chaotic maps. Chaos, Solitons and Fractals 35(2), 408–419 (2008)
[9] Kwok, H.S., Tang, W.K.S.: A fast image encryption system based on chaotic maps with finite precision representation. Chaos, Solitons and Fractals 32, 1518–1529 (2007)
[10] Arroyo, D., Li, C.Q., Li, S.J., Alvarez, G., Halang, W.A.: Cryptanalysis of an image encryption scheme based on a new total shuffling algorithm. Chaos, Solitons and Fractals (available online November 5, 2008) (corrected proof, in press)

Performance of an Optical Packet Switch with Limited Wavelength Converter

Jie Li

Shanghai Research & Development Center, ZTE Corporation, Shanghai, 201203, China
ciomlijie@hotmail.com

Abstract. In this paper, a novel optical packet switch is proposed, which uses limited wavelength converters and recirculation fiber delay lines to resolve optical packet contentions. A control algorithm is proposed to schedule the optical packets. Performance of the proposed switch is evaluated by means of simulation experiments. Packet loss rate and mean delay time of the proposed switch are analyzed in detailed.

Keywords: limited wavelength converter, optical packet switch, optical buffer, fiber delay line.

1 Introduction

Optical packet switching (OPS) [1-9] is an ideal technology to bridge the gap between IP network and optical transportation network. Contention resolution is one of the critical issues in optical packet switching. In general, there are three ways to resolve optical packet contentions, i.e., optical buffering, wavelength conversion and deflection routing [1,2]. Among these three contention resolutions, optical buffer, which always consists of fiber delay lines (FDLs) with present technology, is most effective and has been extensively studied [1-9]. The most commonly discussed buffering schemes are input buffering, output buffering and recirculation buffering. The input buffering is never proposed for purely optical implementation, primarily because of its poor performance[2]. The output buffering can achieve optimal throughput performance and it has been widely applied to optical packet switch architectures [4,5]. However, the structure with output buffers still can't make full use of the output FDLs because at a given instant not all the FDLs are used. In [3], Karol proposes a SMOP switch architecture, which only uses recirculation FDLs to resolve packet contentions. That structure can make full use of the FDLs and achieve good performance, however, large buffered delay is also introduced. In order to reduce the buffered delay, J. Yang [7] propose a SMOP-TWC architecture with recirculation FDLs and tunable wavelength converters (TWCs) to resolve packet contentions. Compared with the SMOP architecture, the SMOP-TWC architecture can reduce the packet loss rate and buffered delay greatly. However, with present technology, the TWC is still an expensive device. Consequently, it is a tradeoff between the performance and the cost.

In this paper, another kind of wavelength converter, i.e., limited wavelength converter is adopted. A novel optical packet switch is proposed, which uses shared

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 223–229.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

recirculation FDLs and limited wavelength converters (SMOP-LWC) to resolve optical packet contentions. The performances of the SMOP-LWC are evaluated by simulation experiments.

This paper is organized as follows. In section 2, the proposed architecture is described. In section 3, the control algorithm is introduced. In section 4, performances of the architecture are analyzed in detail by simulation experiments. Finally, conclusions are drawn in section 5.

2 Switch Architecture

The considered optical packet switch (SMOP-LWC) is shown in Fig. 1. It has *N* input and output fibers. Each input and output fiber consists of *M* different wavelengths. Each wavelength represents a unique channel. At each input fiber, incoming optical packets are wavelength demultiplexed by DMUX shown in Fig. 1. In this paper, we assume that packets have a fixed length and their arrivals on each wavelength are synchronized on a time-slot basis, where a time slot is the time needed to transmit a single packet. Hereafter, the duration *T* of a time slot is assumed as the time unit.

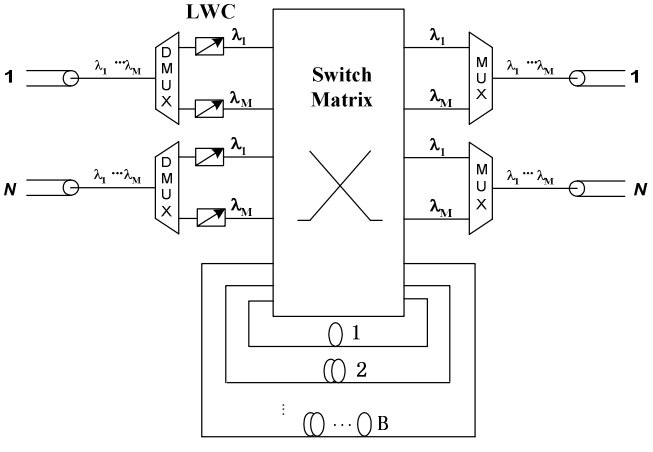


Fig. 1. The architecture of SMOP-LWC

In order to resolve optical packet contention, a limited wavelength converter (LWC) is equipped on each input wavelength channel. For the limited wavelength converters, wavelength $\lambda_j (j \in \{1,\ldots,M\})$ can only be converted into part of the other wavelengths. The conversion ability of LWC is determined by conversion range R_c, which means the number of converted wavelength for λ_j. When the conversion range is equal to R_c, wavelength λ_j can be converted to the following wavelength set $\Lambda_j(R_c)$:

$$\Lambda_j(R_c)=\begin{cases}\{\lambda_{\max}(1, j-R_c/2),\cdots,\lambda_j,\cdots,\lambda_{\min}(j+R_c/2, M)\} & \text{if } R_c \text{ is even}\\ \{\lambda_{\max}(1, j-\lfloor R_c/2\rfloor-1),\cdots,\lambda_j,\cdots,\lambda_{\min}(j+\lfloor R_c/2\rfloor, M)\} & \text{if } R_c \text{ is odd}\end{cases} \quad (1)$$

where M is the number of wavelength channels and $\lfloor X \rfloor$ indicates the largest integer less than or equal to X.

In addition, the switch architecture is equipped with B shared recirculation FDLs. Each FDL is a WDM buffer, which can buffer M optical packets on M different wavelengths at one time slot. The FDLs are arranged in degenerate form [6], i.e., the delays achieved by FDLs uniformly cover the entire range of buffer depth, from T to BT, with increment of one unit.

3 Control Algorithm Description

For the proposed switch, two kinds of packets, i.e., newly arriving packet characterized by Pn(i,k,λj) and recirculation packet characterized by $P_r(m,k,\lambda_j)$, may contend for a given output port. In $P_n(i,k,\lambda_j)$ and $P_r(m,k,\lambda_j)$, $i(i \in \{1,\cdots,N\})$ identifies the input fiber, $m(m \in \{1,\cdots,B\})$ identifies the FDL, $k(k \in \{1,\cdots,N\})$ identifies the output fiber, and $\lambda_j(j \in \{1,\cdots,M\})$ identifies the wavelength which carries the packet.

When a new packet $P_n(i,k,\lambda_j)$ arrives from input fiber i, the packet will then be switched to destination output fiber k if wavelength λ_j on output fiber k is free. However, if λ_j on output fiber k is occupied, the packet will be blocked. In this case, if there is an available wavelength λ_c $(\lambda_c \in \Lambda_j(R_c))$ in output fiber k, through wavelength conversion, $P_n(i,k,\lambda_j)$ will be transmitted on the available wavelength λ_c in output fiber k. Otherwise, $P_n(i,k,\lambda_j)$ will then try to be routed to the pool of FDLs for temporary storage. In this case, if there is an available wavelength λ_c $(\lambda_c \in \Lambda_j(R_c))$ in the FDLs, $P_n(i,k,\lambda_j)$ will be buffered on the pool of FDLs. At this time, FDL 1 will be searched first. If there is an available wavelength λ_c on FDL 1, $P_n(i,k,\lambda_j)$ will be buffered on the first fit available wavelength λ_c on FDL 1. If there is no available wavelength on FDL 1, the next FDL (i.e., FDL 2) will be searched until the longest one (i.e., FDL B). Subsequently, packet $P_n(i,k,\lambda_j)$ stored in the FDLs will recirculate back to a FDL input port for further attempts (no recirculation loops restriction). If there is no available wavelength λ_c $(\lambda_c \in \Lambda_j(R_c))$ in the FDLs, $P_n(i,k,\lambda_j)$ will be dropped.

For recirculation packet $P_r(m,k,\lambda_j)$, it will be switched to destination output fiber k if wavelength λ_j on output fiber k is available. Otherwise, it would be buffered on the former wavelength (i.e., λ_j) and in the former FDL (i.e., FDL m).

4 Numerical Results

The performances of the proposed SMOP-LWC optical switch are evaluated by simulation experiments. In the simulation experiments, the performance of SMOP-LWC is measured by packet loss rate (PLR) and mean buffered delay (MBD) versus the number of shared FDLs. Here, mean buffered delay is the mean time slots that a successfully delivered packet has to be buffered for.

For simplicity, it is assumed that the packet arrivals to each input wavelength channel comply with independent and identical Bernoulli process. That is, in any given time slot, the probability that a packet will arrive at a particular input wavelength channel is ρ_M and each packet is addressed to any of N given output ports with equal probability $1/N$. Based on the Bernoulli model, the probability of having R packet arrivals of a given output can be expressed as

$$P(R) = \binom{N \cdot M}{R} \left(\frac{\rho_M}{N}\right)^R \left(1 - \frac{\rho_M}{N}\right)^{N \cdot M - R} \tag{2}$$

where ρ_M represents the traffic load per wavelength channel at a given input.

In Fig.2, the packet loss rate of the SMOP-LWC versus the number of FDLs B is illustrated. In this set of simulation, the number of input/output port N and the conversion range R_c are respectively chosen to be 16 and 2. In addition, the traffic load per wavelength channel ρ_M is set to be 0.8. For comparison, we also evaluate the performance of the SMOP-TWC, where the number of shared FDLs B and the number of wavelength channels M are the same as those of SMOP-LWC. In Fig. 2(a), the number of wavelength channels M is equal to be 4, and in Fig. 2(b), the number of wavelength channels M is equal to be 8.

From Fig.2, it can be found that with the increase of the number of FDLs B, both the packet loss rate of SMOP-LWC and the packet loss rate of SMOP-TWC decrease. This can be explained as follows. No matter for SMOP-LWC or SMOP-TWC, with more FDLs, more blocked optical packet can be buffered. Consequently, the packet loss rate drops with the increase of B.

From Fig.2, it can also be found that when the same number of FDLs are used, SMOP-TWC always achieves better performance than SMOP-LWC. For example, as shown in Fig. 2(a), when 4 FDLs are adopted PLR of 1.448×10^{-2} is achieved with SMOP-LWC switch. However, for SMOP-TWC structure, with same number of FDLs, PLR of 7.796×10^{-3} can be achieved. In Fig. 2(b), when 4 FDLs are adopted, PLR of 4.659×10^{-3} is achieved with SMOP-LWC switch and PLR of 1.099×10^{-6} is achieved with SMOP-TWC structure. As we know the SMOP-TWC switch has better wavelength conversion capability than the SMOP-LWC switch. When the packet contentions occur, in the SMOP-TWC more blocked packet can either be transmitted with other available wavelength on the output fiber or be buffered with other available wavelength on the FDLs. Consequently, the SMOP-TWC switch achieves lower packet loss rate than the SMOP-LWC switch.

In Fig.2, an interesting phenomenon occurs. When four wavelength channels are used (as shown in Fig.2(a)), the difference of packet loss rate between the SMOP-TWC switch and the SMOP-LWC switch is small. However, when eight wavelength channels are adopted (as shown in Fig.2(b)), the difference of packet loss rate between SMOP-TWC switch and SMOP-LWC switch become obvious. This is also caused by the conversion capability of wavelength converters. For the simulation case in Fig.2, the conversion range R_c is equal to 2. That is to say, the input wavelength λ_j can be converted to the other two wavelengths in the wavelength set $\Lambda_j(R_c)$. For the case in Fig.2(a), the wavelength conversion capability of LWC is similar to that of TWC, thus their performances are almost the same. However, in the case of Fig.2(b), the

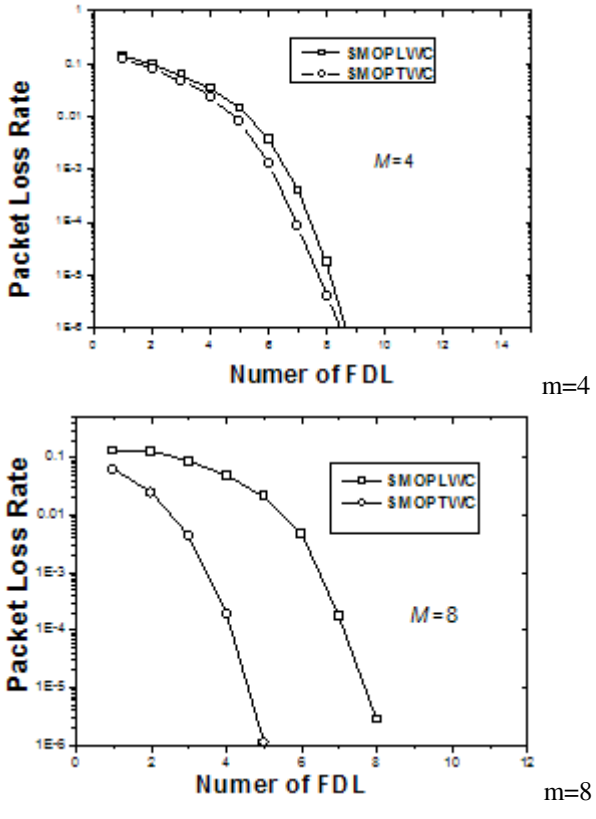


Fig. 2. PLR of SMOP-LWC versus the number of FDL, B

wavelength conversion capability of LWC is far lower than that of TWC, thus the SMOP-TWC switch achieves lower packet loss rate than the SMOP-LWC. Consequently, in order to achieve acceptable packet loss rate, the R_c should not be too small.

In Fig.3, the mean buffered delay (MBD) of the SMOP-LWC versus the number of shared FDLs B is also illustrated. The simulation condition is also the same as that in Fig.2. From Fig.3, it can be found that for both SMOP-LWC switch and SMOP-TWC switch, the same phenomenon occurs. When the number of FDLs B is small, with more FDLs the mean buffered delay (MBD) also increases. However, when B increases to a certain value, the MBD saturates to its maximum. This can be explained as follows. When the value of B is small, with more FDL more blocked optical packets are buffered in the long FDL, which introduces more delay time. Consequently, the MBD increases with the increase of B. When the value of B increases to a certain value, the packet loss rate become very low. In this case, the long FDLs are seldom used, thus with the increase of FDL the MBD does not increase.

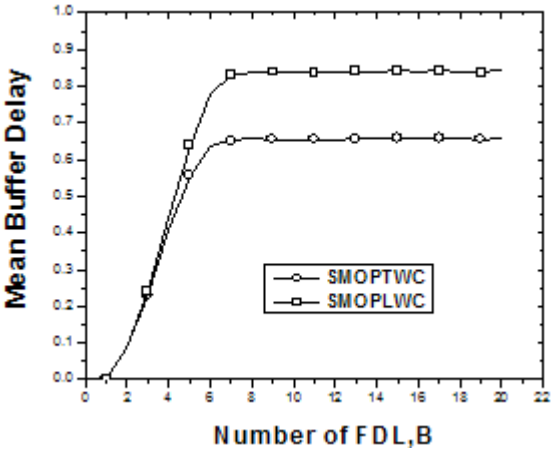


Fig. 3. PLR of SMOP-LWC versus the number of FDL, B

From Fig.3, it can also be found that with the same number of FDLs, the SMOP-LWC achieves larger MBD than the SMOP-TWC. The reason is obvious. Compared with the SMOP-TWC, the SMOP-LWC has less wavelength conversion capability. Among the successfully delivered packets in the SMOP-LWC, more packet contentions are resolved by FDLs, which introduces longer delay time.

5 Conclusions

This paper proposes a novel optical packet switch called SMOP-LWC, which uses limited wavelength converters and recirculation fiber delay lines to resolve optical packet contentions. In order to make full use of the FDLs and the LWCs, a control algorithm is also proposed. Packet loss rate and mean buffered delay of the SMOP-LWC versus the number of FDLs B is evaluated by means of simulation experiments. The simulation results show that the SMOP-LWC also can achieve acceptable packet loss rate and mean buffered delay. When the number of wavelength channels is small, the SMOP-LWC almost achieves the same performance with the SMOP-TWC. However, when the number of wavelength channels is large, the SMOP-LWC achieves higher PLR and larger MBD than the SMOP-TWC. In order to improve the performance of the SMOP-LWC, the conversion range of the limited wavelength converter, R_c should be increased.

References

[1] Yao, S., Mukherjee, B., Dixit, S.: Advances in Photonic Packet Switching: An Overview. IEEE Communication Magazine 38, 84–94 (2000)

[2] Hunter, D.K., Chia, M.C., Andonovic, I.: Buffering in optical packet switches. IEEE/OSA Journal of Lightwave Technology 16, 2081–2094 (1998)

[3] Karol, M.J.: Shared-memory optical packet (ATM) switch. Multigigabit Fiber Commun. Syst., 212–222 (1993)
[4] Danielsen, S.L., Mikkelsen, B., Joergensen, C., Durhuus, T., Stubkjaer, K.E.: WDM Packet Switch Architectures and Analysis of the Influence of Tunable Wavelength Converters on the Performance. IEEE/OSA Journal of Lightwave Technology 15, 219–227 (1997)
[5] Eramo, V., Listanti, M.: Wavelength converter sharing in a WDM optical packet switch: dimensioning and performance issue. Computer Networks 32, 633–651 (2000)
[6] Tancevski, L., Tamil, L., Callegati, F.: Nondegenerate buffers: an approach for building large optical memories. IEEE Photonics Technology Letters 11, 1072–1074 (1999)
[7] Yang, J., Zhang, H., Xiao, S.: Performance Study on Optical Packet Switch with Feedback Optical Buffer. Optical Communication Technology 30, 26–28 (2006)
[8] Yang, J., Chu, F., Lv, J.: Performances of an Optical Packet Switch with Fixed Wavelength Converter Arrays under Bursty Traffic. Applied Optics 48, 4885–4892 (2009)
[9] Eramo, V., Listanti, M., Germoni, A.: Cost Evaluation of Optical Packet Switches Equipped With Limited-Range and Full-Range Converters for Contention Resolution. IEEE/OSA Journal of Lightwave Technology 26, 390–407 (2008)

Time Petri Net-Based Software Reliability Analysis

Maoqiang Yang

Dept. of Computer Science, Zunyi Normal College, Zunyi, China
yangcap888@163.com

Abstract. Petri net is one of effective approaches to analysis of system reliability. In this paper, reliability of a control system for railway intersection exemplifies usage of Petri net. The reliability of the system is analyzed using Petri net. The example is intended to deepen our understanding of Petri net application.

Keywords: Petri net, software reliability, risk, key state.

1 Introduction

With the development of science & technology, the architecture of software system becomes increasingly complicated. As a result, the requirement for system reliability is raised. Reliability engineering has been throughout whole process of design, manufacture and use of a product. Its main techniques include reliability design, reliability test, reliability analysis and estimation etc.

2 Basic Concepts of Software Reliability

2.1 Definition of Software Reliability

Reliability is the probability with which product or system successfully finishes its functionality in prescribed environment, time and condition. In other words, reliability is the capability of finishing prescribed functionality in prescribed condition and time[1].

2.2 Characteristics of Software Reliability

Comparability under prescribed condition:
From the definition of reliability, the reliability of product depends on three regulations. Firstly, reliability varies greatly in different condition and environment; secondly, reliability with differently prescribed time is various; thirdly, reliability will be various if judgment of product functionality is different.

Time quality index:
Time quality of a product can be divided into two categories: t=0 quality and t>0 quality. The former one is the ex-factory quality and the latter one is the quality after the product is used for a period.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 231–235.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Usability:
Reliability has to do with lifespan of a product, but long lifespan is not necessary. Usability refers to functionality of a product within its life [2].

Estimation by statistics and sampling:
Estimation of reliability is carried out only if the product has used or a simulation has been done. In general, analysis and estimation of reliability are executed using statistics and sampling.

Reliability indices:
Generally, reliability of a product can be described by several indices. These indices form an index system. Which index is used depends on discrete product.

2.3 Analysis Methods of Software Reliability

Analysis approaches for software reliability consist of static analysis one and dynamic one. Static approach includes Static Fault Tree Analysis (SFTA) and Static Failure Modes and Effects Analysis (SFMEA). Dynamic approach includes Petri net and Real Time Logic and so on. In this paper, we focus on Petri net analysis approach.

3 Petri Net

3.1 Brief Introduction of Petri Net

Petri net is a mathematical representation of discrete parallel system. It is proposed by German C. A. Petri in 1960. It is well suited for asynchronous, concurrent computer system model. Petri net has not only strict mathematic formulation but also intuitively graphical notation. By means of its rich description methods and analytical techniques for various system behaviors, Petri net provides solid ground for basic conceptions in computer science.

3.2 Definition of Petri Net

Definition 1: A Petri net is a tetrad PN=(S, T, F, M), where S={s1, s2, …, sn} is a finite set of places; T={t1, t2, …, tn} is a finite set of transitions; $F \subseteq (S \times T) \cup (T \times S)$ is a set of directed arcs; M: $S \rightarrow \{1, 2, 3, \ldots\}$ is an initial marking. $S \cap T = \Phi$ and $S \cup T \neq \Phi$ [3].

3.3 Transition Rules of Petri Net

The execution of a Petri net can be defined as the transition relation on its tokens contained in its places. Firing a transition consumes tokens from each of its input places, and produces new tokens in each of its output places.

Definition 2: For a Petri net with marking M PN==(S, T, F, M) and a transition $t \in T$, transition t is enabled if $M(s) \geq 1$ for any s ($s \in S$ and $s \in \bullet t$), where M(s) is the number of tokens in place s in marking M, $\bullet t$ is the pre-set of t and $t \bullet$ is the post-set of t. If transition t is enabled in marking M, it will produce a new marking Mc(s), as follows[4]:

$$Mc(s) = \begin{cases} M(s)+1, & \text{if } s \in t\bullet - \bullet t \\ M(s)-1, & \text{if } s \in \bullet t - t\bullet; \\ M(s), & \text{otherwise.} \end{cases}$$

3.4 Security Analysis Method of Time Petri Net

Time Petri net is a Petri net with generalized time constraints. In addition to theoretical foundations of Petri net, Time Petri net has also extra time constraints [5]. In time Petri net, a time factor is attached on transition t. In one hand, transition t can be fired after at least α time units if it is enabled in marking M. In the other hand, if no other transitions disabled it during the period, transition t will be fired in at most β time units.

There are two steps in security analysis of a software system using time Petri net: firstly, modeling the software system using time Petri net; secondly, determining risk of the software system according to facts. Risk states can be classified into several risk levels, such as simply high level and low level. security analysis is just the analysis of reachability of high risk level. Generally, security analysis can be carried out by forward analysis method and backward analysis method. Forward analysis method determines whether there are high risk levels from initial state, while backward analysis method determines whether there exist conditions of high risk states during progress of state from high risk states.

Since all states of a certain scale of Petri net are usually very large, deciding reachability of high risk states is not realistic through checking all states. Doctor Leveson from Canifornia University proposed a key state method which can study reachability of high risk states without generating all reachable states.

Key state
A state μ is a key state if and only if

a) μ is a low risk state;

b) There exist two non-empty transitions S1 and S2, and two identities $\mu 1$ and $\mu 2$, such that $\delta^*=(\mu, S1)=\mu 1$ and $\delta^*=(\mu, S2)=\mu 2$, where $\mu 1$ is a high risk state and $\mu 2$ is a low risk state.

basic idea of key staes

a) If a high *risk state is* reachable*, there are necessarily key states in path from initial state to high risk states.*

b)For each *member of high risk set, generate its former state and check whether the former state is a key state. If it is not a key state, generate its former state again. If it is a key state, assign a higher priority to the path directed low risk states than that to the path directed high risk states when design a Petri net. In this way, it is guaranteed that high risk states are not reachable.*

4 Case Analysis

There is usually installed an automatic control system of a bulkhead door at a railway intersection. The control system is made of three components: train model, bulkhead

door model and computer model. The states of a train consist of coming, arriving, passing and leaving. Its initial state is coming state. The states of a bulkhead door consist of open and close. Its initial state is open state. The states of computer consist of waiting for arriving signal and waiting for leaving signal, where the initial state is waiting for arriving signal. A Petri net model of this control system is established as Fig. 1.

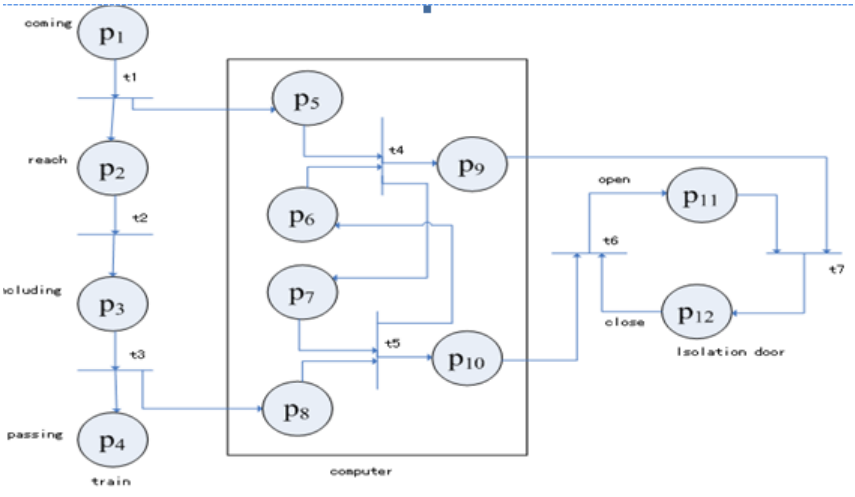


Fig. 1. System of petri net the model

From Fig. 1, it is very dangerous that bulkhead door is open (p11) when train is passing intersection (p3) i.e. a co-occurrence of p3 and p11 is very dangerous. Hence, p3 and p11 are defined as high risk states and other states are low risk states. Reliability analysis determines whether p3 and p11 will occur at the same time and answers how to avoid the co-occurrence of p3 and p11. The key states graph of the system is shown as Fig. 2.

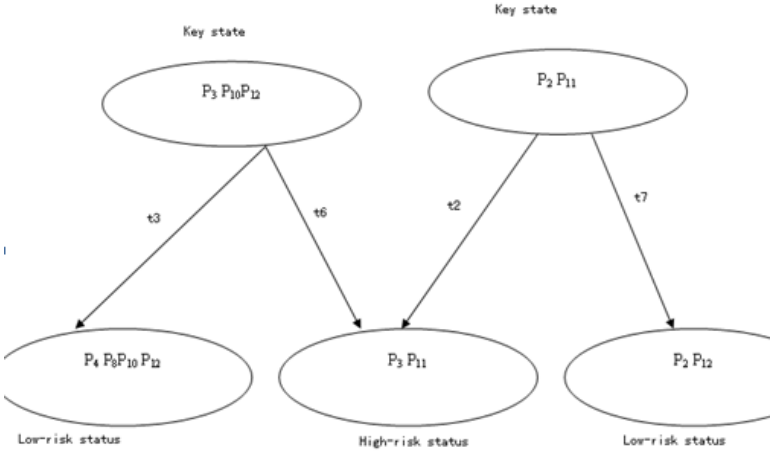


Fig. 2. Critical state figure

From Fig. 2, to avoid occurrence of high risk state, it is only need to be ensured that transition t6 is fired after transition t3 is fired and transition t2 is fired after transition t7 is fired. For this reason, many measures can be adopted to avoid occurrence of high risk states, such as timing restriction method, interlock method and locking position method and so on. Locking position method is based on exclusion of token possess of positions pi and pj i.e. pi and pj can not possess token at the same time. In this way, a security measure is implemented. A security design with locked positions of the system is sketched as Fig. 3.

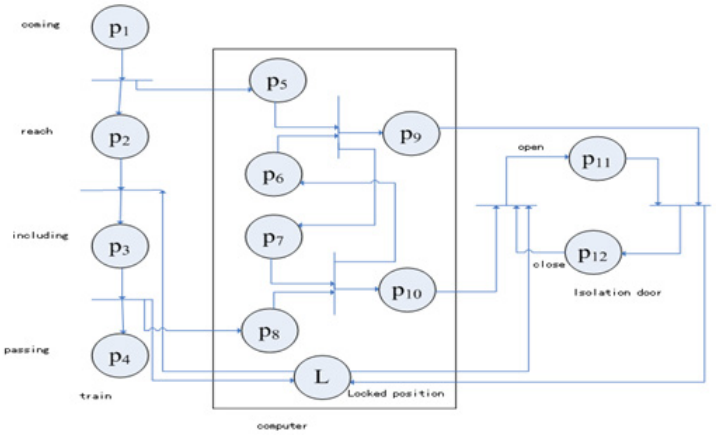


Fig. 3. The system figure with the locked position

5 Conclusion

Petri net is a powerful modeling tool. One of its merits is integrated various elements into a model such that impacts of failure of an element to another element can be found easily. At present, Petri net has become a powerful tool of studying discrete event dynamic system. In this paper, reliability of a control system for railway intersection was analyzed using Petri net. The intention is to deepen our understanding of Petri net application.

References

[1] Fan, L.-B., Zhao, M.: Similarity and distinction between software reliability and hardware reliability. Journal of Guizhou University 24(2), 166 (2007)
[2] Lu, Y., Xu, R.: System reliability. Machinery Industry Press, Beijing (2008)
[3] Wu, Z.: Introduction to Petri nets. Machinery Industry Press (2006)
[4] Li, T., Li, D., Sun, X.: Process scheduling based on Petri net modeling. Computer and Information Technology, 39 (2011)
[5] Liu, C.-Y., Li, Y.-Q., Wen, Y.-F.: A method of power system restoration based on Timed Petri Nets. Journal of System Simulation 19(1) (2007)

Digital Evidence Protection Model with Batch-Verifying and Public Verifiability for Computer Forensics

Zhao-Hui Qi*, Zhan-Feng Gao, Ying-Ming Shen, and Bing-Xin Han

College of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, 050043, China
zhqi_yh2004@yahoo.com.cn

Abstract. The traditional integrity protection of evidence currently available only accounts for the entire time between seizure and presentation. Few people focus their eyes on the time before seizure. Here, through analyzing in detail the relation between digital evidence and intrusion process, a digital evidence protection model with batch-verifying and public verifiability is proposed to protect the generated digital evidence in time. The new protection model can reasonably account for the entire time between generation and presentation. The digital evidence based on the new model is valid and acceptable.

Keywords: cryptography, computer forensics, batch signing, public verifiability, message integrity.

1 Introduction

Computer crime has become a notorious criminal activity. Its frequency and prevalence continue to grow. Traditional criminal investigations need new digital evidence collection tools and technologies in order to punish effectively the computer crimes. The need has led to the development of digital forensic technology, especially computer forensics. Now, computer forensics [1-3] has been around for a while, but it is rapidly becoming a specialized and accepted investigative technique. However, for computer forensics, how to assure the integrity of the digital evidence is an especially difficult problem at present. The gathering of evidence in a computing environment is not only copying files from the suspect's computer and printing them out for presentation in a proceeding. The investigators should establish data integrity assurance.

In cases where seizure is required, most forensic investigation takes place in the forensics laboratory. The investigators are required to present the requisite documentation to prove the integrity of evidence. If the integrity of evidence is broken, the evidence could become compromised and, therefore, unacceptable in a court of law [4,5]. However, at present, the digital evidence is not able to document directly the computer crimes in any court of law though the investigators present the

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 237–242.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

integrity of evidence (i.e., the digital evidence in computer is not direct evidence, but indirect evidence). We think that the traditional integrity of evidence presented by the investigators only accounts for the entire time between seizure and presentation, and it is not possible to explain the time before seizure. If someone vicious had destroyed the digital evidence before seizure, the integrity of evidence from seizure to presentation is unacceptable. In this paper, we firstly analyze in detail the relation between digital evidence and intrusion process. Based on the analysis, then we define a witnessing model on the integrity of evidence that is very important to the capture of digital evidence. The witnessing model allows for the first time the forensic process. Then a proposed batch verifying message integrity scheme with public verifiability is able to provide the integrity of evidence. The scheme is a three-party authenticated protocol, and not only provides efficient integrity assurance for computer forensics, but also has an efficient non-repudiation procedure.

2 The Relation between Digital Evidence and Intrusion Process

There are several different phases in a whole intrusion process. In this paper, we define them as intrusion early phase, intrusion middle phase and intrusion late phase. In intrusion early phase, as for the intruded system, the compromise of the action of an attacker is negligible. In intrusion middle phase, the invader has succeeded in entering the destination, but no one discovers the intrusion behaviors because of his hidden skills. In intrusion late phase, intrusion behaviors are discovered.

During the course of whole intrusion, every phase would generate some important evidences. It is important how to protect these evidences against destruction. We use the notions of security domains (SD) proposed in [6] to account for the relation between digital evidence and intrusion process. SD may be composed of a single host, a single router or a local network. The attacker does not have direct access to the SD group because of some security mechanisms such as firewalls restrict him to specific services.

In order to attack successfully the SD group, the attacker chooses his first target within easy reach. We assume that SD1 becomes the attacker's first aim. At first, what he can only do is to find weaknesses that can be explored remotely, as shown in Fig.1 (a). The attack to SD1 is in the intrusion early phase. After having gained some

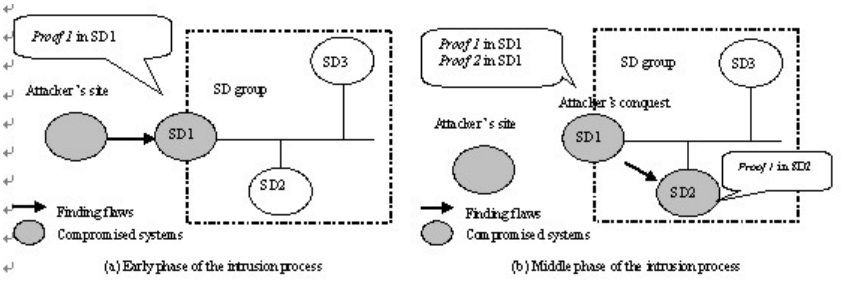


Fig. 1. The relation of digital evidence and intrusion process

level of privilege in the SD1 through exploring remotely but not access privileges, the attacker is going to build up rootkits or backdoors in order to proceed with obtaining access privileges.

Obviously, the attacker's behaviors have generated some important digital evidences though he cannot gain access privileges on the SD1. Here, we call the digital evidences generated in the intrusion early phase as *Proof 1*. At this stage, the attack is not able to destroy the *Proof 1* in SD1. Having gained some level of access privilege, the attacker moves on to the second phase, in which he seeks to obtain the privileges needed to finish his attacks such as destroying *Proof 1* in SD1 and attacking neighbor SDs of SD1. The attack to SD1 is in the intrusion middle phase as shown in Fig.1 (b).

The attacker's some behaviors must have generated some important digital evidences. Here, we call the digital evidences generated in the intrusion middle phase as *Proof 2*. At this phase, there are two types of digital evidences: *Proof 1* and *Proof 2* in SD1. The attacker can modify, delete or forge these digital evidences because of his access privilege. In order to damage the entire SD group, the attacker turns SD1 into a base for the next attack to SD group. SD2 becomes the attacker's next aim. The attack to SD2 is in the intrusion early phase. At this stage, there is digital evidence *Proof 1* in SD2 (Fig.1 (b)).

3 The Proposed Digital Evidence Protection Model

Based on the above analysis of the relation between digital evidence and intrusion process, we think that the messages possibly related to intrusion behaviors should be transferred to a safer place in real time when they are generated. Especially, the data integrity assurance must be established synchronously.

The messages probably related to digital evidence are transferred in time from the Monitored Computer (MC) to a Safe and Trusted Location (STL) when they are generated. Especially, a proposed batch-verifying scheme with public verifiability is used to synchronously establish the data integrity assurance. When the attacker's intrusion is discovered, the legal investigators will seize the intruded system and the STL. The most of digital evidences have been kept in the STL by the timely transfer. The data integrity assurance for digital evidence can be strictly proved by recognized cryptology theory. The digital evidence is valid and acceptable in a court of law.

At present, there are several popular digital signature schemes, which can provide data integrity assurance. They are the RSA signature scheme [7], the ElGamal signature scheme [8] and the digital signature standard (DSS) [9] based on the ElGamal signature scheme, respectively. As for DSS, it requires at least two modular exponentiations to verify each signature. The modular exponentiation over a very large modulus is a very time-consuming computation. The data integrity assurance for computer forensics needs to verify a great number of digital signatures. It is important to speed up the signing and verifying. Here, a batch-verifying scheme with public verifiability is proposed as the integrity protection method of evidence.

4 The Batch-Verifying Scheme with Public Verifiability in the Digital Evidence Protection Model

In the batch-verifying scheme with public verifiability, there are three parties: the Monitored Computer (MC), the Safe and Trusted Location (STL), which is used to store the signed forensics messages, and the Trusted Third Party (TTP). The scheme shares the same parameters as the original DSS [10] and the publicly verifiable authenticated encryption scheme in [10]. Initially, the TTP choose two large primes p and q with $q \mid (p-1)$ and an element g of order q in the finite field $GF(p)$. H is a public one-way hash function. The MC chooses an element x_A in $GF(p)$ as its private key, and computes $y_A = g^{x_A} \bmod p$ as its public key. It publishes y_A which is certified by the TTP and keeps x_A secret. Similarly, the STL generates its private key x_B and public key y_B. The symbol "||" denotes the concatenation of data, such as $m_1 \| m_2$.

4.1 The Signatures of the Monitored Computer (MC) and the Batch Verifying of the Safe and Trusted Location (STL)

The computer perhaps generates a great deal of messages in the process of running. However, the proportion of crime evidences derived from the messages is low. Moreover, they are crucial to computer forensics. It is expensive to search for the crime evidences from a mass of messages in real time. Therefore, the MC transfers in time the messages possibly related with intrusion behaviors to the STL when they are generated. Especially, the MC synchronously signs the messages through interaction with the STL in order to provide data integrity assurance for computer forensics.

For each message m_i to be signed, $i = 1, 2, ..., t$ (to simplify the description, assume that there are t messages in a security time window), the MC does the following.

Step 1: *Chooses a new random integer* k_i, $1 < k_i < q$, *computes* $\lambda_i = ((g.y_B)^{k_i} \bmod p) \bmod q$.

Step 2: *Computes* $r_i = H(\lambda_i, H(m_i))$.

Step 3: *Computes* $s_i = k_i - x_A r_i \bmod q$.

Then the MC sends (r_i, s_i) to STL. After receiving (r_i, s_i), the STL computes $v_i = (g.y_B)^{s_i}.y_A^{r_i(x_B+1)}$. When obtaining all (r_i, s_i), $i = 1, 2, ..., t$, the STL replies with e-bit random message b. Then the MC computes the parameter λ. $\lambda = H((\sum_{i=1}^{t} \lambda_i \bmod p) \bmod q, H(m_1 \| m_2 \| ... \| m_t \| b))$. After computing the parameter λ, MC sends it to STL. After receiving all $(r_i, s_i), i = 1, 2, ..., t$ and λ, the STL dose the following.

Step 1: *Computes* $v = H((\sum_{i=1}^{t} v_i \bmod p) \bmod q, H(m_1 \| m_2 \| ... \| m_t \| b))$.

Step 2: *Checks batch verifying criterion:* $\lambda = v$.

The individual verification criterion is

$((g.y_B)^{k_i} \bmod p) \bmod q = ((g.y_B)^{S_i}.y_A^{r_i(X_B+1)} \bmod p) \bmod q$, i.e., $\lambda_i = (v_i \bmod p) \bmod q$.

And the batch verifying criterion is $\lambda = v$.

4.2 Security and Efficiency Analysis of the Batch-Verifying Scheme with Public Verifiability

A secure batching-verifying message integrity scheme for computer forensics should be able to satisfy the desirable attributions such as unforgeability, original, non-repudiation, and satisfactory efficiency.

The integrity assurance for computer forensics relates to active attack. The proposed scheme can be resistant to this attack by detecting. If a violation of integrity is detected, one can be sure that the received MC's signatures are inauthentic in all probability. Regarding forging MC's signatures, a dishonest STL is in the best position to do so, as it is the only one who knows x_B which is used to verify MC's signatures. Given $((r_1, s_1), (r_2, s_2), ..., (r_t, s_t), \lambda)$ generated by the MC, it is impossible for the dishonest STL to find out other parameters to satisfy both STL's batch verifying criterion $\lambda = v$ and TTP's public batch-verifying criterion $\lambda = \mu$. Therefore, the proposed scheme is unforgeable.

In the scheme, the messages possibly related with intrusion are transferred from the MC to the STL in time after they are generated. The invader has no chance to destroy any evidence. On the other hand, the MC synchronously signs the messages and sends the signatures to both the STL and the TTP. The STL and the TTP can verify the MC's signatures. Hence, the message integrity assurance for computer forensics is synchronously established.

Once the STL computes $\mu_i = y_B^{s_i}.y_A^{r_i x_B}$ and $H(m_1 \| m_2 \| ... \| m_t \| b)$, anyone can verify the batch signature $((r_1, s_1), (r_2, s_2), ..., (r_t, s_t))$ of the messages $m_1, m_2, ..., m_t$. It is computationally feasible for any TTP to settle disputes among the MC, the STL and the invader without divulging STL's private key and the messages $m_1, m_2, ..., m_t$. Especially, the evidences from the messages signed by the MC can legally and truly prove the invader's any lawless behaviors. Thereby the presented scheme has the property of the non-repudiation.

The data integrity assurance for computer forensics needs to verify a great number of digital signatures. The speed of verifying is important. The proposed scheme is a batch verifying message integrity scheme. The batch verifying method greatly improves the efficiency of verifying mass signatures.

5 Conclusion

The traditional integrity protection of evidence currently available only accounts for the entire time between seizure and presentation. Few people focus their eyes on the time before seizure. In this paper, we propose a digital evidence protection model with batch-verifying and public verifiability. The model can begin with the instant in which digital evidence comes into being. Moreover, in order to speed up the signing and verifying, a batch-verifying scheme with public verifiability is proposed as the integrity protection method of evidence. The given proofs show that the scheme has properties of unforgeability, original and non-repudiation.

Acknowledgements. We thank the anonymous reviewers for their valuable comments to improve this paper. This work is supported by Hebei Province science and technology research and development projects (Project No. 09213565D).

References

[1] Marcus, K.R., Kate, S.: The future of computer forensics: a needs analysis survey. Computer and Security 23, 12–16 (2004)
[2] Haggerty, J., Taylor, M.: Managing corporate computer forensics. Computer Fraud & Security, 14–16 (2006)
[3] Ayers, D.: A second generation computer forensic next term analysis system. Digital Investigation, S34–S42 (2009)
[4] Wolfe, H.B.: Computer forensics. Computers and Security 22, 26–28 (2003)
[5] Wang, Y., Cannady, J., Rosenbluth, J.: Foundations of computer forensics: A technology for the fight against computer crime. Computer Law & Security Report 21, 119–127 (2005)
[6] Pal, P., Webber, F., Schantz, R.E., et al.: Survival by defense-enabling. In: Proceedings of the New Security Paradigms Workshop (NSPW 2001), pp. 71–78. ACM Press, New York (2001)
[7] Rivest, R.L., Shamir, A., Adelman, L.: A method for obtaining digital signatures and public-key cryptosystem. Communications of the ACM 21, 120–126 (1978)
[8] El Gamal, T.: A public-key cryptosystem and a signature scheme based on discrete logarithms. IEEE Transactions on Information Theory IT-31, 469–472 (1985)
[9] Federal Information Processing Standards Publication 186, Digital Signature Standard (1994), http://www.itl.nist.gov/fipspubs/fip186.html
[10] Changshe, M., Kefei, C.: Publicly verifiable authenticated encryption. Electronics Letters 39, 281–282 (2003)

Finite Element Analysis of a Bus Skeleton and Structure Improvement

Fachao Jiang[1,*], Yuezhen Fan[2], Entuo Liu[1], and Molin Wang[1]

[1] College of Engineering, China Agricultural University, Beijing 100083, China
[2] School of Technology, Beijing Forestry University, Beijing 100083, China
jiangfachao@163.com

Abstract. Using Finite Element Analysis (FEA) to analyze structure of a bus is a main step of the body design. It can optimize the mechanical structure. A bus sample was researched. Its three-dimensional geometric model was built with CATIA. And a finite element model is setup with basic element of CQUAD4 for future strength analysis in Nastran. When building FEA model, it is essential to consider some factors such as simplification of geometric model, boundary conditions, weight and load. And five working conditions were analyzed. Finally, regional structure with high strength was improved and some improving methods were given based on the result of simulation.

Keywords: bus structure, FEA, Optimizing, static strength, structure improvement.

1 Introduction

Finite element analysis (FEA) is beneficial to improving bus' mechanical performances, structural improvement and design, shortening development period of a new model and reducing cost. Using FEA to analyze structure of bus and present better opinions to meet the requirements of stress and stiffness is a main step of the body design. A bus sample was researched. Its three-dimensional geometric model and finite element model was built. According to analyzing of the five work conditions, some methods of improving the stress distribution were given.

2 Finite Element Model

2.1 Parameters

Finite Element model is the foundation of solution and analysis. The result of analyzing is influenced by quality of the model. Parameters of the bus are shown as Table 1.

Table 1. Parameters

Length(m)	Width(m)	Complete kerb mass(kg)	Maximum pay mass(kg)
11.94	2.53	13250	18000

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 243–247.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Three-dimensional geometric of the bus model (Fig.1) is built with Catia. The mesh generation involves unit type, unit size, the choice of unit amount. Finite Element Model (Fig.2) was generated with Hypermesh. Then it was introduced into Patran to pre-processing. After having setup density, modulus of elasticity, Poisson ratio, load and constraint, Nastran was used to solve to get a result.

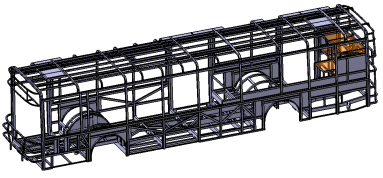


Fig. 1. 3D Model of BUS Skeleton

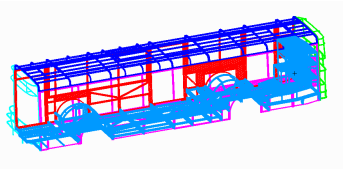


Fig. 2. Finite Element Model

2.2 Different Operating Condition Analyzing

Five operating condition such as bend, bend and twist, brake, turn and climbing was selected to analyze the stress distribution of the bus.

1. Bending condition.When worked on bending conditions, Road surface reaction force makes the body under vertical symmetrical load. So it has a bending deformation. Its deflection depends on the static load and vertical acceleration of the body [1]. The structure must have enough intensity and fatigue strength. The skeleton works at a higher speed and larger dynamic load. So the frame has a basic vertical deformation [2]. In this condition, bending strength of body structure with four-wheel landing under full load was studied. Its stress distribution is shown in Fig.3.

2. Bending-torsion condition.Bending-torsion condition is the most heavy duty condition. It happens when the bus run though a bumpy road. And one wheel maybe lose contact with the ground. If it happened to right front wheel, stress distribution of the bus skeleton is shown in Fig.4.

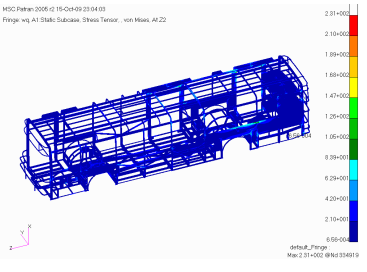


Fig. 3. Stress distribution under bending condition

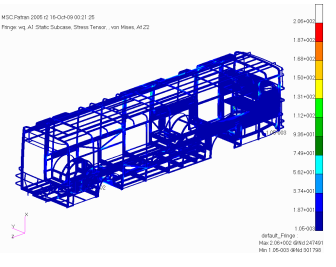


Fig. 4. Stress distribution under Bending-torsion condition

3. Cornering condition. When the bus turns, inertia force will be produced and act on the frame. So vertical acceleration and lateral acceleration should be considered.

After having added the inertia force to the skeleton, stress distribution of bus frame under this condition can be gotten (Fig.5).

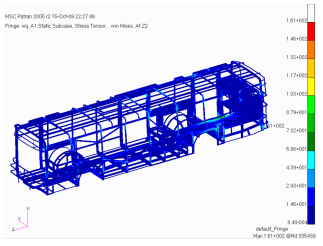


Fig. 5. Stress distribution of bus frame under turning condition

4. Braking condition.In the braking condition, the body suffers the load coming from each unit's static load and inertia force of the vertical [3][4]. In the braking condition, front and rear axle load of bus will change rapidly and internal stress distribution of whole body will have great changes too. So body structure must have sufficient strength to withstand the severe reflection of this condition. When bus turns, seats and passengers will produce inertia force. So the weight of seats and passengers are changed. Under braking condition, stress distribution of whole body frame is shown in Fig.6.

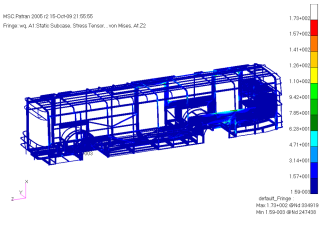


Fig. 6. Stress distribution of frame under braking condition

5. Ramp bending-torsion condition.Ramp bending-torsion condition means a wheel lose contact with the ground when a bus is climbing. Besides the whole body effected by bending and torsion, the force of each device is changed. So the direction of gravity field needs to be changed too. When road grade is 12%, stress distribution of whole body skeleton under the right front wheel hanging condition on ramp is shown in Fig.7.

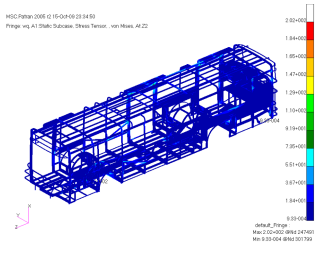


Fig. 7. Stress distribution of hanging right front wheel on ramp

To summarize, maximum stress in these five conditions (table 2) was gotten according to the result of simulation.

Table 2. The maximum stress of the assembly in five condition (MPa)

	Bending condition	Bending-torsion condition	Turning condition	Braking condition	Ramp bending-torsion condition
Front panel	50.9	71	63.2	43	70
Rear panel	231	178	161	173	175
Left panel	130	93.5	125	106	88.9
Right panel	217	173	154	157	172
Top cover	174	161	118	131	160
Chassis	175	206	129	125	202

3 Optimization of Part Structure

According to the analysis of the above different conditions, stress concentrations exist in some parts. So these parts need to be improved. Bending condition was selected to analyze as a simple. After thickening the rear door pillar from 2mm to 4mm, the maximum stress changed from 217MPa to 160MPa. The maps of the stress distribution are shown in Fig.8 and Fig.9.

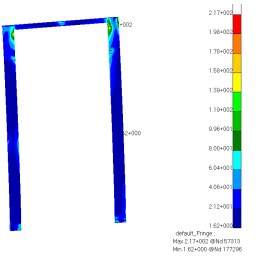


Fig. 8. Stress distribution of rear door pillar before optimized

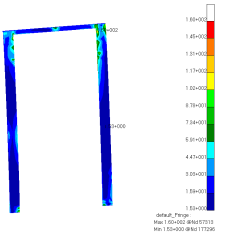


Fig. 9. Stress distribution of rear door pillar after optimized

Furthermore, some methods can be used to improve the structure.

1. Change three doors to two doors.The super capacitor bus is about 12m. So the middle door can be removed and remain the front door and the rear door. Because this scheme not only can meet the needs of passengers but also can increase the intensity of the body.

2. Broaden the door pillar, thicken the rear window pillar or increasing cross section. The simulation results show that larger stress concentration appears in door pillar and window pillar. So the structure in this place needs to be strengthened. So increasing the material thickness or increase the pillar cross section was expected.

3. Reducing the weight of equipment of the rear part of the bus and strengthen the support part. Super capacitor, motors and power control unit are installed in the rear of the bus. The weight of these components needs to be reduced. It can reduce the load received from chassis. Strength chassis's longitudinal beam can solve the problem too.

4 Conclusion

A three-dimensional geometric model of a bus was built with CATIA. And its finite element model was setup with basic element of CQUAD4 for future strength analysis in Nastran. Some methods were given to improve the structure of the bus based on the result of simulation.

Acknowledgements. This paper was supported by Chinese Universities Scientific Fund (BJFU-YX2010-8).

References

[1] Babuskal: Finite e1ement methods for domain with eorners. Computing (1980)
[2] Charalabos, Y.: Doumanidis, Mechanical Analysis of Ultrasonic Bonding for Rapid PrototyPing. Journal of Manufacturing Science and Engineering 124 (May 2002)
[3] Lau, T.S., Lo, S.H.: Finite element mesh generation over analytical curved surfaces. Computer & Structures 59(2), 301–309 (1994)
[4] Tautges, T.J., Blacker, T., Mitchell, S.: The whisker-weaving algorithm: a connectivity based method for con-structing all-hexahedral finite element meshes. International Journal for Numerical Methods in Engineering 39(19), 3327–3349 (1996)

Design on a New Energy Lawn Mower: Fuel Cell Lawn Mower

Fachao Jiang*, Entuo Liu, and Molin Wang

College of Engineering, China Agricultural University, Beijing 100083, China
jiangfachao@163.com

Abstract. The power source of traditional lawn mower is internal combustion engine, it results in noise pollution, leads to an emission problem, and increases the pressure of energy. Fuel cell lawn mower replaces two-stroke or four-stroke gasoline engine of the traditional lawn mower with a fuel cell engine. It has the advantage of energy economy, environmental protection and low noise. Developing the new fuel cell agricultural and forestry machinery have much advantage such as low noise, environmental protection. It is important for the sustainable development. Development of fuel cell lawn mower is introduced. 3D model of lawn mower is set up by using software ProE, it is worth to study and develop Fuel Cell Lawn Mower.

Keywords: fuel cell, lawn mower, design, CAD.

1 Introduction

As lawn industry developing, the demand on the lawn is also growing. Traditional Lawn Mower use two-stroke or four-stroke gasoline engine, which has low fuel efficiency, bring about a serious pollution and cause the problem of energy and emission. Because of size limits, the engine exposed, leading to big noise. Usually lawns situate at the playgrounds or in residential areas, noise and emissions may cause a bad influence on environment and resident. So developing a new fuel cell Lawn Mower, which is low in noise and green, is essential. Energy saving, environmental protection is more important, especially when energy shortage becomes serious.

There are two kinds of electric Lawn Mower in the domestic market: rechargeable and cable. Restricted by the battery performance, the rechargeable electric Lawn Mower can't work continuously, and the cutting torque can't be provided enough. For the cable Lawn Mower, which uses external AC power, its power and speed are low. It's not easy for height adjustment, and has poor security performance. Both types are electric motor-driven, either subjected to limits on providing battery energy, or to the harness. As it is precisely because of these problems, they have not been widely applied.

The fuel cell energy is zero emission. Kept a continuous supply of fuel, the cell will continue to provide electricity. As the product of reaction is water, it won't cause

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 249–253.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

pollution to the environment. If fuel cell technology used in Lawn Mower, it would make it work quietly, green and long, overcoming the shortage of electric Lawn Mower. This paper mainly introduces the fuel cell Lawn Mower, to explore the feasibility of application of fuel cell technology in agricultural and forestry machinery. It is helpful to develop green agricultural and forestry machinery [1] [2].

2 Fuel Cell Technology

The fuel cell transforms the chemical energy stored in the fuel and oxidant direct into electrical energy. When it works, the external supply of fuel and oxidant react. In principle, as long as the reaction continuously, power generation can be kept on. By electrolyte the fuel cell is divided into five types: Alkaline Fuel Cell (AFC), Phosphoric Acid Fuel Cell (PAFC), Solid Oxide Fuel Cell (SOFC), Molten Carbonate Fuel Cell (MCFC), Proton Exchange Membrane Fuel Cell (PEMFC).

The course of the reaction: the wet H_2 and O_2 enter the anode and cathode respectively, by the electrode diffusion layer reach the surface of the catalyst layer and proton exchange membrane. Then under the catalyst redox reactions occur. The anodic product proton (H^+) achieves the cathode through the proton exchange membrane, and the anodic product electron achieves cathode through the external circuit. The water produced is emitted from the cathode in the form of vapor or condensation water [3].

Compared with other fuel cells, Proton Exchange Membrane Fuel Cells has higher thermal efficiency and power density, less noise and vibration, only produces waters, lower reaction temperature. It is easier to start-up and it has been used in automobiles. Canada's Ballard Power Systems Inc, Dalian Institute of Chemical Physics and Shanghai Shenli Company are conducting relevant studies.

3 Fuel Cell Lawn Mower Design

Rotary-type and hob-type are more common among Lawn Mower. The hob-type Lawn Mower is complex, higher in price, suitable for smooth and high quality lawn. Rotary-type is simple, easy to operate, suitable for the common artificial lawn which not need a high quality. It is widely used from large-scale commerce to family daily pruning. The hand-rotary-type scheme is selected. It’s simple and light in structure and saves energy. The fuel cell Lawn Mower makes up of the shell, the power, the cutting, the travel mechanism, the height regulator, the handrail, the grass bag, the control unit [4].

3.1 The Cover

The cover of small and low-power Lawn Mower is normally stamped by using sheet steel or modelled by using high performance organic plastic. The cover of high-power ones is cast by using aluminium alloy.

3.2 The Height Regulator

According to the season and the grasses, the weeding height can be regulated. After been regulated, rotary cutter should be maintained parallel with the ground. So the weeding effect can be improved.

There are two major height regulators. One is to regulate the wheels independently to change the distance between the cutter and the ground. It's a good way to ensure the cutter parallel with the ground plane, but it increases the number of parts, regulated it difficultly and higher costs. The other is a four-wheel interconnected four-bar linkage regulator. It's simple in structure, quick and easy for controls. Most Lawn Mower adopts this structure. So my design adopts the four-bar linkage to regulate the weeding height. Its range is 20-80mm and divided into 11 levels.

3.3 The Power

PEMFC fuel cell has been used in cars, so we can use it in lawn mower. Direct methanol fuel cell can also be considered. The power should match the width of weeding. And it also has something to do with the whole weight and size. According to compare with other Lawn Mowers, a 3kw-power fuel cell was chosen. The safety of hydrogen energy technology is the key to use fuel cell widely. The methods of the hydrogen are as follows: metal hydride hydrogen storage, liquefied hydrogen storage, hydrogen adsorption storage, and compressed hydrogen storage. The way of compressed hydrogen storage is to store the compressed gaseous hydrogen in high-pressure containers. Compared with metal hydride hydrogen storage, liquefied hydrogen storage, hydrogen adsorption storage, it is a widely used method. What's more, it is cheaper. Inflating and deflating is quick, which works at the room temperature. We select the method of compressed hydrogen storage.

Compared with gasoline Lawn Mower, the energy is mainly transferred through flexible wires, but not use mechanical shaft coupling. Therefore it's more flexible to arrange the various units. The motor for the fuel cell Lawn Mower must match the power of the fuel cell. So the design adopts the three-phase asynchronous motor which rated power is 3kW and maximum speed is 3000npm. The motor links the cutter plate directly. The blade was installed in the plate. So replacing the cutter is very convenient.

3.4 The Cutting Device

The cutting device is composed of three parts: the blade, the cutter plate and the linkage. The blade is vertical to the vertical shaft of the engine. The engine rotates the blade. The high-speed rotating blade strikes and cuts the upper leaves of grasses. At the same time the rotating blade produces a certain vacuum absorbing the leaves to form a tower shape, it is benefit for cutting. The cutter plate is the transition between the blade and the linkage, increasing to the binding force of the blade to ensure the stability of the high-speed rotating. The linkage links the blade with the shaft of the engine.

Blade length is the main parameter to decide the efficiency and the size directly and influence safety definitively. The speed of the rated power output of the engine and the length of the blade determine the maximum line speed. Conversely, according

to the maximum line speed determined by the strength of the case, the blade length can be calculated. It is also needed to consider the size. The longer the blade is, then the wider the weeding size. The higher the efficiency and the greater the engine power is. The length of the blade takes for 530mm. A sharp blade will make the stubble cutting smooth, while a blunt blade will make the stubble grow apart and rough. Therefore, to maintain the blade sharp is also important. It is also needed to check the strength of the shaft connecting the engine and the blade.

3.5 The Moving Mechanism

The moving mechanism of self-run Lawn Mower includes the wheel and the power train. The machine adopts the one-way wheel composed of the rubber wheels, the plastic shaft, and double ball bearings. The fuel cell Lawn Mower is pushed by hand, so the running speed is low. As the strength required of the bearing is low, common ones can meet its acquirement.

3.6 The Control Unit

The control system has two major functions: signal monitoring and fault diagnosis. The control unit monitors various operating parameters of the hydrogen fuel cell and the electric motor, such as voltage, current and temperature. The LCD displays the signal, and the controller diagnoses the fault based on the parameters. It will alarm in case of fault to avoid the accident.

3.7 Other Parts

The machine consists of other parts, such as the safety valve, the handling and DC-AC. The 3D explosion of the fuel cell Lawn Mower designed is shown in Figure 1, assembly drawing in Figure 2. We can use Pro/Mechanism to analyze the movement of the four-bar linkage or other mechanisms, observing and recording the analyzing process or some parameters, such as position, velocity, acceleration, force and movement track.

Matching the fuel and the motor is the key to develop the fuel cell Lawn Mower. On the base of meeting the function, the application of new technology, new materials and new technology achieves lightweight design. The lighter the weight is, the smaller force needed and the easier the handling is.

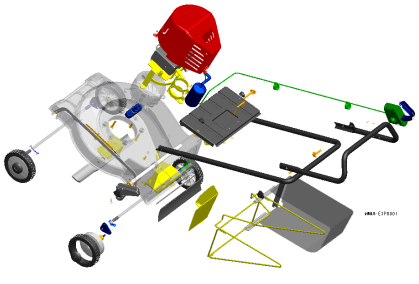


Fig. 1. Exploded view

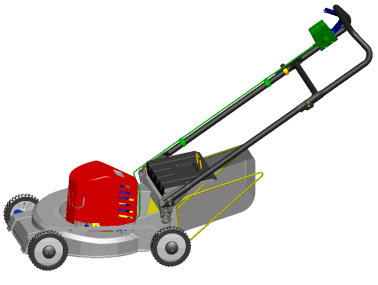


Fig. 2. Assembly drawing

4 Conclusion

In this paper the fuel cell technology was introduced. And a method of fuel cell Lawn Mower was given. The theoretical calculation carried out and a 3D model of fuel cell Lawn Mower with Pro/E was drawn. The fuel cell Lawn Mower is a complex system, which needs improve in further work. The main purpose of the paper is to introduce the feasibility of application of fuel cell technology in agricultural and forestry machinery, and to give an advice for the development of green agricultural and forestry machinery. It is a trend that fuel cell technology is used in agricultural and forestry machinery.

Acknowledgements. This paper was supported by Chinese Universities Scientific Fund (2010JS010).

References

[1] Wang, D.: The Development of Practical Type Multifunction Mower. Forestry Machinery & Woodworking Equipment 31(4) (2003)
[2] Cheng, X.: Garden Machinery. China Forestry Publishing House (December 1995)
[3] Mao, Z.: Full Cell, 1st edn. Chemical Industry Press, Beijing (2004)
[4] Yu, G.: Lawn Maintenance Machinery, 1st edn. China Agriculture Press, Beijing (2004)

Study on the SRI Sensing Characteristics of PFBG and TFBG

Rensheng Shen[1,*], Wen Lin[1], Xiangping Li[1,2], Xiao Zhang[1], and Guotong Du[1,*]

[1] School of Physics and Optoelectronic Technology, Dalian University of Technology, Dalian 116024, China
[2] Department of Physics, Dalian Maritime University, Dalian 116023, China
shjiank@dlut.edu.cn, dugt@dlut.edu.cn

Abstract. The sensing characteristics of surrounding refractive index (SRI) were investigated on the pretreated FBG (PFBG) by hydrofluoric acid (HF) etching treatment and the tilted fiber Bragg grating (TFBG) with an angle of 4° through several experiments, in which the SRI ranges from 1.40 to 1.45. The results demonstrated that the SRI sensing characteristic of TFBG was superior to that of PFBG. Using width changed (ΔL) of high-order modes of TFBG , we can detect SRI whose value ranges from 1.40 to 1.45, between which a good linearity of SRI sensing characteristic can be obtained, and the correlative coefficient can reach 99.543%. The features of TFBG may possess prospects of application in biological and chemical fields.

Keywords: surrounding refractive index (SRI), tilted fiber Bragg grating (TFBG), sensing.

1 Introduction

In the past few years, fiber Bragg grating (FBG) have been utilized as optical sensors to measure a wide range of physical parameters including temperature, pressure, loading, bending, strain, etc[1]. Recently, the surrounding refractive index (SRI) sensing technology based on fiber had received a great deal of attention for their wide range of applications, both in biological and chemical field. To date, a number of SRI fiber sensors have been realized by long period grating sensors (LPGs). LPGs have been utilized for refractive index and concentration sensing because the transmission spectrum that arises from light coupling from the fundamental mode to the forward-propagating cladding modes of the fiber are sensitive to the surrounding refractive index (SRI)[2]-[6]. However, the mechanical strength of LPGs is very poor because grating is long. Besides LPGs, refractive index sensors based on FBGs also have been proposed. Most of them rely on the change of the evanescent field of the guided fiber mode with SRI by etching or side polishing FBGs [7]. This process could significantly weaken the FBGs [8]. So LPGs and the pretreated FBG are not suitable for SRI in practical application.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 255–259.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

In order to solve above problems in the SRI sensing technology based on fiber, tilted fiber Bragg gratings sensor (TFBGs) was proposed to be applied in the SRI sensing technology for its advantages of high-integration, small in size, simplified package and little cross-sensitivity. TFBG is a kind of fiber gratings whose grating pattern has a certain angle perpendicular to the fiber axis. Apart from the characteristics of ordinary FBG that coupling between the forward-propagating fundamental mode and the backward-propagating mode[9]-[11], TFBG further possess properties identical with that of LPG, that is coupling between fiber core mode and fiber cladding mode[2].

In this paper, we manufactured the pretreated FBG (PFBG) which was processed by hydrofluoric acid (HF) etching treatment, and the TFBG with an angle of 4°, respectively. The SRI sensing experiments were implemented on different SRI which ranges from 1.40 to 1.45. The results demonstrated that the SRI sensing characteristic of PFBG was not significant. Compared with PFBG, the SRI sensing characteristic of TFBG has a better linearity, and its correlative coefficient reaches 99.543%.

2 Experiment and Results

Both of the PFBG and TFBG have been manufactured of hydrogen-loaded Corning single-mode fiber (SMF) by means of KrF excimer laser at 248 nm. The required tilted angle of 4° for TFBG was formed by rotating the mask fiber assembly around an axis perpendicular to the fiber axis and to the plane of incidence of the laser light. The PFBG may be obtained by etching normal FBG with hydrofluoric acid having concentration of 40%. Probe light source (Accelink Technologies ASE-3-N-12-FC/APC-B) and OSA (Advantest Q8384) were used to get reflection and transmission spectrum in experiments. The refractive index liquids were provided by Mycro Technologies series AA 18005, in which the value of refractive index ranges from 1.40 to 1.45.

2.1 SRI Sensing Characteristic of PFBG

The PFBG samples obtained by etching normal FBG with hydrofluoric acid having concentration of 40% for 0, 10, 20, 30, 40 and 50 minutes, respectively. The reflection spectra of normal FBG in hydrofluoric acid having concentration of 40% were shown in Fig.1. As shown in Fig.1, we can remarkably observe the red-shift of the reflection spectra of PFBG with time elapsing. The red-shift phenomena were caused by the resonance wavelengths of FBG which depend on the effective refractive indices of core, cladding modes and grating pitch.

We selected a PFBG sample which was treated by HF for 50 minutes. Fig.2 shows the resonance wavelengths of the sample having a diameter of about 60 micrometers which was put into SRI liquids whose SRI ranges from 1.40 to 1.45. It can be seen from figure 2 that the resonance wavelength of PFBG is almost not changed under different SRI. The result demonstrated that the PFBG whose cladding layer about 60 micrometers is not sensitive to SRI. At the same time, the process of HF treatment leads to decreased mechanical strength of PFBG. Thus, PFBG couldn't be practically applied in surrounding refractive index field.

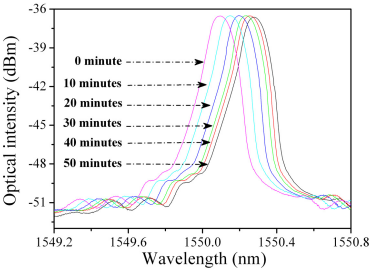


Fig. 1. The reflection spectrum of PFBG in hydrofluoric acid liquid at various times

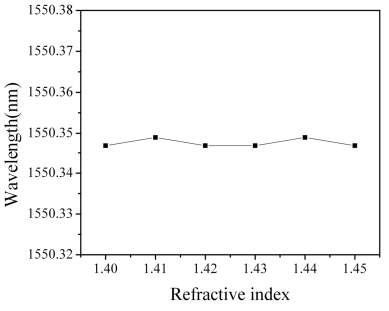


Fig. 2. The resonance wavelength of PFBG with SRI raging from 1.40 to 1.45

2.2 SRI Sensing Characteristic of TFBG

We put the TFBG samples into prepared refractive index liquids with refractive index ranging from 1.40 to 1.45, respectively. The experimental results are shown in Fig.3.

As shown in Fig.3, we found that the resonance wavelength and ghost mode almost are not changed, while the high-order modes changed from short wavelength to long

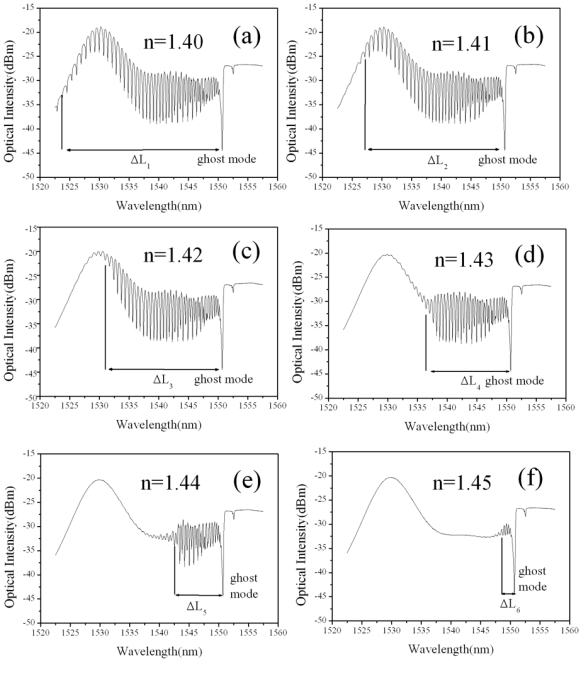


Fig. 3. The transmission spectra of TFBG under different SRI

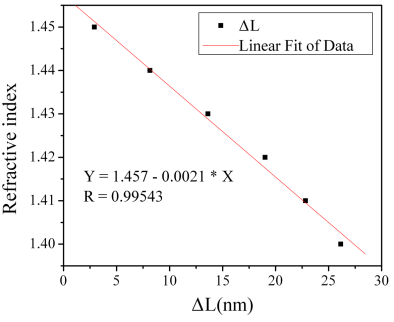


Fig. 4. Relationship between ΔL and SRI

wavelength with the increase of the value of SRI and disappeared gradually. That's because the core-cladding mode is coupled with the resonance of the TFBGs. When the SRI of the liquid is equal to cladding refractive index, there should be no high-order modes resonance [12]. So the curve whose liquid index is bigger than cladding refractive index couldn't change any more. We calculated the distance between ghost mode and higher-order mode, which is indicated as ΔL. Using Origin software, we analyzed the relation between ΔL and SRI, and the results were shown in Fig.4. From Fig.4, it can be seen that the SRI sensitivity of TFBG is 464.8nm/RIU with a good linearity, and its correlative coefficient reaches 99.543%.

3 Conclusion

We have systematically investigated the SRI sensing characteristics of PFBG and TFBG UV-inscribed in SMF. The results of the experiment indicated that the SRI sensing characteristic of PFBG is not significant when its SRI value ranges from 1.40-1.45. However, width change (ΔL) of high-order modes of TFBG can sensitively detect SRI. Compared with PFBG, the SRI sensing characteristic of TFBG has a better linearity. The features of TFBG may prospects of application in biological and chemical fields.

Acknowledgements. This work was supported by National High Technology Research and Development Program (863) of China under Grant No. 2009AA03Z401, the Fundamental Research Funds for the Central Universities under Grant No.2010007 and Scientific Research Foundation for Doctoral program of Liaoning Province of China under Grant No. 20101016.

References

[1] Chen, X., Zhou, K., Zhang, L., Bennion, I.: Optical Chemsensor Based on Etched Tilted Bragg Grating Structures in Multimode Fiber. Photonics Technology Letters 17(4), 864–866 (2005)

[2] Patric, H.J., Kersey, A.D., Bucholtz, F.: Analysis of the response of long period fiber gratings to external index of refraction. Journal of Lightwave Technology 16, 1606–1612 (1998)

[3] Ng, M.N., Chen, Z., Chiang, K.S.: Temperature compensation of long-period fiber grating for refractive-index sensing with bending effect. Photonics Technology Letters 14, 361–362 (2002)
[4] Gwandu, B.A.L., Shu, X., Allsop, T.D.P., Zhang, W., Zhang, L., Bennion, I.: Simultaneous refractive index and temperature measurement using cascaded long-period grating in double-cladding fiber. Electronics Letters 38, 695–696 (2002)
[5] Shu, X., Gwandu, B.A.L., Liu, Y., Zhang, L., Bennion, I.: Sampled fiber Bragg grating for simultaneous refractive index and temperature measurement. Optics Letters 26, 774–776 (2001)
[6] Han, Y.G., Lee, B.H., Han, W.T., Paek, U.C., Chung, Y.: Fibre-optic sensing applications of a pair of long-period fibre gratings. Measurement Science and Technology 12, 778–781 (2001)
[7] Schroeder, K., Ecke, W., Mueller, R., Willsch, R., Andreev, A.: A fiber Bragg grating refractometer, Measurement. Science and Technology 12(7), 757–764 (2001)
[8] Zhao, C.-L., Yang, X., Demokan, M.S., Jin, W.: Simultaneous Temperature and Refractive Index Measurements Using a 3°Slanted Multimode Fiber Bragg Grating. Journal of Lightwave Technology 24(2), 879–883 (2006)
[9] Laffont, G., Ferdinand, P.: Tilted short-period fibre-Bragg-grating-induced coupling to cladding modes for accurate refractometry. Measurement Science and Technology 12(7), 765–770 (2001)
[10] Miao, Y.P., Liu, B., Tian, S., Zhao, Q.D.: Temperature-insensitive refractive index sensor based on tilted fiber Bragg grating. Microwave and Optical Technology Letters 51, 479–483 (2009)
[11] Huy, M.C.P., Laffont, G., Dewynter, V., Ferdinand, P.: Tilted Fiber Bragg Grating photowritten in microstructured optical fiber for improved refractive index measurement. Optics Express 14, 10360–10370 (2006)
[12] Zhang, J.: Research on Blazed Fiber Bragg Grating and the Demodulation Technology of Fiber Bragg Grating, pp. 37–40. Institute of Modern Optics, Nankai University, Tianjin (2007)

E-Commerce Case Study of Fast Fashion Industry

Zhenxiang Wei[1] and Lijie Zhou[2]

[1] Economic and Management Institute, China, University of Petroleum (East China), Qingdao, Shandong, 266580, P.R. China
[2] CNOOC Building, No. 6 Dongzhimen Wai Xiaojie, Dongcheng District, Beijing 100027, P.R. China
weizhenxiang88@163.com, nancy.zhoulijie@gmail.com

Abstract. The article researched into the fast fashion industry worldwide, specifically analyzing the success of Zara, H&M and Gap, followed by an overall analysis of B2C apparel online retailing in China. Then it demonstrates a case of a Chinese fast fashion online retailer, with relevant recommendations for its further prosperity proposed in the end.

Keywords: e-commerce, B2C, online retailing, fast fashion industry, Zara, Vancl.

1 Introduction

The traditional marketing and management of fast fashion industry is experiencing a revolution because of the emerging of e-commerce. Since the birth of e-commerce, businesses have been able to make use of the Internet in reducing costs associated with purchasing, managing supplier relationships, streamlining logistics and inventory, and developing strategic advantage and successful implementation of business re-engineering. E-commerce allows companies to improve communications within the supply chain and enhance service offering, thus providing chances for competitive differentiation. That may explain why so many online retailing companies merged in China recently, such as Joyo Amazon, Dangdang, Taobao, Vancl and 360 Buy. This article takes Zara and Vancl in comparison for case study to analyze the success of the newly merging online retailing company in fast fashion industry and raises suggestions for Chinese fast fashion industry to stay competitive in the future.

2 Brief on the Fast Fashion Industry Worldwide

Consumers are reluctant to spend on discretionary items, and when they do, they expect fashion, quality and low cost items. Fast fashion is a contemporary term used by fashion retailers to acknowledge that designs move from catwalk to store in the fastest time to capture current trends in the market. The apparel products are designed and manufactured quickly and cheaply to allow the mainstream consumer to take advantage of current clothing styles at a lower price.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 261–270.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Since the primary objective of the fast fashion is to quickly produce a product in a cost efficient manner, most companies in fast fashion industry appear as a vertical integration of design, just-in-time production, delivery and sales. Also, it places great emphasis on the efficiency of the supply chain. Because of competitive forces and reducing fashion cycles, retailers have been forced to develop strategies, enabling them to control the supply chain more closely, thus allowing them to source quality products at competitive prices [1].

Table 1. Worldwide Apparel Specialty Stores

Company Name (Flagship Brand)	Country	End of Fiscal Year		Sales (¥ Billions)
Hennes & Mauritz	Sweden	Nov.	2009	1,208.70
INDITEX(Zara)	Spain	Jan.	2010	1,208.40
Gap	USA	Jan.	2010	1,187.00
UNIQLO	Japan	Aug.	2010	814.8

(Source: FAST RETAILING Industry Ranking, Dec. 2010)
*Figures are calculated in yen using the end of November, 2010 foreign exchange rates.

The philosophy of quick manufacturing at an affordable price is used in large global retailers such as Zara, Hennes & Mauritz(H&M), Gap, UNIQLO and Topshop. Zara is a specialist fashion chain and an important example of a fast fashion retailer, with rapid stock turnaround and vertical integration. It belongs to Inditex, one of the largest distribution groups in the world, which operated 1558 stores in 45 countries out of which nearly 550 were Zara stores. 46% of the group's sales were from Spain with France as the largest international market. Zara generated 73.3% of the group's sales and 60% of the Zara sales were from its products for women. Zara gains its success mainly in the following aspects.

2.1 Strong Supply Chain Management

Gap is an America's giant casual-fashion chain. Zara mimics Gap's merchandising strategy of offering differentiated stylish-yet-affordable basic apparel to the masses. Fashion is highly perishable, quickly influenced by the latest thing seen on the catwalk or on the back of a celebrity. Retailers have tried to shorten the ordering cycle and minimize their inventory, so that manufacturers end up carrying more of the risk of managing stock. Many are trying to replace the standard four annual "seasonal" orders with as many as 16 orders a year [2]. Five years ago only Zara followed such a strategy. Zara's designers follow fashion trends closely. In general, a typical clothing company manufacturing in Asia could take six to nine months to get a new design into the shops. With a strong logistics system, an entirely new Zara garment takes about five weeks from design to delivery, while a new version of an existing model

can be in the shops within two weeks. In a typical year, Zara launches some 11,000 new items, compared with the 2,000-4,000 from companies like H&M or Gap.

2.2 Scarcity Value Creation

Apart from the supply chain management, contributing to Zara's success is its focus on a limited range, basic shapes and small sizes, so that it deals with a rather narrow product range. Zara avoids over supply. Although some stock is replenished, its clothing, for both men and women, is deliberately made in small batches. This helps create a scarcity value. It also keeps shops looking fresh and reduces markdowns. At Zara, the number of items that end up in a sale is about half the industry average. The result is that Zara's production cycles are much faster than those of its nearest rival, H&M.

2.3 Low Costs on Promotions

A well-known brand takes a great deal of time to create—partly because, unlike manufacturers, whose products are promoted by shops, retailers must do all the promoting themselves. To compete with Gap, Zara keeps its inventories very lean, meaning it avoids profit-damaging promotions and sales. It avoids advertising in order to cut down on costs. In consideration of promotion, Zara is parsimonious with advertising and discounts. It spends just 0.3% of sales on ads, compared with the 3-4% typically spent by rivals.

Online fashion sales, however, are proving remarkably resilient. Zara and H&M are two more in a long line of fashion stores that have moved into the online retail market. Neither H&M nor Inditex disclose what proportion of sales comes from their online outfits. U.S. rival Gap Inc. generated about 8% of its Gap-branded sales in the U.S. on the Web.

Above all, apparel specialties worldwide tend to gain right balance between fashion, quality and price at the right time.

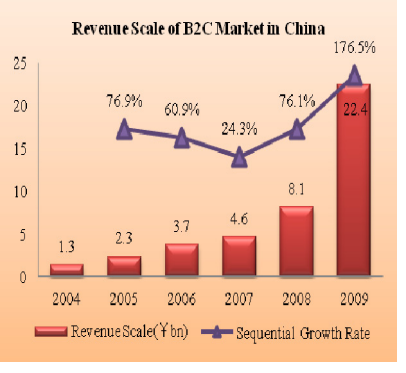
3 An Overall Analysis on Transaction of Chinese Online Fashion

China’s online retailing market value has been roughly doubling in value continuously for several years, and keeps showing a rapid pace of growth. Meanwhile, every six months the population of China’s cyberspace increases by about a third. The annual growth rate of the total online sales market reached 95.9% during the year 2004 to 2010, C2C included, compared to 87.4% of B2C over the same period. However, B2C has accelerated to become segment with the highest growth rate among the total internet transactions forms. Increasingly, confidence in retailers and payment methods, as well as more well-rounded information privacy policies are driving more consumers to shop online, especially towards those identifiable as bone fide companies.

For existing bricks-and-mortar retail companies, online sales are creating new avenues to new consumers in regions of China where they have not yet managed to cover. Without the need to invest in the building physical store chains or management of franchisees, online retailers financially bear less. What does lag, and stalls the more

rapid growth of retailer penetration across more of the country via the Internet, is the lack of product delivery infrastructure and on-going weaknesses in the transaction clearing infrastructure.

For emerging online-only retailers, there is the prospect of creating the kind of strong competition with the traditional retailers that can act as a strong lever for increasing capital investment, or negotiating into partnerships with the traditional retailers, where established online retail expertise and technologies are hugely valuable.



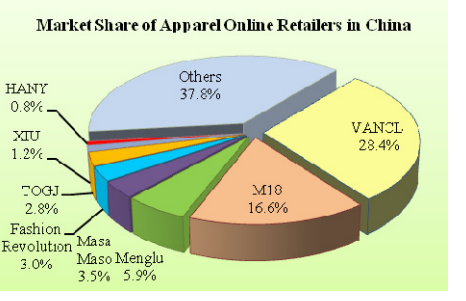


Fig. 1. Revenue Scale of B2C market in China **Fig. 2.** Market Share of Online Retailers in China

(Source: "2010 Chinese B2C Market Investigation", Zero2IPO Research Centre, 2010)

As we can see in Fig.1, the transaction scale of Chinese e-commerce has grown steadily from 2004 to 2009, 35% of which in terms of quantity is mainly engaged by apparel consumption("2009-2010 Chinese E-commerce Industry Report" by IResearch), indicating that the apparel industry in China provides an enormous space for e-commerce. On the basis of the figures gained in the previous years, IResearch estimates the prosperous future trend of Chinese apparel e-commerce transaction scale.

Online fashion has had a 'budget' reputation from the beginning. But there are significant costs for online retailers in logistics and the costs of handling returns. These increased costs will have a high impact on the retailer's low margin, high quantity business model. Also, the average delivery costs, to be paid by the consumer, can easily be as high as, or even higher, as the price of a single piece of budget clothing [3]. Consequently, clothing manufacturers and retailers must establish a good feedback system and improve levels of supply chain management, logistics, product identification, standardization and automation. Different from traditional channels, e-commerce appeals as a significant move to target more customers for Chinese apparel industry. Table 2 shows the revenue of several major apparel online retailers in China from the year 2007 to 2009.

Table 2. Revenue of Transactions of Major Apparel Online Retailers in China

Brand Name	Year Founded	Revenue (￥bn)		
		2007	2008	2009
Vancl	2007	/	0.88	1.22
Masomaso	2008	0.02	0.05	0.08
Moonbasa	2006	0.02	0.06	0.1
M18	1996	0.11	0.1	0.35

(Source: "2010 Chinese B2C Market Investigation", Zero2IPO Research Centre, 2010)

The apparel retailing industry is investing large sums of money in the Internet technology, delivery infrastructure and transaction systems which will give allowance to secure a significant slice of this market into the future. According to a survey conducted by the China Internet Network Information Center (CINIC), major factors affecting online purchasing decisions involve price and quality of goods, brand recognition, after-sale services, customers' feedbacks, speed and quality of delivery, promotions, advertisements and user interface of websites.

4 Case Study of Chinese Fast Fashion Online Retailing—Vancl

Vancl, rated as the largest online and phone-based own-brand apparel retailer in China, was formally established in 2007, led the Deloitte Technology Fast 500 Asia Pacific 2009 ranking with a three-year revenue growth rate of 29,577%. It took over its rival---PPG in terms of sales only 5 months after its establishment, creating a commercial marketing myth in the Chinese garment industry [4]. Vancl has fast become China's first (online) business-to-consumer brand for men's clothing. Through a series of brand and product positioning, Vancl successfully extended the range of products from men's shirts to POLO shirts, pants, shoes and other clothing and household goods which mainly target at the white collar workers and students. In 2009, the company continued to expand its product lines by entering the women's clothing market and by marketing footwear. Vancl has quickly made itself a familiar name to netizens in China under a unique marketing model using online advertisements. The company sells more than 30,000 items ranging from shirts and shoes to home textiles. The company currently has more than 800 employees.

According to "2010 China's Online Shopping Market Monitoring Report" raised by IResearch Consulting Group, Vancl took a share of 5.3% of independent sales turnover of Chinese B2C online shopping market in the 3rd quarter of 2008 and enjoyed 28.4% of the domestic apparel e-commerce share (as is shown in Fig.2) in 2009, thus ranking the top in the fast fashion online retailing in China. Sales of Vancl was recorded worth RMB 2 billion in 2010 and expects the figure to grow to RMB 6 billion in 2011, according to Sohu report in Jan. 2011.

(Source: "2009-2010 Chinese B2C Apparel Transaction Report", IResearch, 2010)

Yet Vancl's success is no accident. It gained success primarily based on the proper product and brand positioning, B2C online direct-selling model, multiple channels

promotion, high-quality products, fair price, efficient logistics, favorable user experience and good team management.

4.1 Product Positioning-Men to Women

What leads to success in the first place is Vancl's right choice of the first entry point into the market--- Men's shirts.

Since users have indirect contact with the online products, Vancl must provide standardized products to improve customer satisfaction and try its best to meet needs of user experience. As criteria for men's shirts are relatively simple, shirts are easier to reach the requirement.

Most of the time, shirts appear as a kind of essential clothing for men, resulting in a high demand in men's shirts.

Trends of urban white-collar clients are subject to foreign influence. Men get used to wear a different shirt every day, thus increasing consumer demand in shirts, which provides Vancl a broad market.

Most of men would haste in shopping, making the demands for shirts relatively simple. Thus, shirts may have similar designs but distinguish by patterns and colors.

In view of the above characteristics of men's shirts consumer market, Vancl creates a model in full compliance with the consuming habits of men on shirts, while providing a more abundant and convenient option. According to the feedbacks of its customers, Vancl began to target at the women clothing later in 2009. Survey showed that over 70% of the visitors on Vancl website were female and apparently there exists a great need to promote women clothing. Currently, Vancl could make total revenue of around 10 million Yuan every day.

4.2 Brand Positioning-Simple, Easy and Natural

The brand name "Vancl" means all customers are ordinary in Chinese. Vancl is always promoting a lifestyle: getting rid of a vanity of things and have a most decent and comfortable way of life. Vancl's slogan goes like "Cotton style is the new formula of life". Clothes of it have common grounds namely simple, classic style, natural and environmental-friendly fabric and easy ways of cutting. And the company is trying to convey a value system to consumers, which is pursuing the brand in a liberate way, being in a style of low profile and simple, seeking calm and elegance and focusing on the details. With a good brand positioning, Vancl gradually establishes a favorable corporate image and forms stable intangible assets.

4.3 Business Model-B2C Direct Selling

Vancl applies the B2C (Business to Customers) model for online retailing. It doesn't possess any fabrication plants or any physical sales channels. Vancl's direct selling model mainly consists of raw material suppliers, garment processors, shopping sites and logistics. In this model, Apart from its brand's design and part of the logistic delivery, Vancl outsources the production process to OEM or other plants. Vancl sells goods directly through the website, phone orders and catalog mail, hence eliminating the need for the traditional maternity shops. Hence, consumers could enjoy a much lower price of commodity at Vancl than other conventional purchasing channels.

That's how Vancl shortens the production cycle and save amounts of money on stock, leading to reduction in cost of products. Therefore, B2C online direct selling model could be regarded as one key ingredient of Vancl's success.

4.4 Marketing Strategy-Ads in Multiple Channels

Vancl attempted to deliver ads through various channels. Take network advertising for instance, Vancl's network media advertising is believed to be successful considering its depth, breadth and accuracy. Here are some tips in its marketing strategies:

- Vancl makes use of large banners to cover the largest area on the page to attract attention when making advertisements on Sina, Tencent, Netease, Sohu and other portal websites in mainland China,.
- Ads are posted on the most popular and frequently visited mainstream communities, forums and some instant communication tools to achieve deep penetration of key target groups.
- Vancl's pay per click advertisement operates just as Google AdWords and gives permission to sellers to bid on relevant keywords used by target audiences to search for desired products. Pay per click advertisements will mostly be shown on the right up side of the front-page.
- The search engines keyword ads and content ads are utilized to seize the potential customers on the "long tail"[5].
- Popular celebrities were once invited to be the image prolocutors of Vancl's products, arousing a "Vancl introduction style" in the summer of 2010.

Despite the amount of the ads, clear selling points, nice arrangement and high qualified pictures of ads play important role in catching the attention of consumers to some extent, thus enhancing sales simultaneously.

There are many types of physical ads as well. Billboards of Vancl are placed aside the passengers of subways and even on the buses in Beijing. Notably, all the products (accessories and packages included) would have an obvious print of "Vancl" logo, making it easy for buyers' acquaintance to remember the brand name.

With a wide range of strategies on publicity and promotion, Vancl is bound to win its market.

4.5 High-Quality Products

Vancl have greatly invested in the selection of suppliers in the upstream, the establishment of the tracking system and the introduction of the design team. Highly trained quality checking staff, usually stationed in the plant of OEM to monitor the entire production process and make product sampling tests occasionally. Then dedicated quality control personnel would conduct a comprehensive examination towards the finished products stocked in the warehouse. Only the products that got passed the check of the appearance, cuffs, buttons, threads and other aspects could be sent to the warehouses.

4.6 Fair Price

Vancl deems to provide good quality products with fair price. The most typical example is its famous "VT" requesting just29RMB. The T-shirts it offers are all of high-quality and dedicated designed by professional designing team abroad. The same applies to its canvas shoes and shirts for men and women. For top consumers, Vancl provides Vjia (V+) products to satisfy their consuming needs.

4.7 Efficient Logistics

Most online retailing companies would naturally outsource logistic to TLC so as to lower down the total operation cost [6]. But who should be blamed for and pay for the loss of customers in case that delivery was delayed? There suppose to be no definite answer in most cases. To solve the problem, Vancl combined the wholly-owned logistic department and third party partners, with the order tracking system ensuring the company receives fewer complaints for delivery.

4.8 Favorable User Experience

Vancl has improved a lot in its user experience services. Vancl firstly activated a service of trial before payment and cash on delivery (C.O.D). And customers would have the right to reject purchase without any reason within 30 days in the validation of orders [7]. Additionally, the site offers enhanced services such as tracking of goods and management of inventory. Apart from that, Vancl provides spaces for customers to comment on the goods they purchased as feedbacks on the website. In order to better fit the needs of customers, the company has bought a 3D virtual model system and is planning to develop its own trial-online animation system based on that.

4.9 Good Team Management

It is the team management, which forms the fundamental difference among other direct marketing online projects, plays a key role in the whole project in Vancl. CEO-Chen Nian, who was also the founder and former executive vice president of Joyo Amazon, led the following job-hopping members form the core team that later became backbones of Vancl. Their experience of network marketing, an accurate grasp of China's Internet environment and the actual operation experience of B2C e-commerce are the best explanations for the Vancl miracle [8].

5 Suggestions for Vancl's Future Prosper

5.1 Strengthen Supply Chain Integration and Supervision

It has always been a hot potato as how to ensure the quality of clothing from upstream suppliers as well as the quality of delivery service provided by the TPL companies. The most effective way to solve it is to strengthen the supervision and control of all aspects during the process. There should be appropriate supervision in the process from the fabric transported into the plant, onto various assembly lines and to final

order fulfillment. Whether outsourcing quality supervision to a third party or establishing its own monitoring team, Vancl must strengthen the integration of the supply chain and all aspects of supervision during the process from production to the final delivery. It could establish relevant integrity systems and service quality system for the cooperative enterprise, in which it judges the integrity and quality of service companies according to their business performance, and selects the more suitable enterprises to cooperate for further development.

5.2 Enhance Quality and Service of Products

Quality is quite crucial towards the enterprise itself. To guarantee the quality of goods sold online, Vancl has to seek creditworthy channels for manufacturing. Feedbacks of consumers could also reflect the quality of goods. Consequently, greater emphasis should be placed on user experience services. Vancl could devote to designing user-friendly page and interactive forum so as to take advantage of the feedbacks to ensure that consumers could acquire a variety of information about products either displayed on the website or the catalogue [9]. Meanwhile, the website must refresh its information regularly and in time to inform consumers about its new fashion and stock.

5.3 Innovate and Enhance Core Competitiveness Constantly

Large numbers of domestic clothing market or network marketing companies are founded to follow the steps of Vancl recently. And the traditional clothing manufacturers rapidly transformed to e-commerce platform, such as the Saint Angelo, Shanshan, Youngor and so on. Therefore, Vancl should continue to innovate to meet consumer demands to lead the industry [10]. On the one hand, innovation helps to build its powerful competitive advantages, making it difficult for competitors to surpass. On the other hand, Vancl's endeavor for innovation would attract more people's attention in the entire industry, which helps to enhance its fame in the future.

Products and direct sales model could be easily imitated by others. Yet, execution of team is exception. Fast fashion B2C industry changes fast so that the enterprise couldn't just rely on the previous experience of the team. Thus, Vancl has to build a learning team to enhance its core competitiveness and provide the learning team a chance to manage and operate the company, hence allowing the enterprise to turn into a learning-oriented enterprise, so as to ensure that businesses compete in an invincible position in the future [11].

5.4 Build Strong Brand Credibility

Fast fashion consumption can be also reckoned as spiritual consumption, core spirit of which is embodied in the brand. Customer loyalty will be established afterwards once the brand is established [12]. Otherwise, it is impossible for consumers to trust the brand. PPG's failure lies in the facts that it paid attention to innovative business models, while ignoring to form strong barriers for its rivals using brand and the real core competitiveness relying on brands. Vancl definitely should learn the lesson and enhance its brand management in the future, paying every effort and operate carefully

in the market research, product positioning and design, advertising and so on. Only through this way could Vancl build strong brand credibility.

Acknowledgements. Special thanks to the support from IResearch institute and Vancl Corp..

References

[1] Bruce, M., Daly, L.: Innovative process in E-commerce fashion supply chains. In: Innovative Quick Response Programs in Logistics and Supply Chain Management, 227–241 (2010)
[2] MacCarthy, B.L., Jayarathne, P.G.S.A.: Fast Fashion: Achieving Global Quick Response (GQR) in the Internationally Dispersed Clothing Industry. Innovative Quick Response Programs in Logistics and Supply Chain Management (2010)
[3] Newbery, M., ter Meulen, K.: Tomorrow's clothing retail: sectors, markets and routes – forecasts to 2016. Just-style. In: International Handbooks on Information Systems (June 2010)
[4] Wu, X., Bai, X.: From PPG to Vancl: Business Model Evolution of Online Clothing Retailing. Journal of Management Case Studies 3, 403–414 (2010)
[5] Liu, Z.: Research on the Cyber Marketing Strategies on basis of 'the Long Tail Theory. Market Modernization 567, 108 (2009)
[6] Heinemmann, G., Schwarzl, C.: New Online Retailing, pp. 187–209
[7] Ran, Z.: Why could Vancl develop so fast——research on the hints Vancl brought for online retailing apparel industry. Business Strategy Case Study, 64–66 (2009)
[8] Ran, H.: Vancl: Say goodbye to PPG. New Economy Guide 11, 56–57 (2009)
[9] Nielsen, J., Molich, R., Snyder, C., Farrell, S.: E-commerce User Experience Search. Nielsen Norman Group (2000)
[10] Chellappa, R.K., Sin, R.G.: Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. Information Technology and Management 6, 181–202 (2005)
[11] Peili, L., Bo, W., Jian, W.: Exploration towards Management Patterns of Chinese Garment Industry—Basis on Vancl Marketing Case. Economic Research Guide 83, 166–169 (2010)
[12] Otim, S., Grover, V.: E-commerce: a brand names curse. Electron Markets 20, 147–160 (2010)

Rapid Impulsive Noise Denoising in Range Images

Jian Wang[1,2,*], Zhen-qiang Yao[2], Lin Mei[1], Yao-jie Zhu[3],
Yuan Yao[4], and Yang Zhang[5]

[1] Cyber Physical System R&D Center,
The Third Research Institute of Ministry of Public Security, Shanghai, 201204, P.R. China
wjconan@ieee.org
[2] State Key Laboratory of Mechanical System and Vibration,
Shanghai Jiao Tong University, Shanghai, 200240, P.R. China
[3] Shanghai Yanfeng Visteon Automotive Trim Systems Co., Ltd. Shanghai 200233,
P.R. China
[4] Information Center, Nanjing University of Aeronautics and Astronautics, Nanjing,
Jiangsu, 210016, P.R. China
[5] Clolrectal Department, Nanjing Hospital of T.C.M, Nanjing, Jiangsu, 210001, P.R. China

Abstract. A novel rapid impulsive noise denoising method for range image is presented. In a 2D scan line acquired by a laser range finder, distribution features of impulsive noises (INs) are analyzed, and then a mathematical representation of the features is provided by defining a few new coefficients. Subsequently, a rule-based distinguishing criterion is formulated to detect two types of INs: dropouts and invaders. The traditional mean filter is improved by an automotive non-IN neighbor searching procedure. A compositive algorithm with a very low computational complexity has been implemented as an embedded module in our self-developed software with copyright. An experiment on real range image is performed, and the results indicate that the proposed method can detect all the impulsive noises accurately and denoise them with a significant efficiency. It is proven that the method is suitable for practical applications on industrial or other fields.

Keywords: Range image denoising, Impulsive noises detestation and removal, Computer vision, Pattern recognition, Laser radar scanning.

1 Introduction

In a digital image, noise is caused by a variety of factors. No matter how good the sensor is, noise is unavoidably introduced into an image within acquisition, compression or extraction. Therefore, image denoising and improvement is always treated as an indispensable preprocessing step prior to image understanding and other actions.

By extensive investigation on popular image denoising methods, the researchers classified them into three categories: local smoothing filters, frequency domain filters and statistical neighborhood approaches[1]. Due to poor performance on preserving

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 271–279.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

edges and textures, the classical Gaussian smoothing approaches have been improved by many local smoothing filters, such as the anisotropic filter (AF)[2, 3], total variation minimization[4, 5], neighborhood filters[6], etc. Providing a superior performance due to properties such as sparsity and multiresolution structure, wavelet transform has become the most popular method for image processing in past two decades. Resultingly, a variety of algorithms for denoising in wavelet domain has been developed. Wiener filter[7] yielded optimal results when the signal corruption can be modeled as a Gaussian process, considering the mean square error (MSE) as the accuracy criterion. As for the nonlinear filter, most of the wavelet shrinkage methods focused on the problem of choosing the optimal threshold. Depending only on number of data points, VISUShrink[8] has asymptotic equivalence suggesting best performance in terms of MSE when the number of pixels reaches infinity. Combining the universal threshold and the Stein's Unbiased Risk Estimator (SURE) threshold, SUREShrink[8] performs better than VISUShrink. Furthermore, BayesShrink[9] outperforms SUREShrink, since it minimizes the Bayes' Risk Estimator function assuming Generalized Gaussian prior. Due to its assumption of signal to be Non-Gaussian, Independent Component Analysis (ICA)[10] has gained wide spread attention recently in denoising images with Non-Gaussian as well as Gaussian distribution. However, the extra computational cost of using a sliding window and difficulty of obtaining the noise free training data become its main drawbacks.

Compared to the large number of literature on intensity image denoising, rather few studies were addressed to range image denoising methods. Limited research focused on analysis of the mathematical model, characteristics and effects of dropouts and range anomalies in range images[11]. With prior knowledge of spatial noise distribution[12], range image denoising is often realized using the combination of a mathematical morphology filter and a mean filter[11].

To the pulsed time-of-flight (TOF) laser radar, the most significant non-system noises are impulsive noises (INs), which look like unexpected pulses in the scan line. These INs destroy the continuity of range images and decrease image quality obviously. Unlike the system noise, INs are difficult to be removed through traditional filters and easy to be confused with normal structures on the surface, e.g. sharp edges, by wavelet transform methods. Some improved method can recognize INs, but with a huge computational cost due to the complexity of the method.

For high efficiency, we proposed a rapid large noise detection method, in which a few easy criteria are established to distinguish the INs and then remove them. Experiments results on real data indicate the effectiveness and efficiency of the proposed method.

2 Distribution Analysis of Large Noise

Basic mind of our approach is to discriminate IN points according to their discontinuities in a single scan line, row line or column line. A large noise appears an unexpected pulse in the scan line. As Fig. 1 shown, a IN point marked with a circle is obviously far away from neighboring points.

There are two main reasons to cause a IN. The first one is dropout phenomenon that occurs when transmitted laser pulse never comes back to receiver. For instance,

dropouts could happen, when there is no reflecting surface in the pulse's path, or reflective laser energy is so low to cannot be detected. Another cause is some unexpected objects invade into scanning region. Inevitably, range finders in outdoor environment occasionally encounter these invaders, e.g. snowflakes, raindrops or big dust particles, which reflect the laser beam at a random position irrespective of the profile of scanned targets. For indoor applications, INs are often caused by dark (or specular) points on the surface or other random spots in the air.

However, all the INs have a common distribution feature. That is, they give rise to obvious uncontinuities in the scan line and have large distances between themselves and the neighboring points. Regarding the dropouts, their range values should approach to the maximum range in current configuration of the range finder. These distinct features make it possible to distinguish IN in a simple way.

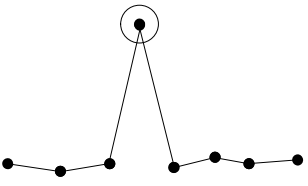


Fig. 1. A large noise point in a scan line

3 Detection and Denoising Algorithm

3.1 Coefficients Definition

Two coefficients are defined to give a simply mathematical representation of IN's features, as Fig. 2 illustrated. Given a point P_i, D_{ai} denotes the distance between P_i and the line connecting its two nearest neighbors, $P_{i\text{-}1}$ and P_{i+1}. Another coefficient, $\Delta\theta_i$, stands for a swerve angle that describes direction variation of the scan line at the given point P_i. Observed from Fig. 1, it is obvious that a IN has a very huge $\Delta\theta_i$, and a rather large D_{ai}, different from other non-noise points.

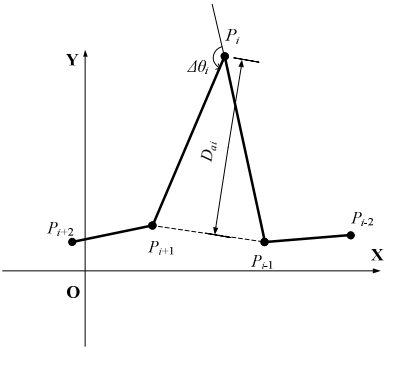


Fig. 2. Coefficients definition on denoising

3.2 Dropout IN Detection

According to the characteristics of the dropouts, range value of a given point P_i can be used to form a distinguishing criterion.

(a) Compute R_{dp};
(b) For $i = 2$ to n-1 do
(c) If $\rho_i >= R_{dp}$, mark P_i as a IN;
(d) Next i

Where,

ρ_i: the measured range value of P_i;
R_{dp}: a dropout threshold which is computed by Eq.(1);
λ_{dp}: an offset coefficient;
R_{max}: range maximum in current configuration.

$$R_{dp} = \lambda_{dp} R_{\max} \tag{1}$$

Note that an empirical range of 0.75~0.95 can be chosen to adjust λ_{dp}, and it is easy to form an adaptive edition.

3.3 Invader IN Detection

Based on preceding qualitative analysis, two thresholds must be provided to obtain a quantitative invader IN detection algorithm in a scan line. Basic principle of the algorithm is given as follows:

(a) For $i = 2$ to n-1 do
(b) Compute $\Delta\theta_i$ of P_i;
(c) If $\Delta\theta_i < \Phi_{in}$, P_i is not an IN, go to (e);
(d) Else Compute D_{ai};
  i. If $D_{ai} < DA_{in}$, P_i is not an IN, go to (e);
  ii. Else, mark P_i as an IN;
(e) Next i

Where, Φ_{in} and DA_{in} are thresholds of swerve angle and point-to-chord distance, respectively.

3.4 IN Denoising

In order to obtain the high accuracy and efficiency, an auto-avoiding mean filter is proposed for IN denoising. Starting from the IN point, the algorithm searches other points at two directions until two nearest non-IN neighbors (NNN) are found. IN mark check is main task in this process. The procedure for searching left NNN of an IN P_i is provided as follows:

(a) For $j = i$-1 to 0 do
(b) Check IN mark of P_j;
  a) If there exist IN mark, P_j is an IN, go to (c);
  b) Else, record point P_j as a nearest non-IN left neighbor;
(c) Next j;

Apparently, the right NNN can be found with a similar approach by convert searching direction. Subsequently, the range value of P_i, ρ_i, is replaced by an average of the range values of left NNN and right NNN, ρ_j and ρ_k, respectively.

$$\rho_i = \frac{\rho_j - \rho_k}{2} \tag{2}$$

It is worthy of noting that all the effective structures of the scan lines, e.g. edges or endpoints, should be prevented from destruction by the denoising method. Practically, a rapid ruled-based edge detection, i.e. RIDED-2D[13], is performed prior to the denoising procedure, and an iterative process is accompanied to refine the scan line.

4 Experiment

In our laboratory, a SICK LMS-291 laser range finder is hanged up to scan the targets on the ground. All the related parameters and thresholds in the experiments are presented as Table 1. Under indoor environment, we utilize many small mirrors on a desk to generate dropout INs and a few floating paper patches to form invader INs.

The proposed algorithms have been implemented as a denoising module in our self-developed software named LMAS[1] (Lidar Measuring and Analyzing System). Fig. 3 plots a panoramic view of a 2D scan line involving 401 scan points. There are 4 dropout INs and 7 invader INs in this scan line.

By observation on Fig. 4 that is an enlarged graph of region A of Fig. 3, we can find that two dropout INs abut each other that will be denoised by a mean filter. Meanwhile, there are two clusters of adjacent invader INs in Fig. 5, an enlarged graph of region B. The amounts of these invader IN clusters are 2 and 3. The scale of the cluster is subject to the size of the invader. Invasion of a big object will lead to laser beam occlusion at more than one azimuth angles for the laser radar. Related to IN denoising algorithm, it is unwise to search non-IN neighbors without radius limitation, just because a large searching range will damage some subtle structures, e.g. a thin erected rod. In our algorithm the searching radius is set to 3 points. Apparently, wider the radius is, more the computational time costs. From Fig. 4 and Fig. 5, it is obvious that all the INs are distinguished accurately. LMAS utilizes dark green boxes to mark dropout INs, and light red boxes for invader INs. Fig. 6 and Fig. 7 illustrate the original region C and a denoised one, respectively. By contrast, it is evident that the scan line becomes considerably smoother after mean filtering. Thereafter, the desk can be recognized by a simple curve fitting or similar approximation approach.

Table 1. Thresholds in experiment

	Parameter	Unit	Value
Configuration	R_{max}	mm	8183
	A_{max}	deg.	100
	A_{res}	deg.	0.25
Thresholds	λ_{dp}		0.90
	Φ_{in}	deg.	120
	DA_{in}	mm	200

[1] LMAS v2.09 is registered software copyright of China with No. 2009SR039116 on September 14, 2009.

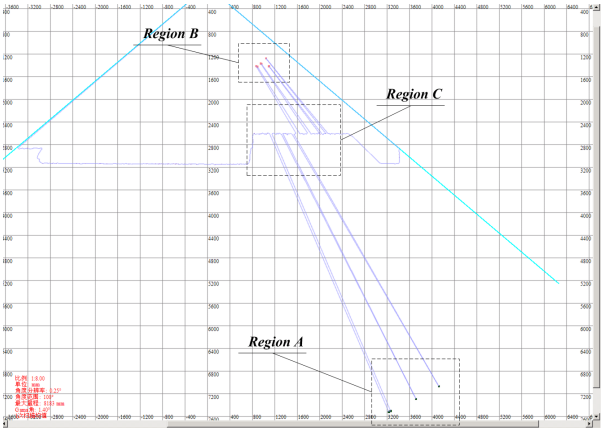


Fig. 3. A scan line with marked INs

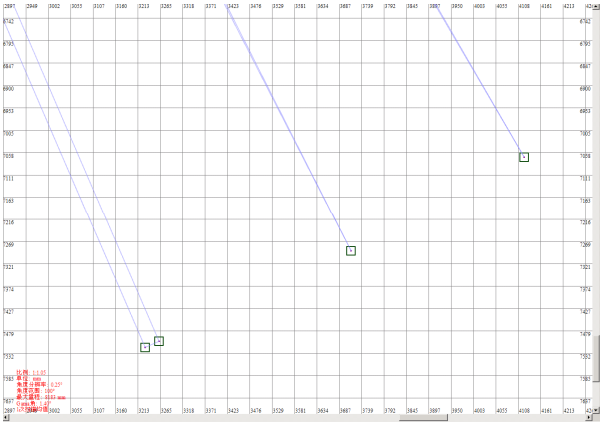


Fig. 4. Enlarged region A with 4 dropout INs

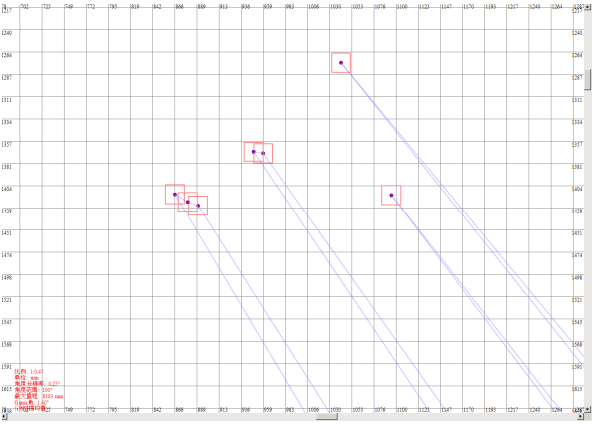


Fig. 5. Enlarged region B with 7 invader INs

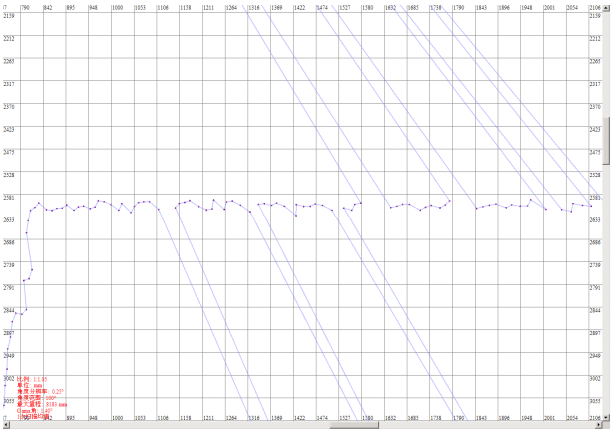


Fig. 6. Enlarged region C (Original)

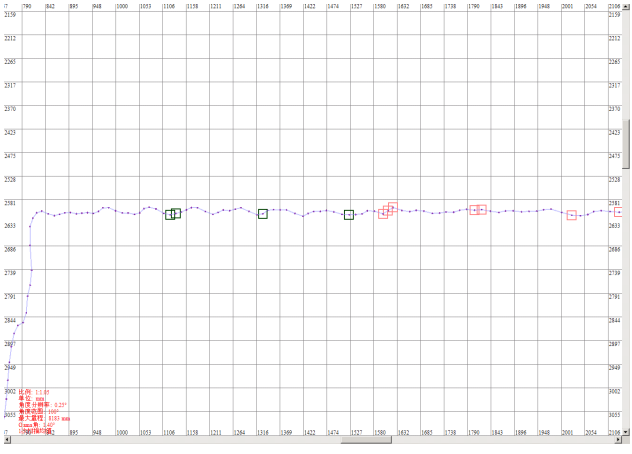


Fig. 7. Enlarged region C (Denoised)

5 Evaluations

As the chapter 3 discussed, there are three main components involved in the proposed method. Since the two detection components can operate the same point synchronously, they may be merged into one loop whose computational complexity (CC) is $O(n)$. The CC of a pure IN denoising step can be regarded as $O(1)$ due to the trivial amount of INs. Therefore, in general, the total CC of the proposed method is $O(n)$. Sometimes, IN denoising is embedded into a mean filtering process on all the points in the scan line; consequently, its CC is $O(n)$.

On the basis of precise time expenditure method by recording CPU clock periods[13], computational cost of the proposed method is evaluated during 20 times independent tests. On the aforementioned sample scan line, the gross average time

expenditure is 23.458098 ms with 64370995 CPU ticks. Concerning the average CPU occupy rate is 25.35%, the net time cost is 5.946628 ms with 16318047 CPU ticks.

6 Conclusions

First of all, we review a lot of popular method for image denoising. Due to high complexity of the algorithm, these methods usually need large computational costs for data processing and are difficult to be put into practice. Especially, denoising impulsive noises that often have a small proportion in an image have not attracted enough attention. For high efficiency, we proposed a novel rule-based method for detecting and denoising impulsive noises rapidly in range images. Our method is established upon a solid foundation including feature analysis on impulsive noises and coefficient definitions. A compositive algorithm has been implemented with a very low computational complexity. Experimental results on real data indicate that the proposed method can detect all the impulsive noises accurately and denoise them with a significant efficiency. It is proven that the method is suitable for practical applications on industrial or other fields.

For the sake of simplicity, this paper only provides the principles how to detect and denoise impulsive noises in 2D range scan lines. The method has been extended to 3D Cartesian ordinate system and spherical ordinate systems, and will be discussed in future papers.

Acknowledgement. Our research was supported by following projects:

- National Science and Technology Support Projects of China (No. 2012BAH07B01);
- Major program of Science and Technology Commission of Shanghai Municipality (No. 09DZ1122202, 10511500700);
- National Natural Science Foundation of China (No. 51075271, No. 51121063, No. 50805094).

Thanks to Mr. Zheng-song XU for helping us to design and install the range finder platform, and are grateful to some classmates at our laboratory, Dr. Xing-guang LIANG, Dr. Wei-bin GU, Dr. Kang-mei LI, for providing us many advices.

Special appreciation to Dr. Quan-zhang AN for deliberate review and constructive suggestions on this paper.

References

[1] Buades, A., Coll, B., Morel, J.M.: A review of image denoising algorithms, with a new one. Multiscale Modeling & Simulation 4(2), 490–530 (2005)

[2] Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12(7), 629–639 (1990)

[3] Lindenbaum, M., Fischer, M., Bruckstein, A.: On Gabor's contribution to image enhancement. Pattern Recognition 27(1), 1–8 (1994)

[4] Rudin, L.I., Osher, S.: Total variation based image restoration with free local constraints. In: Proceedings of IEEE International Conference on Image Processing, ICIP 1994, Austin, TX, USA, November 13-16, vol. 1, pp. 31–35. IEEE Press (1994)
[5] Osher, S., Burger, M., Goldfarb, D., et al.: An iterative regularization method for total variation-based image restoration. Multiscale Modelling and Simulation 4(2), 460–489 (2005)
[6] Yaroslavsky, L.P.: Digital picture processing: an introduction. In: Information Sciences. Springer, Berlin (1985)
[7] Strela, V.: Denoising via block Wiener filtering in wavelet domain. In: Proceedings of the 3rd European Congress of Mathematics, Barcelona, Spain. Birkhäuser (July 2000)
[8] Donoho, D.L., Johnstone, I.M.: Ideal spatial adaptation by wavelet shrinkage. Biometrika 81(3), 425–455 (1994)
[9] Chipman, H.A., Kolaczyk, E.D., McCulloch, R.E.: Adaptive Bayesian wavelet shrinkage. Journal of the American Statistical Association 92(440), 1413–1421 (1997)
[10] Jung, A.: An introduction to a new data analysis tool: Independent Component Analysis. In: Proceedings of Workshop GK Nonlinearity, Regensburg, Germany, March 18, pp. 1–16 (2002)
[11] Qi, W., Qi, L., Zhe, C., et al.: Range image noise suppression in laser imaging system. Optics & Laser Technology 41(2), 140–147 (2009)
[12] Edeler, T., Ohliger, K., Hussmann, S.: Time-of-flight depth image denoising using prior noise information. In: Proceedings of 2010 10th International Conference on Signal Processing, ICSP, Beijing, China, October 24-28, pp. 119–122 (2010)
[13] Wang, J., Yao, Z.-Q., An, Q.-Z., et al.: RIDED-2D: a rule-based instantaneous denoising and edge detection method for 2D range scan line. International Journal of Pattern Recognition and Artificial Intelligence 25(6), 807–833 (2011)

Study on Race and Hazard of Combinational Logic Circuit

Wang Xiuling[1,*] and Wang Haocheng[2]

[1] College of Information Engineering, Inner Mongolia University of Technology, Inner Mongolia, 010051, China
wxltt3756@yahoo.com.cn
[2] Inner Mongolia Branch Company, Beijing Jingneng New Energy Co., Ltd., Inner Mongolia, 010070, China

Abstract. The reasons and types, detection methods, elimination and simulation of the race and hazard phenomenon in combination logic circuit are introduced and analyzed. The removing methods of race and hazard are pointed out with examples.

Keywords: combinational logic circuit, race hazard, detect, elimination.

1 Introduction

In recent years, the research on digital logic has attracted attention of logic designers because it increase the information density of circuit, and thus leading to the reduction of the number of connections and pins, and the saving of on-chip area of IC[1]. However, it also leads to the complexity in designing and analyzing circuit along with the increase of the number of signal level. For example, the race and hazard can occur in electronics systems, especially combinational logic circuits. Race and hazard is flaw in an electronic system, it brought on the interference pulse, which will make the latter circuit goes wrong, and damage original function [2]. To ensure the stability of combinational logic circuits, when designing and analyzing circuits, we must consider the reason of race hazard, test existence, and remove it.

2 Reasons and Types of Race and Hazard

In digital circuits, a gate as long as any two input signals change into the opposite direction simultaneously (i.e. from 01 into 10, or vice versa), its output may interfere with a pulse. Two input signals change into the opposite logic level transition at the same time is called race; the output peak pulse because of signals race is called hazard. The reasons of race and hazard are following.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 281–286.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2.1 Signals Can't Break

The input signals changing state need a very short transition time (skew), and different signals have different transition time, which make signals reach the logic gate not in the same time.

2.2 Signals Transmit through Different Paths

In the combinational logic, the input signals of a logic gate may be transmitted through different paths, which leads to a time difference that signals reach the output. The average gate delay will be the root cause.

The outputs are same pre and post input signals change, and the interference pulse occurred only when signals state transition, which know as static race. Static race is divided into 0-hazard and 1-hazard. Static 0-hazard is that $F = A + \overline{A}$, in theory, the output identically equals 1, in fact, there is a "0" interference pulse. Static 1 race is $F = A \cdot \overline{A}$, in theory, the output identically equals 0, in fact, the output has a "1" interference pulse [3] - [4].

Fig.1 and Fig.2 show respectively 0-hazard race and 1-hazard. In Fig.1, we input the same signals from A and B, in theory Y=1, in fact, from Fig.1, we can see there is a "0" pulse in output. In Fig.2, we input the same signals from A and B, in theory Y=0 all along, from Fig.1, however, we can see there is a "1" pulse in output. Interference pulse "0" and "1" in output will make the latter circuit goes wrong. When designing and analyzing combinational logic circuits, we should to test, prevent, and remove it.

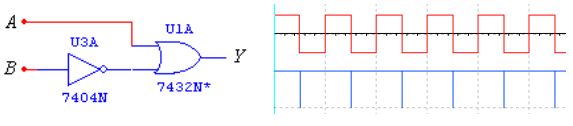


Fig. 1. (a) Example of a static 0-hazard logic; (b) static 0-hazard wave

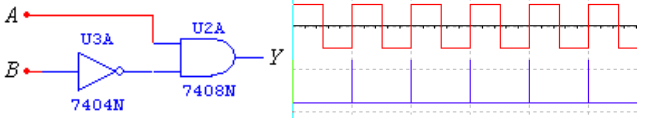


Fig. 2. (a) Example of a static 1-hazard logic; (b) static 1-hazard wave

3 Detection Methods of Race and Hazard

There are number of methods to detect race hazard of combination logic circuit. For example, logic algebra, karnaugh diagram, computer simulation, experimental observation and so on.

3.1 Logic Algebraic Method

If the two input signals, A and $\overline{A}$ of gate come from A through two different transmission paths, signal A changes state, there must exist race and hazard. So, as long as logic function of combinational circuit can be simplified as $F = A + \overline{A}$ or $F = A \cdot \overline{A}$, or rather, can be simplified as $F = \overline{A + \overline{A}}$ or $F = \overline{A \cdot \overline{A}}$ in some senses, we can judge there must be race hazard.

Using logic algebraic method to judge whether there is race and hazard of combinational circuit, we need to represent logical relationship between input and output with algebraic first, simplify the logical relationship [5]. For Fig.3, logical relationship is represented as $Y = AB + \overline{A}C$.

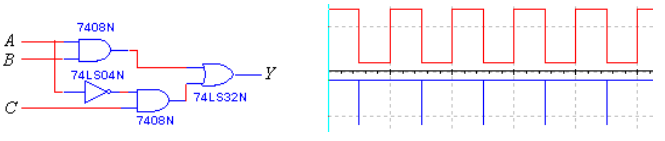


Fig. 3. (a) Example of logic algebraic method; (b) logic algebraic method wave

When $B = C = 1$, output $Y = A + \overline{A}$. According to logic algebraic method, we can judge the output will be "0" interference pulse. Using oscilloscope, we get output waves.

Using the same method, we determine whether there is race hazard of function $F = (A + B)(\overline{B} + C)$. This problem can be resolved: Let $A = C = 0$, get $F = B \cdot \overline{B}$. From logic algebraic, we conclude there is race and hazard in the combinational circuit.

3.2 Karnaugh Diagram

In the combinational logic circuit for multiple input variables, we use Karnaugh diagram method to determine whether the race and hazard phenomenon exist at two or more variables change simultaneously[6].

Using Karnaugh diagram to determine raced and hazard, checking whether there are product terms bordering on geometry, if no product terms bordering on geometry, we judge there no race and hazard, and vice versa. For example, according to the function, $F = \overline{A}\,\overline{B}\,\overline{C} + BD + AC\overline{D}$ of a combinational circuit, we get the Karnaugh diagram, Fig.4, from which we know product terms, $\overline{A}\overline{B}\overline{C}$ and BD, $AC\overline{D}$ and BD are bordering. Therefore, there are race and hazard in circuit.

According to $\overline{A}\overline{B}\overline{C}$ and BD bordering, when $\overline{A} = \overline{C} = D = 1$, then $F = B + \overline{B}$, B changing state, the output may produce interference pulse. According to $AC\overline{D}$ and BD bordering, when $A = B = C = 1$, $F = D + \overline{D}$. Likewise, D changing state, the output produces transitional interference pulse. If increase product term consisting of product terms bordering on geometry, such race and hazard can be overcame [7].

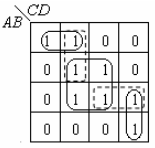


Fig. 4. Example of Karnaugh diagram method

3.3 Computer Aided Analysis

Using computer aided analysis of combinational logic circuits, running simulation program on the computer of digital circuit, we can judge whether the race and hazard phenomenon of circuit quickly. Race hazard waves of Fig.1, Fig.2 and Fig.3 are simulation waves based on simulation environment Multisim10.

4 Elimination of Race and Hazard

It has been pointed out above that race and hazard can produce interference pulses that, if fed to circuit system can also make system metastable.

There some common methods, careful design can minimize or eliminate race and hazard of circuit. Such as, introducing blocking pulse, strobe pulse, accessing filter capacitor, addition of redundancy items in design circuits, etc…

4.1 Accessing Filter Capacitor

Generally, peak pulse produced by race hazard is narrow pulse(less than tens of nanosecond). So, connecting a small capacitor (C_f) in parallel at the output end only, we can easily eliminate or minimize race and hazard, which is shown in Fig.5.

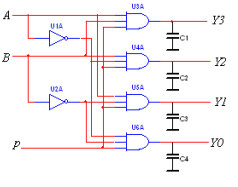


Fig. 5. Accessing filter capacitor

This method increases rise time and fall time of the output, the output waveform deteriorate.

4.2 Introducing Strobe Pulse

If the input gate of the circuit introduced a strobe plus, p, such as Fig.5, let the high level of p latter and state transition of input signal, and the blockade of the pulse width is greater than or equal to the circuit from one stable state to another, the time needed to stabilize, which will lockup the output of all gates. Then, output peak pulse

of the gate blocking the output of the circuit does not appear the phenomenon of race and hazard during signals changing state time range [8].

4.3 Addition of Redundancy Items

Modify the logic method is usually called the method of adding a redundancy items. If we use properly, we can effectively eliminate the phenomenon of race and hazard.

As shown in Fig.3, the original logic expression of circuit is $Y = AB + \overline{A}C$. According to the redundancy theory, we add a redundancy item, BC in logic function expression of circuit. The logic function expressed as following, and increased relevant gate in design the circuit.

$$Y = AB + \overline{A}C = AB + \overline{A}C + BC$$

Thus, when $B = C = 1$, no matter how A changes, there no change in output end, output maintains $Y = 1$. That is A changes the state that will no longer cause the race and hazard.

The corresponding logic circuit and the simulation output with the modify expression shown in Fig.6.

In Fig.4, using addition of redundancy items to eliminate the race and hazard, we can add redundancy items $\overline{A}BD$ and ABC. The logic function expressed as $Y = \overline{A}\overline{B}\overline{C} + BD + AC\overline{D} + \overline{A}\overline{C}D + ABC$ .It is not most simple, but optimal [9].

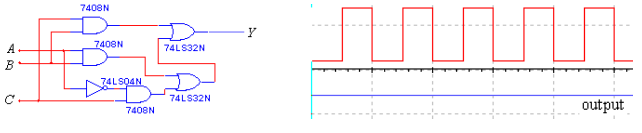


Fig. 6. Adding a redundancy items

5 Conclusion

Race and hazard phenomenon of combinational logic circuit was studied in detail. Timing in combinational logic circuits are generally caused by different input signals to a gate changing in opposite directions. Depending on path delay, this may or may not produce an interference pulse (0 or 1) at the gate output. The pulse constitutes a faulty response, which transmitted to the circuit output, results in a transient error. The result can be that the system goes wrong for an uncertain time interval. By introducing blocking pulse, strobe pulse, accessing filter capacitor, addition of redundancy items, race and hazard can be removed.

Acknowledgements. The authors gratefully acknowledge the support of College of Information Engineering of Inner Mongolia University of Technology. The paper has been supported by the Natural Science Foundation of Inner Mongolia University of Technology (ZS201006).

References

[1] Bredeson, J.G., Hulina, P.T.: Elimination of static and dynamic hazards for multiple input changes in combinational switching circuits. Information and Control 20, 114–124 (1972)
[2] Eichelberger, E.B.: Hazard detection in combinational and sequential switching circuits. IBMJ 9, 90–99 (1965)
[3] Huffman, D.A.: Design of hazard-free switching circuits. J. ACM 4, 47–62 (1957)
[4] Tietze, U., Schenk, C.: Electronics Circuit – Design and Application. Springer (1991)
[5] Hurst, S.L.: Multiple-valued logic - Its status and its future. IEEE Trans. on Computer, 1160–1179 (1984)
[6] Garrod, S.A.R., Borns, R.J.: Digital Logic – Analysis, Application & Design. Holt Rinehart and Winston, Inc. (1991)
[7] Hurtado, M., Elliott, D.L.: Ambiguous behavior of logic bistable systems. In: Proc. 13th Ann. Allerton Conf. Circuit and System Theory (October 1975)
[8] Butler, S.W., Butler, J.T.: Profiles of topics and others of the International Symposium on multiple valued logic for 1971-1991. In: Proecding of 22nd IEEE Int. Symp. on Multiple Valued Logic, pp. 372–379 (1992)
[9] Seitz, C.: System Timing, ch. 7

Approach to Analyze Time Series Similarity Pattern Mining Based on Haar

Li Yi

International Institute, Qingdao University, 266071, Qingdao, China
lyqgx@126.com

Abstract. Time series data mining is a rather complex branch, which is a technique that extracts the most valuable information from large amount of history time series data. Haar wavelet technology will be used in time series similarity-pattern mining. This dissertation discusses the problem on time series similarity-pattern mining from similarity measurement, storage structure, searching integrality. is a rather complex branch, which is a technique that extracts the most valuable information from large amount of history time series data. Haar wavelet technology will be used in time series similarity-pattern mining. This dissertation discusses the problem on time series similarity-pattern mining from similarity measurement, storage structure, searching integrality. Finally simulation results show that the Haar wavelet technology used in time-series similar pattern mining in a good performance. Finally simulation results show that the Haar wavelet technology used in time-series similar pattern mining in a good performance.

Keywords: Time Series, Haar Wavelets, Similarity Pattern Mining.

1 Introduction

With the rapid development of information, both in scientific research or in daily life will accumulate a large amount of data. Start urgently requires a large number of data from these findings as soon as possible valuable information. Thus, the database management system to store data, using machine learning methods to analyze data, this combination led to knowledge discovery in databases (KDD: *Knowledge Discovery in Databases)* came into being. Data mining is 80 years from the 20th century began, and gradually become a knowledge from the data mining tools. Data mining can also called knowledge discovery in databases, is the current international research on artificial intelligence and databases the most dynamic emerging field. The goal is to satisfy user demand, automatic processing of large amount of raw data, identify important and meaningful from the model, be expressed as knowledge of its. However, it has been found stored in the data, not just digital data, some special types of data.Increasing the proportion accounted for, which generate time database, spatial database, multimedia databases, application-specific database system. Time-series data is one very important branch of time series data in all areas of social life have a wide range of applications, such as stock information, shopping malls, supermarkets, sales information, weather information, the information

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 287–296.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

contained in ECG, the use of power system electrical connection information. Thus, time-series data mining is especially important.

Similar patterns of time series time series data mining is an important branch of mining. Time Series Similar Pattern Mining based on time series data to explore similar relationships as the main task of data mining can be used most of the methods used, such as rough sets, clustering and so on. Similar patterns of time series data mining in reality has a wide range of applications, such as astronomers can record spatial database of stars and planets similar time series data analysis, can help astronomers discover new stars; on the power system of the time Sequence similarity analysis of data to help power system technologies and scheduling personnel found to have similar consumption pattern of the load, which will be the time to properly adjust the load of electricity to meet power system economic operation [1]; in a period of time, according to similar relation between the stock of A shares rise, with the Department and its B shares a plate will be, or not in the same plate with the A shares of C stock would happen? Even possible to analyze the A shares rose 10%, B shares will increase or decrease the percentage? Who can buy shares for reference; analysis of a supermarket to buy a season of data, such as 10% of customers buy milk at the same time to buy bread, 20% of both the purchase of milk bought eggs. Analysis of the relationship through such similar businesses can give help and so on. Therefore, time series data similar to the pattern mining as a promising urgent practical need, a hot research topic.

2 Similar Pattern of Time Series Data Mining Method Statements

2.1 Time Series and Sequence Data Mining Classification and Related Methods

Time-series data and sequence data mining has several important aspects: trend analysis, similarity search, time-related data, sequential pattern mining and pattern mining cycle.

1) Trend Analysis

How to deal with time series data, at present there are four kinds of general major changes or components for special time series data: a) Long-term or initial change: It is used to reflect the general direction of change, the timing diagram shown in Figure 1, it describes a point of changing circumstances over time, it reflects the longer time interval of data changes. This change reflected a trend line or trend curve in Figure 1 using dotted lines.

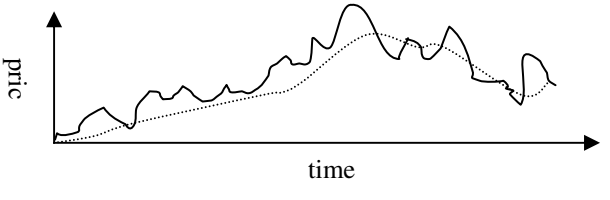


Fig. 1. Time-series data

b) cyclic changes or cycle changes: mainly refers to cycle, that is the trend line or curve in the long period of time showed signs of swinging, it may or may not be cyclical. That is, in such time interval between the cycle of evolution do not follow the same pattern.

c) seasonal variation or seasonal variation: the same is the same or similar model years in a row during the month of repeated

d) irregular or random change: it reflects the random or accidental events as sporadic due to timing changes.

Timing analysis can also refer to the time series is divided into four basic movement in front of the analysis. Thus, timing variable Y can be manifested as: Y = T * C * S * I, or Y = T + C + S + I, their choice is usually rule of thumb of. Where T, C, S, I represent the trend, cycle, seasonal, irregular variables. Through the trend, cycle, seasonal, irregular changes in the system analysis components, people can be more reasonable circumstances, to develop a long-term or short-term the forecasts (ie forecasting time series).

2) The similarity search in time series analysis

Usually a database query is to identify the exact data match the query, but for timing analysis similar to the search is to find the closest to a given query sequence data series. Subsequence matching is to find a given sequence similarity with all the data series, while the overall sequence matching is to find sequences similar to each other. Similar time series data analysis, commonly used Euclidean distance as the basis for calculating similarity. When there is enough between them similar to non-overlapping sequence of timing on when the two sequences are considered similar. It has several basic ways:

a) data transformation methods, transformation from the time domain into the frequency domain, this is because the technology needs of many signal analysis of data from the frequency domain. Common data transformation is independent of the discrete Fourier transform (DFT) and discrete wavelet transform (DWT).

Sequence matching pairs, each sequence is first divided into windows of length W "fragment." Feature space for each sequence mapped to a "clue." Pair sequence analysis, the clues for each sequence is divided into "sub-thread", each a minimum bounding rectangles.

b) Increased similarity search method to deal with offset and amplitude of the gap and differences. Most of the practical application does not necessarily require a sub-sequence matching exactly the same timeline, that sub-sequence with the same shape, but the memory in sequence or in the gap or offset differences in amplitude, we can think that they are match. This difference of situation can be judged despite similarities, is an improvement of the similarity model, which is to allow users or experts some parameters, such as the sliding window Chicun, similar to the scope of the Kuan Du, Zuitaijianxi, Pipei fragments and so on. Processing shift and amplitude differences between the gap and similar search steps are: atomic matching (to identify all non-clearance of the small window on the same), combined with the window (the same window, forming a large sub-sequence similarity, which allow a gap between the atomic match), sort sequence.

c) the index of similarity search method for large databases to improve the efficiency of similarity search, indexing have been proposed various methods. Such

as R-tree, R^* - tree, and so on, they are used to store the minimum bounding rectangle in order to speed up the similarity search. In addition, the proposed KDB tree method, which is used in high-dimensional space similar to the connection point to improve the speed, also raised the suffix trees.

3) Sequential pattern mining

Sequential pattern mining is the mining events or other modes of relatively high frequency mode. Since many business transactions, fax records, weather data and production data over a sequence of events, in targeted markets, attract customers, weather forecasts, data analysis, sequential pattern mining is useful purposes.

a) sequential pattern mining conditions and parameters

Many of the sequential pattern mining pattern aimed at symbols as digital curve model usually statistical timing analysis is trend analysis and forecasting areas.

For the sequential pattern mining, there are some parameters, how its value will seriously affect the mining results. The first parameter is the duration of the time series T. Sequential pattern mining is therefore limited to a specific duration of the excavation. The second parameter is the event overlapping window W. Occur within a specified time period a group of events, can be considered as an analysis of the events occur together. The third parameter is the pattern found in the interval between the time of the incident int.

b) sequential pattern mining method

Association rule mining properties used in the Apriori mining sequential patterns can be used, therefore, sequential pattern mining methods are used most of the class variant of Apriori algorithm. Another excavation of such models is based on the projection of sequence database growth model.

2.2 Similar to the Pattern Mining Problem

Time Series Similar Pattern search process is also known as time series similarity search.

Time series similarity search problem can be summarized as follows: for a given search sequence S, when searching sequence Q, the similarity measure function sim (), similarity search strategy for find (), similarity search process is to be searching sequence Q, search sequence S to find all similar sequences set R, namely [1]

$$R = (x \quad Q, \text{find } (\text{sim } (S, x), Q)) \tag{1}$$

From (1), we can see, the time series similarity search problem mainly involves three aspects, namely, similarity measure, search and find the complete storage strategy of [1].

Similarity measure is a measure of both the standard and basis of sequence similarity is the basis of similarity search problem. But not any of the conditions can become the standard similarity measure, similarity measure must meet a condition that the two sequence similarity measures calculated by the results of less than or equal to the real situation of the two sequences calculated by the same value. For example, we Euclidean distance formula, for example, the two series A, B of the Euclidean distance must be less than equal to sequence A, B the distance between the original, that is $D_{(Euclidean)(A, B)} \quad D_{(True)(A, B)}$. This avoids "False alarms" of the place, and can reduce the cost of similar search of work.

At home and abroad are using similar metrics as well as specific measure has the following main ideas.

1) Distance-based similarity measure [2]
Given a computing formula for the distance between time series and to determine a distance threshold for any given two sequences, when the distance is less than equal to the threshold, then that the corresponding similarity between the two, otherwise, that the two are not. Because the length n of the time series can be seen as n-dimensional space point, so space can be used to sequence distance function distance formula. Mainly in the following categories (see formula 2-6).

- Mean Character Difference:

$$\frac{1}{n}\sum_{i=1}^{n} | x_i - y_i | \tag{2}$$

- Minkowski Metric:

$$D(X,Y)=(\sum_{i=1}^{n} | x_i - y_i |^r)^{\frac{1}{r}} \tag{3}$$

Manhattan Distance(r=1)

$$D(X,Y)=\sum_{i=1}^{n} | x_i - y_i | \tag{4}$$

Euclidean Distance(r=2)

$$D(X,Y)=(\sum_{i=1}^{n} | x_i - y_i |^2)^{\frac{1}{2}} \tag{5}$$

Chebyshev Distance(r=∞)

$$D(X,Y)=\max_{1\le i\le n} | x_i - y_i | \tag{6}$$

2) The sequence of transformation [2]
The basic idea of sequence change is to allow more sequences to be due to the transformation sequence similarity measure. Sequence is a class effect real change in the sequence of simple functions. First translation (Shifting) and scaling (Scaling) transform applied to time series, the six kinds of sequence change are summarized as follows (Figure 2). To illustrate the convenience of the sequence as a function of time t, with f (t) expressed sequences.

Shifting: SHK (f (t)) = f (t) + k, where k is a constant.
Uniform Amplitude Scaling: UASK (f (t)) = kf (t).
Uniform Time Scaling: UTSK (f (t)) = f (kt), where k is the normal number.
Uniform Bi-scaling: UBSK (f (t)) = kf (t / k), where k is the normal number.
TimeWarping (Non-uniformTimeScaling): TWg (f (t)) = f (g (t)), where g is a monotone increasing function.
Non-uniformAmplitude caling: NASg (f (t)) = g (t), which for any t, g '(t) = 0 if and only if f' (t) = 0.

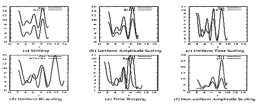


Fig. 2. [5] Sequence transform

3) The longest common subsequence metric [2]
Longest common subsequence metric (Longest Common Subsequence measures abbreviated as LCS) The basic idea is to use two time series have non-overlapping sub-sequences of the length of the public to measure the degree of similarity between the two series. Simply put: If the sequence X = 3,2,5,7,4,8,10,7 and sequence Y = 2,5,4,7,3,10,8,6, then the LCS = 2, S, 7,10, X and Y sequence similarity Sim (X, Y) = | LCS | = 4.

Practical research in the multi-use and sequence Transform class-LCS metric, the basic idea is: if two time series with a sufficiently long sequence of non-overlapping sub-paragraph similar, then the two sequences can be considered similar.

4) Image-based slope angle of the similarity measure [3]
The sequence is divided into many segments, if the difference between the two fragments of the absolute value of inclination is less than the threshold, we think that the two fragments are similar. When two sequences are not similar to the number of small fragments of a particular given threshold, we think the two sequences are similar.

5) Based on the slope of tangent value of similarity measure [4]
First time series is divided into N sub-sequences, each sequence contains the t lines segments, time-series curve with piecewise linear representation. And calculate the slope value of each line, if the slope of the two sub-sequences arctangent calculated to meet a certain threshold during a given that the two sub-sequences are similar.

6) The shape features based on time series similarity measures [5]
The basic idea is that the sequence of piecewise linear and symbolic. Size according to given time series curve of the linear simulation, the adjacent two line segments connected by a line under the sequence, the folded line segment along with a representative of each of the shape of the symbol. This sequence of similar problems put into question the similarity between symbols.

3 Haar Wavelet Time Series Similar to the Model Used in the Basic Principles of Mining

Haar wavelet function is a set of mutually orthogonal normalized set of functions. Haar wavelet is the set of functions derived from this group obtained, is supporting the domain in the $t \in [0,1]$within a single square wave. As However, Therefore, $\int \psi(t)dt = 0$, then $\int t\psi(t)dt \neq 0$. Haar wavelet function as the specific structure, namely:

$$\psi(t) = \begin{cases} 1 & 0 \le t \le 0.5 \\ -1 & 0.5 \le t < 1 \\ 0 & \text{others} \end{cases} \tag{7}$$

Scaling function as:

$$\phi(t) = \begin{cases} 1 & 0 < t < 1 \\ 0 & \text{others} \end{cases} \tag{8}$$

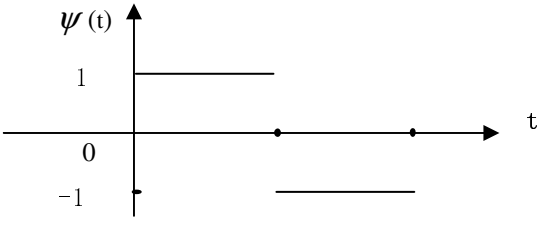


Fig. 3. The mother function of Haar wavelets

Shown in Figure 3, Haar wavelet basis function is horizontal in the [0,0.5] range for the vertical to a horizontal line 1, in the [0.5,1] range for the vertical as a horizontal line -1 the other part is 0. The original data and the role of the wavelet function, image and positive and negative absolute value of the two longitudinal coordinates of the same linear function, can offset the image in a very short period of time frequently repeated shocks (which is part of the image details), while Image Central Africa will not be part of frequent shocks offset, to retain (ie, scale part). This is the use of Haar wavelet time-series data reduction ideas and principles.

4 Performance Evaluation

The data source used in the experiments is a province of China from 1979 to 2001, continuous meteorological data, the data is typically time-series data, covering all sites across the province, various meteorological parameters, the data is very large. The following sequences were from the match and nearest neighbor search on the search conducted two experiments, the results are as follows.

4.1 Matching for Search Sequence

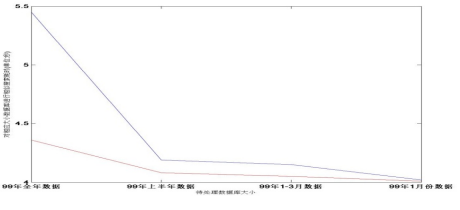


Fig. 4. The compare of efficiency in difference of S length in All-pairs Query

In Figure 4 red line is the result Haar wavelet time series data reduction sequence matching the search after the time-consuming situation, the blue line is not the result of wavelet transform time series data on the direct reduction of the same match sequence search time-consuming situation. First of all, from this experiment can be seen through the Haar wavelet time series data reduction, whether it be the size of the number of search sequences, to the subsequent search for matching sequences on the time to bring increased efficiency is obvious, and be accompanied Search sequence of larger size, the degree of efficiency more clearly. So wavelet transform for high capacity data reduction efficiency is high. In this experiment also received another conclusion, that is, whether for the first time using the Haar wavelet series data reduction, during the same match sequence search to the conclusion is the same. This is further evidence of the use of wavelet transform time-series data reduction does not change the time series similarity search results, also confirmed the use of wavelet transform time-series data reduction in the accuracy and feasibility.

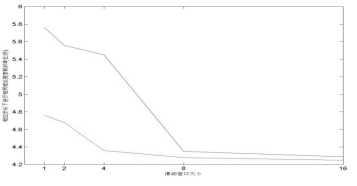


Fig. 5. The compare of efficiency in difference of step length in All-pairs Query

Figure 5, two lines of the meaning of an experiment Ibid. First of all, from this experiment can be seen through the Haar wavelet time series data reduction, regardless of the number of step size to match the sequence was carried out by the time the search has brought efficiency are obvious, and the smaller step , the greater the degree of efficiency. Another experiment will also be given in: whether to first use the Haar wavelet time-series data reduction, during the same match the sequence of the search when the results obtained are the same.

4.2 Nearest Neighbor Search

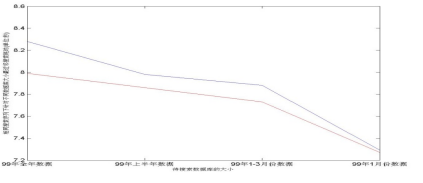


Fig. 6. The compare of efficiency in difference of S length in Nearest Neighbor Searches

First of all, from this experiment can be seen through the Haar wavelet time series data reduction, whether it be the size of the number of search sequence length, was carried out to bring the time of the nearest neighbor search efficiency are visible, and

to be search sequence length increased, the greater the degree of efficiency. Another experiment will also be given: whether or not the first to use Haar wavelet time-series data reduction, during the same nearest neighbor search, the results obtained are the same.

5 Conclusion

Brief statement of the time series similar pattern mining methods, and meteorological data on a large number of experimental results show that the Haar wavelet in Time Series Mining mining can greatly improve the efficiency of matching.

References

[1] Gedik, B., Singh, A., Liu, L.: TangLiang. Time Sequence Mining and Similarity Search Technology (Master degree theses). Shanghai Teachers University (2004)
[2] Yang, M.: . The Research and Application of Similarity in Hydrological Time Series Model. (Master degree theses). Hohai University (2002)
[3] Zheng, C., Cai, Q.S.: A Algorithm of Multi-scale Time Sequence Similar Pattern Matching. Journal of Mini-Micro Systems (2003)
[4] Cai, Z., Yue, L.H., Wang, X.F: A Research on Sequential Patterns Found Algorithm. Journal of Computer Research and Development
[5] Yang, M., Wang, Z.J., Yin, Y.M.: A Research on Time Sequence Comparability Search Algorithm. Journal of Shandong Teachers University
[6] Keogh, E., Smyt, P.: A Probabilistic Approach to Fast Pattern Matching in Time Series Databases
[7] Das, G., Lin, K.-I., Mannila, H.: Rule discovery from time series. In: Proc. of the Fourth int'l Conf.on KDD, pp. 16–22 (1998)
[8] Jiawei, H., Guozhu, D., Yiwen, Y.: Efficient Mining of Partial Periodic Patterns in Time Series Database
[9] Zhang, J., Zhu, M., Papadias, D.: Location-based Spatial Queries (2003)
[10] Chan, K.-P., Fu, A.W.-C.: Efficient Time Series Matching by Wavelets (1999) [7], Lee, S., Kwon, D., Lee, S.: Efficient Similarity Search for Time Series Data Based on the Minimum Distance, 5 (2002)
[11] Liang, J., Chen, L., Sharad, M.: Efficient Similarity String Joins in Large Data Sets Efficient Similarity String Joins in Large Data Sets
[12] Miguel, A., Ari Ari, O., Brani Vidakovic, O.: On Wavelet Scalograms and Their Applications in Economic Time Series On Wavelet Scalograms and Their Applications In Economic Time Series
[13] Mong, L.L., Wynne, H., Jensen, C.S.: Supporting Frequent Updates in R-Trees:A Bottom-Up Approach. In: Proc.the 29 th VLDB Conference (2003)
[14] Kahveci, T., Singh, A.: Optimizing Similarity Search for Arbitrary Length Time Series (2003)
[15] Kahveci, T., Singh, A.: Variable Length Queries for Time Series Data (2001)
[16] Tang, L.: Time series mining and similarity search technology research [MS Thesis]. Shanghai Normal University (2004)
[17] Tsai, C., Yue, L., Wang, X.L.: Sequential Pattern Discovery Algorithm. Computer Research and Development 37(9), 1107–1113 (2000)

[18] Wu, Y.-L., Divyakant, A., Abbadi E.: A Comparison of DFT and DwT Based Similarity Search in Time—Series Databases. In: Proceedings of the 2000 ACM CIKM International Conference on Information and Knowledge Management, Mclean, VA, USA, pp. 488–495 (November2000)
[19] Jin, X., Lu, Y., Shi, C.: Similarity Measure Based on Partial Information of Time Series (2002)
[20] Zhang, X., Thearling, K.: Non-Linear Time-Series Prediction by Systematic Data Exploration on a Massively Parrallel Computer
[21] Yang, M.: Hydrological time series similarity model and application [MS Thesis]. Hohai University (2002)
[22] Min, Y., Zhijian, W., Yanmin, Y.: Time series similarity search algorithm. Shandong Normal University 16(4), 348 (2001)
[23] Faloutsos, C., Ranganathan, M., Manolopoulos, Y.: Fast Subsequence Matching in Time-Series Databases; Moon, Y-S., Whang, K-Y., Loh, W-K.: Efficient Time-Series Subsequence Matching using Duality in Constructing Windows (2000)
[24] Zhang, J., Zhu, M., Papadias, D., [11] Yufei, T., Dimitris, P., Xiang, L.: Reverse kNN Search in Arbitrary Dimensionality. In: Proceedings of the 30 th VLDB Conference (2004)

Design of a Four-Annulus Integrated Disk Electrode for Bioelectrical Impedance Spectroscopy Measurement

Kening Wang*, Xiongxing Zhang, and Wei Wang

School of Optoelectronic Engineering Xi'an Technological University
N0.4, Jin hua bei Road, Beilin District, Xi'an City, Shaanxi Province
Xi'an, China
wangkening2000@sohu.com

Abstract. In order to study the relationships between pork quality and the BIS of porcine tissues, a novel four-annulus integrated disk (FAID) electrode was designed and manufactured. The contrastive measurement experiments on two fresh isolated animal tissues, namely porcine underback and porcine liver using the FAID electrode were performed. The results show that the impedance amplitude-frequency characteristics and phase-frequency characteristics of the two kinds of tissues have significant differences. The Cole-Cole parameters fitting are then performed using the measured BIS data. The fitting results show that the arrangement of the BIS data on the complex plane is in good concordance with the Cole-Cole plot respectively, and the extracted results, R0, R∞, α and fC, also indicate the difference between the two tissues.

Keywords: bioelectrical impedance spectroscopy, pork quality, four-annulus integrated disk electrode, Cole-Cole model.

1 Introduction

Pork quality has an essential effect on people's life and health, thus the incidence of quality problems such as diseased or decayed pork should be effectively controlled, and pork quality parameters must be quantified quickly and incorporated into the market information. Therefore, it is important to develop fast and accurate instrumentation to detect pork quality.

As a quick, affordable, portable and harmless technique, bioelectrical impedance spectroscopy (BIS) is proved in recent studies that it is a reliable diagnostic tool to monitor pathological status of biological tissues [1-3]. The BIS technique is also used to detect meat quality such as porcine spongiform encephalopathy (PSE) [4] and lamb carcass composition[5].

In the project for fast detection of pork quality, we have developed a portable spectrometer which can make BIS measurement of porcine tissue impedance over frequencies from 20 kHz to 1MHz. Since the capacitive nature of cell membranes accounts for the frequency dependence of porcine muscle impedance, BIS measurements may provide insight into the structure and viability of porcine muscle tissue. Low

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 297–303.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

frequency BIS measurements will have a large reactive component and result in greater impedances with larger phase angles, while high frequency BIS measurements will have a low reactance of the cell membrane capacitance[4].

BIS measurement is usually based on the four-electrode method, as shown in Fig. 1. For this method, the two outer electrodes are used to inject the current into the tissue. The injected current then forms equipotential surfaces within the tissue. The inner electrodes then measure the resultant voltage. The measured tissue impedance is then calculated as: Z=(VIN+–VIN-)/IOUT [6]. Separation of the current injection and signal detection electrodes virtually eliminates the influence of the electrode tissue impedance. The current injected to the tissue is independent of the measured impedance[7].

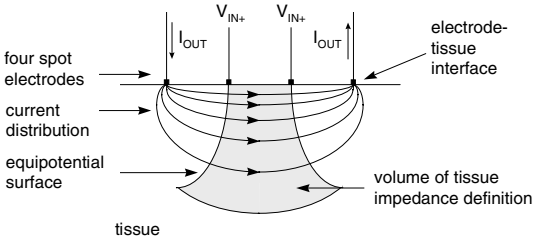


Fig. 1. Four-Electrode Method of BIS Measurement

According to the published literature, the four electrodes can be arranged linearly [8], quadrangularly [9, 10] or annularly. Osterman et al [11] designed compound annular electrodes, and Nebuya et al [12] designed ring-shaped electrodes, but both were separated into two discrete electrodes. This paper proposes a four-annulus integrated disk (FAID) electrode for non-invasive BIS measurements, and the preliminary experiments on isolated porcine underback and liver are performed.

2 Electrode Disign

The designed FAID electrode is shown in Fig. 2, which is composed of four concentric annuluses in a disk plane. The excitation current source is injected through the outermost annular electrode HC, and the innermost annular electrode LC is connected with the power ground to call back the current source. The two annuluses in the middle, HV and LV, are the voltage sensing electrodes and usually connected with an instrumental amplifier. The area of the current electrodes HC and LC are bigger than the voltage electrodes HV and LV, which will helpful for current injecting. The distance between the current electrode and voltage electrode is 1 mm and about ten times wider than the interval between biological cells, which is designed to make the current density in the measurand under the two voltage electrodes be homogeneous [13].

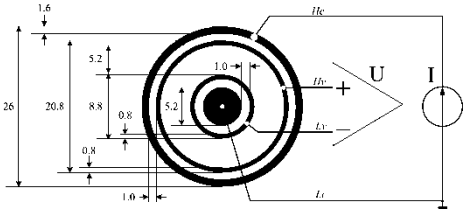


Fig. 2. The structure and dimension of the FAID electrode (unit: mm)

The FAID electrode is designed in PCB (printed circuit board) software Protel 99Se according to the dimension as shown in Fig. 2 and manufactured as a double layer PCB. On the bottom layer of the PCB, the four annuluses are tinned and used as the electrodes, each of which is connected to the top layer through four via holes. The four via holes are bounding pads which can be welded with wires and connected to BIS instrument. A plastic cylinder handle is also designed to fix the FAID electrode, and four interface connectors are extended. The practical photo of the FAID electrode is shown in Fig. 3.

Fig. 3. Photo of the FAID electrode

3 Measurement Experiments

3.1 Experiment Materials

In order to test the FAID electrode, we prepared two types of isolated biological tissues, the porcine underback (UB) and liver as shown in Fig. 4, for BIS measurements. The two porcine tissues were fresh and the death time was less than 4 hours. Each of the tissues was about 0.5 kilogram and under room temperature (about 25℃).

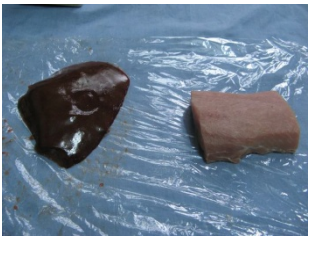


Fig. 4. The porcine liver (left) and underback (right) tissues used for BIS measurements

3.2 Experimental Results

The BIS measurements were performed at 22 frequencies as listed in Tab.1 in the frequency range from 20 kHz to 1 MHz. The measurement results were displayed as amplitude-frequency curve and phase-frequency curve, as shown in Fig. 5 (a) and Fig. 5 (b) respectively. In the figure, the measurements results of porcine liver were marked by circular points, and the measurements results of porcine underback were marked by square points.

Table 1. The 22 Measurement Frequencies from 20 kHz to 1 MHz

fi	Frequency (kHz)	fi	Frequency (kHz)	fi	Frequency (kHz)
$f1$	20	$f9$	100	$f17$	500
$f2$	30	$f10$	120	$f18$	600
$f3$	40	$f11$	140	$f19$	700
$f4$	50	$f12$	160	$f20$	800
$f5$	60	$f13$	180	$f21$	900
$f6$	70	$f14$	200	$f22$	1000
$f7$	80	$f15$	300		
$f8$	90	$f16$	400		

In Fig. 5 (a), we can observe that the amplitudes of porcine liver at all frequencies are obviously bigger than those of the porcine underback; while in Fig. 5 (b), the phases of the porcine underback are bigger than those of the porcine liver at low frequencies (below 120 kHz), and vice versa at high frequencies (above 120 kHz).

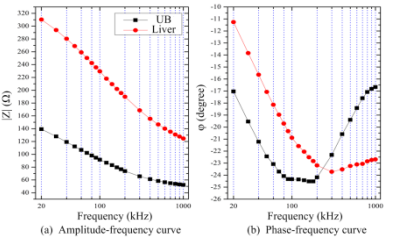


Fig. 5. Amplitude-frequency curve and phase-frequency curve of the porcine liver and underback (UB) tissues

4 Cole-Cole Model Parameter Analysis

4.1 Cole-Cole Model

The Cole-Cole model has been used successfully to describe the experimental data for the dielectric constant of many materials as a function of frequency[14]. The Cole-Cole model can be determined as:

$$Z = R_\infty + \frac{R_0 - R_\infty}{1 + (j\omega\tau)^\alpha} \quad . \quad (1)$$

Where:

R0 = impedance at zero frequency;
R∞ = impedance at infinite frequency;
τ = characteristic time constant;
α = empirical dimensionless constant from 0-1.

In Eq. (1), the time constant τ is usually replaced by another more meaningful parameter, namely the characteristic frequency:

$$f_C = \frac{\omega_C}{2\pi} = \frac{1}{2\pi\tau} \quad (2)$$

According to Eq. (1) and Eq. (2), the Cole-Cole model can be expressed as:

$$Z = R_\infty + \frac{R_0 - R_\infty}{1 + \left(j\frac{\omega}{\omega_C}\right)^\alpha} = R_\infty + \frac{R_0 - R_\infty}{1 + \left(j\frac{f}{f_C}\right)^\alpha} \quad (3)$$

So in Cole-Cole model, the dielectric constant depends mainly on four parameters, namely R0, R∞, α and fC.

4.2 Parameter Extraction of Cole-Cole Model

Kun [14] proposed an iterative least square fitting method to extract the four parameters R0, R∞, α and fC. According to this algorithm, we can obtain the concrete value of R0, R∞, α and fC from the raw data of BIS measurement experiments.

The measurement results as shown in Fig. 5 (a) and Fig. 5 (b) can be changed into real part R (resistance) and imaginary part X (reactance) at every frequency point according to the equation:

$$R = |Z|\cos\varphi \quad (4)$$

$$X = |Z|\sin\varphi \quad (5)$$

Then an iterative least square fitting algorithm was performed in Matlab software using a group of BIS data (Ri, Xi) (i=1, …, 22), and the four Cole-Cole parameters R0, R∞, α and fC were obtained. Two sets of the BIS measurement data on the porcine liver and underback were plotted in Fig. 6, and their Cole-Cole parameters were extracted by this algorithm. Two sets of Cole-Cole parameters R0, R∞, α and fC were obtained and listed in Tab. 2.

In Fig. 6, the two sets of the BIS measurement data on the porcine liver and underback match their Cole-Cole plot well, and both Cole-Cole plots have a similar locus, which is a minor arc of a circle whose center lies below the real impedance axis. It has a zero frequency intercept R0 and an infinite frequency intercept R∞.

From Fig. 6 and Tab. 2, one can observe that there are obvious differences between the Cole-Cole parameters of the porcine liver and those of the porcine underback. The impedance results have a high agree with the literature reports [15], which validates the availability of the newly designed FAID electrode.

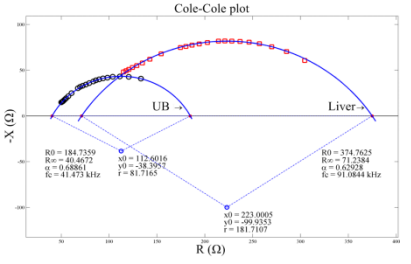


Fig. 6. The Cole-Cole plot of the BIS data on the porcine liver and underback (UB) tissues

Table 2. The Cole-Cole parameters of the porcine liver and underback (UB) tissues

measurand	Fitting results			
	R0（Ω）	R∞（Ω）	α	fc（kHz）
UB	184.7359	40.4672	0.6886	41.473
Liver	374.7625	71.2384	0.6293	91.084

5 Conclusions

The relationships between pork quality and the BIS of porcine tissues are to be studied. This paper designed and manufactured a novel four-annulus integrated disk (FAID) electrode. The contrastive measurement experiments on isolated porcine underback and porcine liver using the FAID electrode were performed. The Cole-Cole fitting results show that the arrangement of the BIS data on the complex plane is in good concordance with the Cole-Cole plot, which validates the availability of the newly designed FAID electrode. Future work will focus on studying the quantitative relationships between the BIS of porcine tissues and the pork quality such as moisture content and fresh degree.

Acknowledgment. This work has been supported by a grant from the Scientific Research Plan of Education Bureau of Shaanxi Province, China (No. 2010JK588).

References

[1] Passauer, J., Petrov, H., Schleser, A., Leicht, J., Pucalka, K.: Evaluation of clinical dry weight assessment in haemodialysis patients using bioimpedance spectroscopy: a cross-sectional study. Nephrol. Dial. Transplant. October 6, (2009)

[2] Tonkovic, S., Tonkovic, I., Kovacic, D.: Bioelectric impedance analysis of lower leg ischaemic muscles. In: Proceedings of the 22nd Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, vol. 1, pp. 757–760 (2000)

[3] Gonzalez-Correa, C.A., Brown, B.H., Smallwood, R.H., Stephenson, T.J., Stoddard, C.J., Bardhan, K.D.: Low frequency electrical bioimpedance for the detection of inflammation and dysplasia in Barrett's oesophagus. Physiol. Meas 24, 291–296 (2003)

[4] Whitman, T.A., Forrest, J.C., Morgan, M.T., Okos, M.R.: Electrical measurement for detecting early postmortem changes in porcine muscle. J. Anim. Sci. 74, 80–90 (1996)

[5] Altmann, M., Pliquett, U., Suess, R., von Borell, E.: Prediction of lamb carcass composition by impedance spectroscopy. J. Anim. Sci. 82, 816–825 (2004)

[6] Palko, T., Bialokoz, F., Weglarz, J.: Multifrequency device for measurement of the complex electrical bio-impedance-design and application. In: Proceedings of the First Regional Conference in Engineering in Medicine and Biology Society, 1995 and 14th Conference of the Biomedical Engineering Society of India. An International Meeting, pp. 1/45–1/46. IEEE (1995)

[7] Songer, J.: Tissue ischemia monitoring using impedance spectroscopy: Clinical evaluation. M.S. thesis of Worcester Polytechnic Institute in USA (2001)

[8] Aaron, R., Huang, M., Shiffman, C.A.: Anisotropy of human muscle via non-invasive impedance measurements. Phys. Med. Biol. 42, 1245–1262 (1997)

[9] Seoane, F., Lindecrantz, K., Olsson, T., Kjellmer, I., Flisberg, A., Bagenholm, R.: Spectroscopy study of the dynamics of the transencephalic electrical impedance in the perinatal brain during hypoxia. Physiol. Meas. 26, 849–863 (2005)

[10] Jones, D.M., Smallwood, R.H., Hose, D.R., Brown, B.H., Walker, D.C.: Modelling of epithelial tissue impedance measured using three different designs of probe. Physiological Measurement 24, 605–623 (2003)

[11] Osterman, K.S., Hoopes, P.J., DeLorenzo, C., Gladstone, D.J., Paulsen, K.D.: Non-invasive assessment of radiation injury with electrical impedance spectroscopy. Phys. Med. Biol. 49, 665–683 (2004)

[12] Nebuya, S., Noshiro, M., Brown, B.H., Smallwood, R.H., Milnes, P.: Detection of emboli in vessels using electrical impedance measurements—phantom and electrodes. Physiological Measurement 26, S111 (2005)

[13] Paulson, K.S., Pidcock, M.K., McLeod, C.N.: A probe for organ impedance measurement. IEEE Transactions on Biomedical Engineering 51, 1838–1844 (2004)

[14] Kun, S., Ristic, B., Peura, R.A., Dunn, R.M.: Real-time extraction of tissue impedance model parameters for electrical impedance spectrometer. Med. Biol. Eng. Comput. 37, 428–432 (1999)

[15] Faes, T.J.C., van der Meij, H.A., de Munck, J.C., Heethaar, R.M.: The electric resistivity of human tissues (100 Hz-10 MHz): a meta-analysis of review studies. Physiological Measurement 20, R1–R10 (1999)

Development of Frequency and Power Control System Hardware-in-Loop Simulation Platform for Ship Power Plant

Sun Bin[1,*], Zeng Fan-ming[1], Zhang Wei-dong[2], and Zhao Hua[3]

[1] College of Naval Architecture and Marine Power, Naval University of Engineering, Wuhan, China
[2] Department of Early Warning Detection Equipment, AFRA, Wuhan, China
[3] Naval Aeronautical Engineering Academy Qingdao Branch, Qingdao, China
sun544@163.com

Abstract. In order to research the dynamic and static performance of ship power plant, the status quo of frequency and power control are analyzed. A hardware-in-loop simulator was built based on LabVIEW platform and Programmable Logic Controller (PLC), and the hardware-in-loop simulations were performed for frequency and power control in ship power plant. Before choosing the frequency and power controller, the value of the deviation of the real load from the desired load is basically big, diesel error is 39.5%, and turbine error is 19.5%. When controller is adopted, diesel error is 12.8%, and turbine error is 5.3%.The experiment results show the platform provides a new effective method for development of the frequency and power control, because of its merit such as shortening lead time and saving development funds.

Keywords: Ship power plant, Hardware-in-loop Simulation, frequency and power controller.

1 Introduction

Along with deep research of all-electric system and extensive use of new high-power electronic equipment, ship has rapidly developed toward large-scale and automation [1]. It brings challenge for the capacity of ship power plant, power quality and reliability. For the Ship Power Station control system, the traditional mathematical simulation hardly connects with actual hardware facility, so it is difficult to reflect the actual situation of the system. The other simulation way, the physical simulation is high cost and long cycle, for some extreme trials, it can not proceed.

With the rapid development of the computer technology and software technology, Hardware-in-the-loop Simulation is by progress development, which integrates the advantage between physical simulation and mathematical simulation[2-4]. Within HIL environment, a real hardware to be developed interacts with a virtual system that

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 305–312.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

replaces part of a real system or component based on the computer interface technology[5,6]. Currently, popular tools of hardware in the loop simulation are mainly: dSPACE[7,8], xPC Target[9], and LabVIEW-RT[10]. As a software development integration environment, LabVIEW-RT contains a wealth of subroutines colluding data collection, data analysis, and data signal control. It makes modular program more simple and intuitive, and easy to debug, understand and maintain, so it has the most widely application and the most strongly function in the automotive industry.

In this paper, the method of HIL simulation is explored as a tool for analyzing frequency and power control of ship power plant systems. The turbine-diesel power station control system is taken as an object to establish the hardware-in-the-loop simulation platform, with which the frequency and power control strategy is investigated. The remainder of this paper is organized as follows. Section 2 provides an overview of system structure and system design. Section 3 details the development of hardware to represent frequency and power controller. Section 4 goes on to develop software design concluding system model, plc Control program and human machine interface. Section 5 then discusses the system simulation and experiment in the LabVIEW environment. Section 6 presents our conclusions and suggests possible areas of further research.

2 Design Project of System

2.1 Structure and Principle of Control System

Turbine-diesel power station control system consists of steam turbine speed control system, diesel engine speed control system and frequency-power control system. The core of frequency and power control system is to control the two governor control unit. Therefore, the two subsystems (turbo-generator speed control system and diesel generator speed control system) not only are the speed control system to run their own separate, but also are control objects of frequency-power controller in the parallel operation. Frequency-power controller controls two prime movers through the two subsystem for achieving stability grid frequency and balanced power distribution.

2.2 Design of Hardware-in-Loop Simulation System

Hardware-in-loop simulation platform of the frequency-power control system is illustrated in Figure 1. The HIL simulation platform consists of three parts: the simulation computer, interface circuit and programmable logic controller (PLC). Within the simulation computer, the model of diesel generator set and turbine generator set are developed, and the speed (frequency), voltage, power and other values of system are simulated and calculated. Interface circuit consists of PCI6221 acquisition card, USB6008Signal detection and Signal conditioning board. It has analog channels, digital channels, the counter output channels. It gathers and processes signals from the computer, the processed signals then send to PLC device. In the PLC, the integrated error signal of frequency and power is calculated according to their values. It turns to PWM signal by control algorithm.

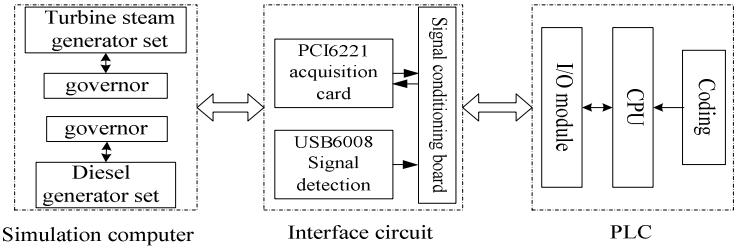


Fig. 1. Structure of HIL Simulation platform of frequency and power control system

As a feedback signal, the PWM signal which feed directly to model of the diesel and turbine set in the simulation computer in order to achieve frequency and power control. Hardware-in-loop simulation system make use of the LabVIEW virtual instrument function, at the same time make use of the LabVIEW Simulation Module to model, which achieves an industrial grade (anti-jamming), real-time hardware-in-loop simulation.

3 Hardware Design

Hardware design include: simulation computer, programmable logic controller (PLC), acquisition card and signal conditioning board.

3.1 Selection of Computer

Advantech Industrial Computer is selected which install Windows XP, it equips a 2.2 GHz Intel Pentium 4 processor, 512 MB DDR memory and 10/100BaseTX Ethernet. Advantech Industrial Computer is mainly used for initial preparation, information exchange, mathematical modeling, running the real-time simulation model, and data input and output through multi-board I/O channel.

3.2 Selection of PLC

Mitsubishi FX2N is selected, because it is the most powerful in FX series and the highest rate of micro-PLC. It mainly consists of CPU module, input modules, output modules, and the composition of the programmer. There are two basic working modes: run mode and stop mode. In the run mode, PLC implements control functions through repeating implementation of the user program. This cycle of work was known as the scanning work. FX2N basic instruction execution time is up to 0.08μs. User memory is 8K steps. And it can be expanded to 16K steps. It can be expanded to 256 I/O points, and there are a variety of special function module and a function expansion board.

3.3 Selection of Acquisition Card

Acquisition card adopts NI's PCI6221 data acquisition card as mainly output and USB6008 signal acquisition module as an auxiliary output. As Multifunction data

acquisition board, PCI6221 comprises 16-channel analog inputs (16-bit A/D), 2-channel analog outputs, 24 channel digital input and output, and two counter/timers. USB6008 is a plug and play data acquisition module based on USB bus, which have 4-channel analog inputs, 2-channel analog output, 12-channel digital input and output, and a counter.

3.4 Signal Detection and Processing

The voltage range of data acquisition card input and output signal is -5V-+5V or -10V-+10V, the voltage of PLC input signal is 24V, therefore 7406 gate circuit is adopted in order to connect data acquisition card with PLC. This circuit upgrades the data acquisition card voltage to 24V. Circuit principle is shown in Figure 2.

To utilize PLC high-speed pulse output, PLC output is adopted transistor output. When Mitsubishi FX2N transistor output is ON, its output voltage is about 1.5V. Data acquisition card can not distinguish between high and low level. Therefore CMOS gate circuit is adopted in order to reduce lower level to 0V. Circuit principle is shown in Figure 3.

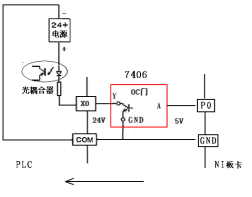


Fig. 2. Voltage changeover circuit

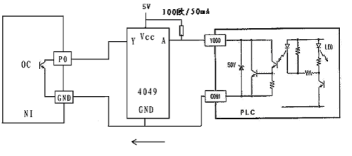


Fig. 3. Voltage changeover circuit

Frequency signal detection and processing: the timer interrupt is adopted in order to obtain speed pulse signal, and then the engine speed is calculated using regular time intervals. Firstly speed pulses are transformed, and then puts into the PLC high-speed counter inputs, the PLC accounts pulse times, speed signal is accounted from the counter module.

Power signal detection and processing: PLC analog input module is a key component of generator power detection, FX2N-4AD Analog Input module provides high-precision 12bit resolution, 4-channel voltage input (-10V ~ 10V DC) or current input (-20mA ~ 20mA DC). Each channel can be defined as voltage input or current input. It exchanges data with the main PLC unit through the buffer memory. In the buffer memory, channel, sampling rate and average mode are chosen, and sampled values can be read from the specified memory cells.

Turbine and diesel generator power signals export to the PLC analog input module (A / D module) Through the 6008 module; 6221 data acquisition card counters can generate two-way 1 ~ 60kHz square wave signal, it not only can simulate speed signal from normal magnetic-electric speed sensor, but also simulate speed signal from analog digital encoder. This speed signal is sent to the PLC high speed counter interface X0, X1. By this way, the PLC devices can obtain power and frequency error signal.

The integrated error signal is calculated in PLC. The PWM signal is sent to NI acquisition card analog input ports from the high-speed pulse output Y0, Y1. It adjusts turbine and diesel generators speed in order to complete the adjustment of frequency and power.

4 Software Design

Software design is core in the hardware in the loop simulation platform. In this paper, software design includes turbine - diesel generator set model, PLC control procedures, as well as human-computer interface design.

4.1 System Model

Turbine and diesel generator set model is basis in system simulation platform, but how to connect simulation model with hardware is the key to building a simulation platform. Modeling In the LabVIEW environment can be relatively easy to implement I / O. Structure diagram of ship power plant simulation model is shown in Figure 4.

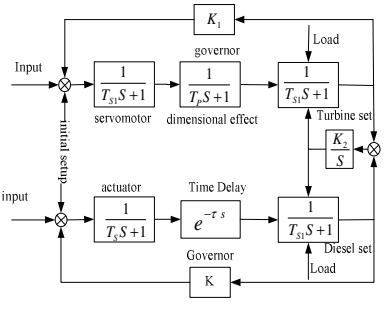


Fig. 4. Model of the diesel-turbine generation set

4.2 PLC Control Program

CFX-GP/WIN-C software is adopted to achieve programming and debugging in Mitsubishi FX series PLC. This software can be carried out under Windows3.1 and Windows95/98/xp operating systems. It has off-line programming, document management, program transmission, operation and monitoring function.

4.3 Design of Human-Machine Interface

Human-machine interface is designed in the labview environment. It includes "Dynamic Test", "Results", "Back" and others control button. Press the "dynamic test" button, and enter the test interface. it is shown in Figure5. In the test interface, dynamic change curve and relevant data of power - frequency are displayed when in parallel running, load and disturbance can be also set. At the same time, in test control interface, press the "Start Test" button, the computer applies step load disturbance to

system in order to complete a test. For more tests, press the "repeat" button for the next step disturbance tests. Industrial computer calculates and extracts characteristic parameters of diesel and turbine-generator set, and on the computer screen, debug parameter and debug curves is displayed.

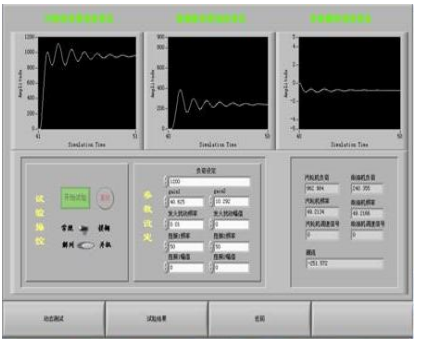


Fig. 5. Dynamic Test Interface

So it can accurately analysis performance and intuitive represent ability, it provides an effective means to accurately determine controller performance and control strategy effect.

5 Simulation and Experimentation

The prime mover of ship power station can not do exactly the same speed droop, otherwise, after long term work, the working conditions of governor had deteriorated. Those are reason that sharing load is uneven. In hardware in the loop simulation platform of ship power plant, in order to verify the project validity, the speed droop of prime mover governor is adjusted in human machine interface to simulate the uneven sharing load operating conditions,

Test design: steam turbine rated power is 1200kW, and diesel engine rated power is 600kW. At 0s, simulation system runs at empty load. At 1s, a load of 1200 kW was applied to power plant. At 5s, the frequency and load of simulation system go steady. For frequency of electric system, the simulation value is 49.2Hz; compare to 50Hz at the theoretical value, the error is 1.6%. Turbine set bears a 956kW load; compare to 800kW at the theoretical value, the error is 19.5%. Diesel set bears a 242kW load; compare to 400kW at the theoretical value, the error is 39.5%. At 11, load-frequency controller works, frequency is still 49.2Hz, and changes little. Turbine set bears a 842kW load; compare to theoretical value, the error is 5.3%. Diesel set bears a 348kW load; compare to 400kW at the theoretical value, the error is 12.8%.

Results analysis: when load-frequency controller does not work, the frequency of electric system declines, and the load distribution was highly uneven. After the controller runs, frequency adjustment result is basically satisfactory, the load distribution improves. The simulation results show the conventional PID control strategy is basically contented for frequency control. Simulation result is shown Figure 6.

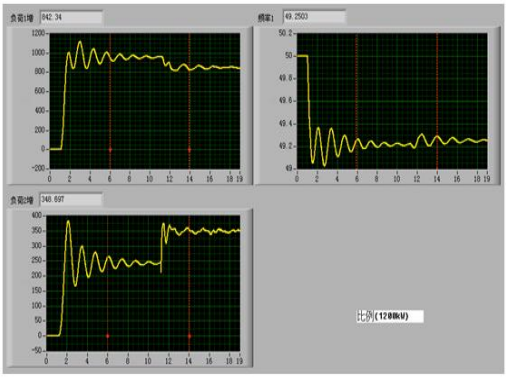


Fig. 6. Change curve when load sudden on 1200kW

6 Conclusions

This paper reviewed both structure and principle of power plant control System. Subsequent sections designed the Hardware-in-Loop Simulation System what concluded Hardware design and software design. Based on PLC, Hardware parts was designed. and software parts were designed in the LabVIEW. The hardware-in-loop simulations were performed for frequency and power control in ship power plant.

The simulation results show the conventional PID control strategy is basically contented for frequency control. The intent of this paper is to provide a new effective method for development of the frequency and power control, because of its merit such as shortening lead time and saving development funds.

References

[1] Chen, G., Zeng, F.: Modern marine engine. National defense science and technology university press, Changsha (2001)

[2] Ren, W., Steurer, M., Woodruff, S.: Progress and Challenges in Real Time Hardware-in-the Loop Simulations of Integrated Ship Power Systems. In: Universities Power Engineering Conference, Newcastle, pp. 669–673 (2006)

[3] Brendecke, T., Kucukay, F.: Virtual real-time environment for automatic-transmission control units in the form of hardware-in-the-loop. International Journal of Vehicle Design (28), 84–102 (2002)

[4] Gawthrop, P.J., Virden, D.W., Neild, S.A., et al.: Emulator-based control for actuator-based hardware-in-the-loop testing. Control Engineering Practice 16, 897–908 (2008)

[5] Bullock, D., Johnson, B., Wells, R.B., et al.: Hardware-in-the-loop simulation. Transportation Research Part C 12, 73–89 (2004)

[6] Ayasun, S., Fischl, R., Vallieu, S., et al.: Modeling and stability analysis of a simulation–stimulation interface for hardware-in-the-loop applications. Simulation Modeling Practice and Theory 15, 734–746 (2007)

[7] Li, S., Wang, J., Li, K.: XPC Technique Based Hardware-In-the-Loop Simulator for Driver Assistance Systems. China Mechanical Engineering 18(16), 2012–2015 (2007)
[8] Chiu Feng, L., Chyuan Yow, T., Tsai Wen, T.: A hardware-in-the-loop dynamics simulator for motorcycle rapid controller prototyping. Control Engineering Practice 14, 1467–1476 (2006)
[9] Li, C., Zhang, F., Huang, Y.: Hardware-in the-loop Engine Simulation for Electronic Control Units Based on the DsPACE System. Acta Armamentarii 25(4), 402–406 (2004)
[10] Fang, X., Yang, M., Peng, X.: Design and implementation of hardware-in-the-loop simulation system for intelligent vehicle. Journal of Huazhong University of Science and Technology (Nature Science Edition) 36(s1), 258–261 (2008)

Control of Sludge Recycle Flow in Wastewater Treatment Plants Using Fuzzy Logic Controller

Wangyani

Guangzhou Institute of Railway Technology, Guangzhou, China
ch_j_11994@163.com

Abstract. Sludge recycling system is an important part of wastewater treatment plants, because the lack of control model, almost all of the sludge return system with wastewater treatment plants is simply the ratio by the amount of water to set the sludge return rate. In order to ensure the water quality with a high ratio, resulting in high energy consumption, handling capacity decreased. This paper presents the application of a fuzzy logic controller to sludge recycle flow processes to predict the amount of return sludge, adjustment return sludge dynamic real-time, saving energy costs .The Simulation results have shown the robustness of the implementation and the efficiency of the proposed control strategies, which have been operated automatically in a safe, stable and optimum operating point, improving effluent quality and reducing energy costs.

Keywords: sludge recycle, fuzzy logic controller, wastewater treatment plants.

1 Introduction

In recent years, the increase in environment restrictions has led to an increase in efforts aimed at attaining higher effluent quality form wastewater treatment plants(WWTP).Activated sludge process is the most common method of wastewater treatment. However, the control of activated sludge processes using classical control tools is often impossible because of the existence of strong nonlinearities, time-variant parameters and complex multivariate. In order to meet these demands, the use of advanced monitoring and modeling methods is required. Sludge recycling system is an important part of WWTP, which can ensure the required reactor sludge concentration, the maintenance of secondary sedimentation tank and the reactor, the dynamic balance between the amount of sludge, Its wastewater treatment plant effluent quality, system stability operations and Operating costs have a major impact. At present, the control of sludge return flow usually two kinds of strategies. [1] The first strategy is to keep sludge recycle flow unchanged ; the second strategy is to have recycle flow increasing in the proportion of the system's influent flow .Traditional control method of sludge return are extensive mode of management, resulting in return sludge concentration range, and low concentrations, there is wasted energy.[2] This paper presents the application of a fuzzy logic controller to sludge recycle flow processes to predict the amount of return sludge. Through real-time sampling and calculation, we have obtained the variation both in the inflow and outflow of all structures and the suspended solids concentration in each layer of secondary clarifier. Reasonable control of return sludge.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 313–318.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Process Description

2.1 The Pilot Plant

The pilot unit configuration is given in figure 1.When these wastewaters enter the plant they are driven to an aeration tank where they mixed with a sludge made up of bacteria. After they have “eaten” most of the organic matter, water and sludge are driven to a settling tank where they are separated. The sludge flows downwards while the water stays at the top of the tank and flows over an overflow weir to be released to nearby rivers. The sludge is withdrawn from the tank and then split into two streams: one is driven to the aeration tank to keep the sludge concentration at an appropriate level while the other is eliminated. [3]

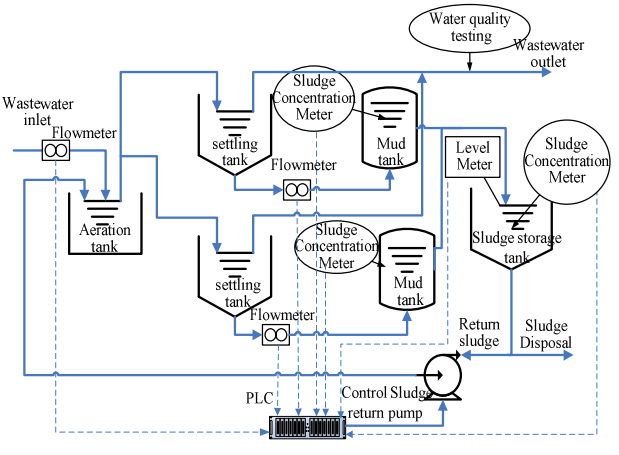


Fig. 1. Diagram of a Wastewater Treatment Plant

2.2 Process Control

Aeration tank of sludge from the secondary settling tank return The sludge return system, biological reactor and secondary settling tank linked to a balanced system . Biological reaction tank MLSS Concentration (MLSS) and the return sludge concentration (RSS) relationship, see equation (1). According to the system dealing with water, return sludge concentration, the concentration of control over biological and chemical mixture pool target calculations obtained the amount of return sludge, see equation (2).

$$MLSS = \frac{R}{1+R} RSS \tag{1}$$

$$Q_R = \frac{Q_P \times MLSS_0}{RSS - MLSS_0} \tag{2}$$

In this model the input variables are the "Influent Flowrate" Q_P , the "Sludge Concentration" in the aeration tank $MLSS_0$ and The "return sludge concentration" in the sludge storage tank, RSS , is assumed to be a process parameter that will be provided at every time step. The output variables is the "return sludge flow" Q_R .It defines the sludge concentration in the settler, a portion of which will be recalculated to the aeration tank to keep the sludge concentration at an appropriate level. This recirculation is defined by a flowrate, Q_R , which is adjusted by the technician supervising the plan dynamics. It is represented by a parameter, R, which provides the ratio between this "Recirculation Flowrate" and the "Influent Flowrate". [4] The control of this parameter is the aim of this work. The methods for calculating the set-point for the sludge flow, Q_R , Control variable is the system influent flowrate By the above formula.

3 Design of a Fuzzy Logic Controller for the Return Sludge

3.1 Architecture of the Fuzzy Logic Controller

A fuzzy logic controller for a sludge recycle flow processes, as figure 2 depicts, comprises of the following components :fuzzification module, inference engine , knowledge base, and defuzzification module. A fuzzy logic controller is operated by repeating a sequence of the steps. First, crisp physical measurements are taken from process variables that represent relevant states of the controlled system.. Next, these crisp values are converted by the fuzzification module into linguistic terms, which are specified by the appropriate membership functions of the fuzzy sets. The fuzzified measurements are then matched to the ruleantecedent by the inference engine to assess the adequacy of the control rules stored in the knowledge base . The corresponding output of this evaluation is still a fuzzy set (or several fuzzy sets) defined on the universe of possible control actions. [5] Finally , the so-called defuzzification module converts. Finally ,the so-called defuzzification module converts.

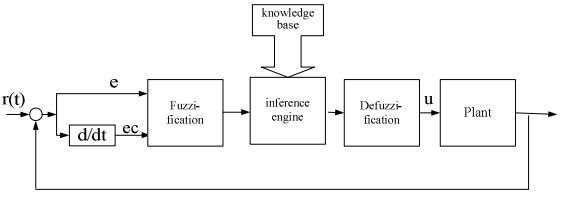


Fig. 2. Architecture of the fuzzy logic controller

3.2 Design Procedure of a Fuzzy Logic Controller for the Sludge Recycle Flow Process

Designing a fuzzy logic controller requires specifying all of the fuzzy sets and defining their membership functions for each input and output variable, compiling a

complete set of control rules that operate on these fuzzy sets, and determining the method to select a crisp output action based on the fuzzy result generated by the control rules. [6]

Step 1. Identify the and output variables (rule base). An sludge recycle flow system comprises primarily of two units: (1) a biological reactor with anaerobic/oxic zones (2) a setting tank and (3) a sludge storage tank. Process performance of a sludge recycle flow system, Influent Flowrate and return the sludge concentration are two key factors. Accordingly, two input variables, Q_P an and $MLSS_0$, are determined to be the fuzzy logic controller's inference result is then transmitted to be the fuzzy logic controller's input variables. The controller's inference result is then transmitted to effectively control the output variable: sludge recycle rate[7].

Step 2. select linguistic state for each variable (data base). After identifying the fuzzy logic controller's relevant input , the designer must select meaningful linguistic states for each variable. In this study, we selected eight linguistic states for each of the three variables, as figure 3:

PB-positive big PS-positive small PM-positive medium
NS-negative small NB-negative big NM-negative medium
NZ-negative zero PZ-positive zero

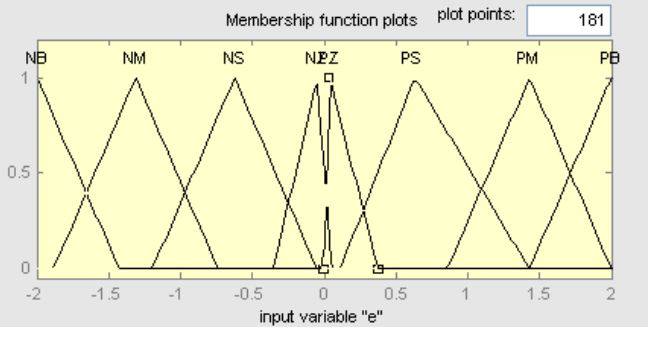


Fig. 3. Fuzzy sets of variable

Step3. construct control rules (rule base). In this step, the knowledge pertaining to the control system is formulated in terns of a set of fuzzy rules. With input variables and output variable, the inference rules have the following form similar to the one below:

If MLSS set error is NB and change in error is NB then output is NB . [8]

Since each input variable in this study has eight linguistic states, the total number of possible nonconflicting fuzzy inference rules is 8^2 =64. These rules can be conveniently represented in three-dimensional surface, as Figure. (4) depicts.

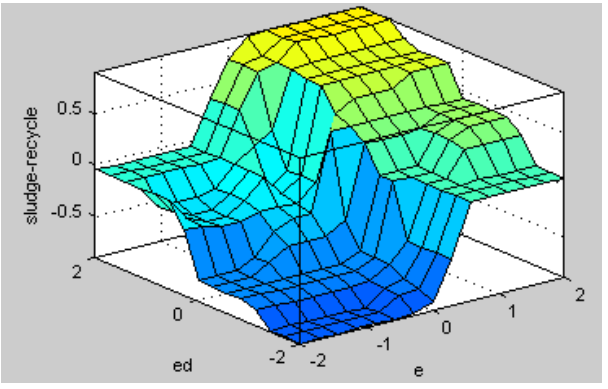


Fig. 4. Fuzzy rules three-dimensional surface

Step 4. select defuzzification strategy. In the last step of the design procedure, the designer of fuzzy controller must select an appreciate defuzzification method. Defuzzification attempts to convert cach conclusion in terms of a fuzzy set, to a crisp output value. [9] The crisp output value u is the abscissa under the centre of gravity of the fuzzy set. The precise control value u is calculated by the following formula:

$$u = \frac{\sum_i \mu(x_i) x_i}{\sum_i \mu(x_i)} \tag{3}$$

4 Simulation and Discussion

We use MATLAB 7.6 to do the experiment. By adjusting the sludge recycle flow rate, the sludge concentration was maintained at the set value, set the value determined under dynamic operating conditions, The simulation result was shown by Figure 5. The results show that the fuzzy control method can effectively stabilize the biochemical pool of sludge concentration. The control strategies have the robustness of the implementation and the efficiency of the proposed, which have been operated automatically in a safe, stable and optimum operating point, improving effluent quality and reducing energy costs. [10]

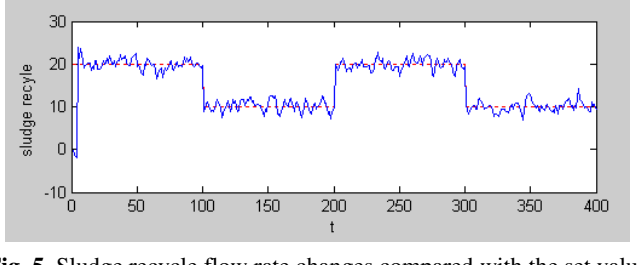


Fig. 5. Sludge recycle flow rate changes compared with the set value

5 Conclusion

The present study concludes that the application of fuzzy logic controller can improve Sludge recycling system, providing low effluent discharges with minimum operational costs. The simulation and experiments of the proposed control strategies for sludge recycling and wastage and robust operation of the biological process improving the effluent quality, saving energy costs and reducing hydraulic disturbance of the final settler. In conclusion, a combination of these strategies may more effectively reduce effluent variability and provide better sludge quality during normal plant operation. Also, it is believed that the results obtained can be used in order to improve or develop an actual wastewater treatment plant. [7]

Acknowledgment. This work is get a helping hand from 'Fund of Innovation Creation Academy Group' established by the Guangzhou Education Bureau. Key Project of Guangzhou Scientific Program of China under Grant #2010Z1-E301.

References

[1] Ma, Y., Peng, Y., Wang, S.: New automatic control strategies for sludge recycling and wastage for the optimum operation of predenitrification processes. Journal of Chemical Technology and Biotechnology 81, 41–47 (2006)
[2] Henze, M., Grady, C.P.L., Gujer, W., Marais, G.V.R., Matsuo, T.: Activated Sludge Model No.1, IAWQ Scientific and Technical Report No. 1. IAWQ, London (1987)
[3] Banadda, E., Smets, I., Jenné, R., Van Impe, J.: Dynamic modeling of filamentous bulking in lab-scale activated sludge processes. In: Proceedings of the 7th International Conference on Dynamics and Control of Process Systems, Dycops 7, CDROM, Cambridge, MA (USA), p. 6 (2004)
[4] Koumboulis, F.N., Kouvakas, N.D., Tzamtzi, M.P., Stathaki, A.: Metaheuristic control of substrate concentration for an activated sludge process. In: Proceedings of the International Conference on Modelling, Identification and Control (ICMIC 2008), Shanghai, China (2008)
[5] Campello, R.J.G.B., Von Zuben, F.J., Amaral, W.C.C., Meleiro, L.A., Maciel, F.R.: Hierarchical fuzzy models within the framework of orthonormal basis functions and their application to bioprocess control. Chemical Engineering Science 58(18), 4259–4270 (2003)
[6] Lech, R.F., Leslie Grady Jr., C.P., Lim, H.C., Koppel, L.B.: Automatic Control of the Activated sludge Process-II. Efficacy of control strategies. Water Res. 12, 91–99 (1978)
[7] Olsson, G., Newell, B.: Wastewater Treatment System. IWA (1999)
[8] Zadeh, L.A.: Fuzzy Logic, Neural Networks and Soft Computing, One page course description of CS 294-4. The University of California at Berkley (1992)
[9] Tomiello, M., Perrin, E., Roubens, M., Crine, M.: Fuzzy control of an activated sludge process. In: Second European Congress of Chemical Engineering –Montpellier (1999)
[10] Tong, R.M., Beck, M.B., Latten, A.: Fuzzy Control of the Activated Sludge Wastewater Treatment Process. Automatica 16, 695–701 (1980)

Research and Application of Embedded Real-Time Operating System

Zhangbo

ChongQing Industry Polytechnic College, ChongQing 401120, China

Abstract. In this paper, based on the analysis of existing embedded real-time operating system, the architecture of an operating system is designed and implemented. The experimental results show that the design fully complies with the requirements of embedded real-time operating system, can achieve the purposes of reducing the complexity of embedded software design and improving the maintainability, reliability, flexibility. Therefore, this design program has high practical value.

Keywords: embedded system, real-time operating system, EDF, SJF.

1 Introduction

With the information technology development, more and more embedded real-time systems are applied to social, economic, educational and military fields. For these systems, real-timing is one of the most pivotal performance indicates. Embedded real-time systems often have very high reliability requirements, such as industrial control, aerospace, bio-medical and everyday vehicle control systems, which needs to be guaranteed in many ways. The task real-time scheduling in time and space, QOS quality control and multi-task management often tend to increase the complexity of system design[1], making the design process of such a system needs to experience a long time, lots of manpower and material resources to achieve the final design requirements. And the system such designed often is based on proprietary hardware and software, making the whole system upgrade, migration, and maintainability are greatly affected.

The emergence of embedded real-time operating system (ERTOS) provides a powerful tool for embedded real-time system design. The initial embedded real-time system does not need ERTOS support. At that time, due to memory size and processor capacity constraints, program design is often achieved with the assembly code, but the concurrent implementation of multiple tasks is entirely ensured by the manual arrangements of the staff. With the development of chip manufacturing technology and programming languages, many high-level language-based programming technologies and ERTOS gradually emerged in. A lot of practical experiences have proved that ERTOS can greatly simplify the system design process, save human and material resources invested, and achieve a real-time concurrent multi-task operation, thus has a high application value.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 319–325.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Embedded Real-Time Operating System Design

Operating system architecture is the key to ensure the system real-timing, reliability, flexibility, portability and scalability. At present, the architecture of the operating system can generally be divided into: simple structure, single-core structure, hierarchy structure, micro-kernel and the outer core structure. These structures have their own advantages and disadvantages, application background[5], so it should select the appropriate architecture based on the specific application scope and requirements. The ERTOS this paper designed is an embedded real-time operating system focusing on reliability and flexibility. So it uses partial order hierarchy kernel. In addition it uses modular design from the aspect of software engineering idea. This has formed a compromise program with hierarchy structure as the core, while slightly tilting to the monomer structure, the architecture as shown in figure 1.

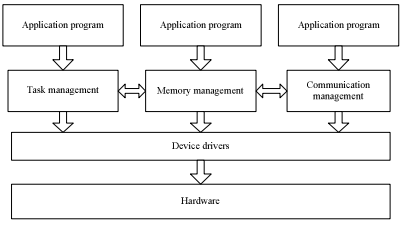


Fig. 1. Embedded real-time operating system architecture

Through the use of hierarchical structure, the organizations and dependencies between modules are clear. The upper function is built based on the lower functions, and then system's readability and adaptability are enhanced. In addition, by the isolation between the layers, the user space, kernel space and the hardware abstraction layer are separated, thus the system is more secure and reliable. Furthermore, to modify or replace for a certain layer only affects the neighboring two layers at most, thus is easy to modify and expand. The disadvantage is that which layer the function modules of the operating system should be placed, how to effectively stratify must be considered, in particular, the operating system using this architecture style must also consider some special layered principles[2]. In order to enhance its adaptability, the software closely related to the machine characteristics must be placed on the bottom. Secondly, the most common operating mode should be put in the most inner layer, and the part changing with these operations on the outer. In addition, the current operating system design is based on the concept of the process, so the system call module to provide services for the process is usually put on the system inner layer, and many other factors are necessary to consider.

3 Embedded Real-Time Operating System Implementation

3.1 Task Management

Task management is the core of real-time operating system, which is to mainly provide mission life cycle management, state control and real-time scheduling capabilities. Mission life cycle management task is to achieve the task creation, task deletion, task

deletion lock, task restart and other functions. Task state control is to provide switching between multiple finite states, which is achieved mainly through the finite state machine. The task scheduling is mainly to coordinate the fight use of computer system resources. Process management, also known as CPU scheduling, the fundamental task is to allocate CPU resources for the process in a ready state according to a certain principle. As the embedded system memory and I/O devices are generally attributable to a particular process with CPU, so the concept of task management and process management are similar, without distinction on many occasions.

3.1.1 Mission Control Mechanism

Task is the basic unit in the system to access resource and carry out scheduling. A program to run must first create a task. Then the system will allocate stack space for it, including the user stack and system stack. If its priority is higher than the running task or there is no currently running task, it will also be allocated processors to put into operation. When the program is finished[6], a deletion operation on the task should be implemented to release all the resources applied in the execution process. From a structural point of view, real-time task is composed by task code, task data, stack space and TCB four parts, the structure as shown in figure 2.

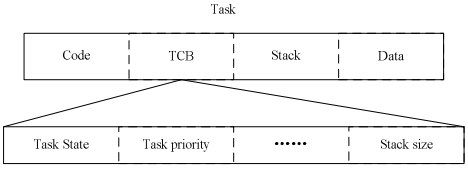


Fig. 2. Task composition structure

The definition and conversion of the task state is the basis of multi-task scheduling. In multitasking operating systems, the task must be in the system with a specific state. On different systems, task state definitions vary, a typical task transforming relationship as shown in figure 3. With the running of the real-time embedded operating system, each task will be transformed from one state into another state according to the relationship of the task state transformation. Clear and detailed division of task states, a clear distinction of the system calls having an impact on the callers themselves and the calls having an impact on other tasks, on the one hand can make the task transformation more clear, the system calls easier to understand, and on the other hand, in many systems the blocking state overlapping is easy to make the state management complex, difficult to understand.

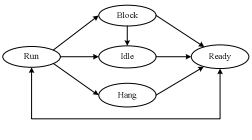


Fig. 3. Task state transformation relationship

3.1.2 Task Scheduling

For multi-task real-time kernel, scheduling algorithm has a great impact on response time. The existing embedded real-time operating systems commonly use priority-based scheduling algorithm, the system always chooses the highest priority current task for execution[3]. However, with in-depth development of real-time system applications, many types of real-time constraint tasks are concurrent, and the task dynamical join and exit in changing environment, such hybrid real-time system demands also increased, which also makes the embedded systems have increasing demands for the task scheduling, the original method only according to fixed-priority to account for scheduling has been difficult to meet the needs of current embedded real-time operating system. So how to design an efficient, reasonable scheduling algorithm to make the system can quickly complete the task scheduling and switching, reduce the load impact on the performance, is a very important issue. In this paper, the scheduling strategy combining the EDF with SJF algorithms to achieve the task scheduling is adopted, and the scheduling algorithm as shown in figure 4.

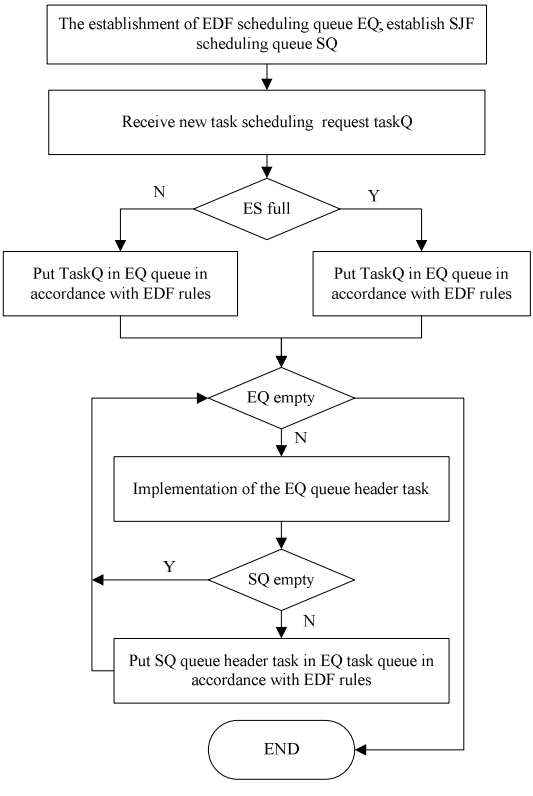


Fig. 4. Task scheduling process

3.2 Memory Management

In embedded systems, physical memory is the most basic and important resource in addition to CPU. Physical memory space distribution and collection efficiency in the system plays an important role in the decision of the system efficiency. In addition, reasonable memory allocation also is the key to ensure the system reliability, avoid memory leaks and conflict[4]. Therefore, how to efficiently allocate and reclaim memory space, how to reduce the loss of resources memory fragments caused, and how to implement memory management security, all these problems must be considered in ERTOS memory management module design and implementation.

In ERTOS, memory usage should follow the principle which is to best use static memory allocation, such as variable, array advance statements, because the use of static memory allocation can reduce the memory operation and memory fragments, to increase system speed. The so-called memory fragments refer to that the memory has some non-continuous free memory blocks. As these free memory blocks are too small, so in each subsequent memory application, they can not be used. The presence of memory fragments not only makes the available memory space reduced, but also will increase the computational burden of memory management unit, thus reduce the system's real-timing. In order to ensure the real-timing, ERTOS generally has not function of clearing the memory fragments,

because that ERTOS can not find a suitable time to pause a running program to clean up the memory. However, the dynamic application of application to the memory is inevitable[7]. Therefore, the performance of memory allocation algorithms directly affects the ERTOS real-timing and stability.

Taking into account the relatively small memory of embedded systems can reduce the cost of first-fit algorithm finding and has capacity of retaining large space partition, ERTOS uses first-fit algorithm to allocate memory[8]. At the same time the modular design also makes that when facing different demands it can use better and more suitable allocation algorithm or improve the original algorithm. ERTOS memory allocation algorithm flowchart is shown as in figure 5.

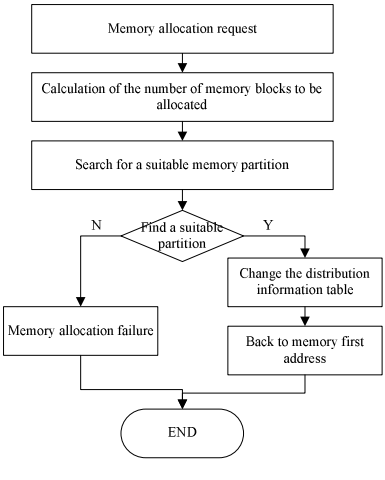


Fig. 5. Memory allocation algorithm processing

4 Experimental Results and Analysis

To ensure that ERTOS has broad versatility and portability, here ARM920T-based ARM platform is used for development and testing, the experimental results as shown in table 1:

Table 1. Experimental data

Test scale	Test type	Response time (μs)	Conclusion
5 tasks	Memory allocation	6.37	Normal
	Task switching	10..67	Normal
	Task interruption	9.43	Normal
8 tasks	Memory allocation	6.98	Normal
	Task switching	11.34	Normal
	Task interruption	9.78	Normal
10 tasks	Memory allocation	7.43	Normal
	Task switching	12.31	Normal
	Task interruption	9.97	Normal

The above experimental data show that the program fully adapts to the open embedded system requirements, task switching time and interruption response time is shorter, which can better reflect its real-timing performance, thus fully meets the requirements of embedded real-time operating system.

5 Conclusion

As embedded systems become increasingly complex and large, the use of ERTOS in embedded devices has become an inevitable trend. Using ERTOS can decompose the complex tasks into multi-tasking system to simplify the system design, and then to improve the speed of product development and market. Meanwhile, the use of ERTOS also has a great help for improving the system stability and reliability. In this paper, based on the analysis of existing embedded real-time operating system, the architecture of an operating system is designed and implemented. The experimental results show that this design can fully meet the requirements of embedded real-time operating system, thus has higher practical value.

References

[1] Li, Q., Yao, C.: Real-Time Concepts for Embedded Systems. CMP Books, Lawrence (2004)
[2] Rivas, M.A., Harbour, M.G.A.: Ada RTS POSIX-Extending Ada's Real-Time System Annex with the POSIX Scheduling Services. Ada Letters 13(1), 15–18 (2001)
[3] Stallings, W.: Operating systems: Internals and design principles, 4th edn., pp. 46–68. Prentice Hall, America (2004)
[4] Bovet, D.P., Cesati, M.: Understanding the Linux Kernel, 2nd edn., pp. 53–75. O'Reilly (2002)
[5] Biroli, G., Monasson, R., Weigt, M.: A variational description of the ground state structure in random satisfiability problems. Journal of Engineering Computer Structure (11), 1–12 (2000)
[6] Cervin, A.: Merging Real Time and Control Theory for Improving the Performance of Embedded Control Systems. Department of Computer Engineering and System Science, vol. 9(15), pp. 455–460. University of Pavia, Italy (2004)
[7] Colin, A., Puaut, I.: Worst-Case Execution Time Analysis of the RTEMS Real-Time Operating System. In: 13th Euromicro Conference on Real-Time Systems Delft, pp. 191–198. IEEE Computer Society, Netherlands (2001)
[8] Silberschatz, A., Galvin, P.B., Gagne, G.: Operating system concepts, 7th edn. John Wiley& Sons, Inc., Hoboken (2004)

A PCA-Based Automated Method for Determination of Human Body Orientation

Weihe Wu[1,2], Aimin Hao[1], Wentao Wang[2], and Yiqiang Wang[2]

[1] State Key Laboratory of Virtual Reality Technology and System, Beihang University Beijing, China
[2] School of Art & Design, Beijing University of Technology, Beijing, China
wutongtong@hotmail.com

Abstract. Human body orientation is a fundamental preprocessing task in surface registration, motion analysis, and data driven animation for human body. Previous human body orientation determination methods suffer from pose dependence and non-robustness. In this study, we propose an automated pose-independent approach to human body orientation by combining geometric characteristics of human body based on principal component analysis (PCA). We first analyze the relations between the bones of lower limbs and the body orientation, and obtain a feature vector consisting of the angles between the bones and the body orientation. We then use PCA to analyze the training samples of the feature vector and construct a classifier based on the first principal component to determine the human body orientation. Our experimental results show that the linear separability of the feature vector is perfect, and the classifier trained by PCA can be used for orienting human body efficiently in a pose-independent manner. This method can also be easily integrated into existing approaches that extract skeleton from human body shape.

Keywords: human body orientation, pose-independent, skeleton extraction, PCA.

1 Introduction

The 3D model based on surface meshes has become a popular geometry representation of the human body, and widely used in the fields of character animation, motion analysis, and shape recognition. The model, however, contains little semantic information about human body. Therefore, the process and analysis for extracting semantic information have recently drawn the interest of the shape analysis community. Numerous methods have been developed for human body surface mesh segmentation or skeleton extraction, for instance, mesh contraction [1], shape decomposition [2], Reeb graph [3-4], distance transformation [5], medial axis transform [6], clustering technique [7-9]. However, these methods are not able to determine the orientation of human body and recognize limbs after segmentation or skeleton extraction, thus limiting their further applications, such as to human body recognition and animation.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 327–337.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

The orientation of human body can be determined based on the tip of feet or the centroid of human body in a standard pose. Hu et al. [10] propose an orientation determination method for matching the virtual human model and motion data based on the observation that the centroid is in a slight forward direction. Although the observation is valid for some specific poses, the centroid usually varies with the human body pose and is not always in a forward direction. Therefore, the observation may not be applied to human body orientation in arbitrary postures. Yu et al. [11] orient the human body based on the concavities of waist contours of human body near the spine. Yet, the concavities do not always exist for human body models. In fact, it is difficult to orient the human body in an arbitrary posture using a traditional approach that has only one parameter of angle or position. For example, the orientation of human body in a kneeling posture can not be determined by the tips of feet. Many researchers often have to confine the human body to some specific poses in their studies. For instance, Baran et al. [12] require that the character must be given in approximately the same orientation and posture as the given skeleton in the automatic rigging method, which is a big limitation to its application. To solve this problem, we present a method of orienting human body in arbitrary postures using the relation between the body orientation and the bones of lower limbs. We first construct a feature vector consisting of angles between each bone of the lower limbs and the body orientation, and then train a classifier of human body orientation by analyzing the feature vector using PCA. Finally, we determine the orientation of human body using the classifier after extracting the skeletons from models by existing methods.

The rest of the paper is organized as follows: in section 2, we present the approach to extract skeleton, and then define the human body orientation and the feature vector. In section 3 and 4, we explain the algorithmic details for orienting human body using PCA. In section 5, we show the experimental results. Finally, we conclude the paper and briefly discuss the future research in section 6.

2 Feature Vector Construction

2.1 Human Body Model

Fig. 1(a) shows a surface model of human body M, which is called a human body model for short in this paper. In the viewpoint of topology, M can be expressed with a two-tuple $< K, V >$, where K is a simplicial complex, representing the connection relations among vertices, edges and faces, and V is the set of vertices on M, $V = \{v_i \in R^3 \,|\, i = 1, 2, \cdots, |V|\}$, and $|V|$ is the number of vertices. Therefore, a human body model M corresponds with a point in $R^{3\times|V|}$.

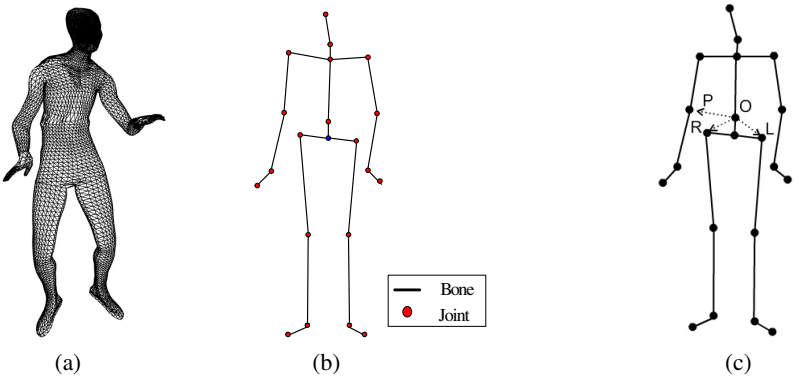


Fig. 1. (a) Human body surface model; (b) Human body skeleton model; (c) Orientation of human body

The skeleton model of human body S, shown in Fig. 1(b), consists of bones and joints. Let J be the set of joints, which contains the top point of head and the end points of limbs in this paper, and $|J|$ denotes the number of joints, $|J| = 21$ in Fig. 1(b). Extracting skeleton from human body model can be taken as a map D: $\{M\} \rightarrow \{S\}$, and human body models with different numbers of vertices can be mapped to skeleton models with the same number of joints. A skeleton model also corresponds to a point in $R^{3\times|J|}$, where $|V| >> |J|$. Owing to the skeleton model containing information about the posture and orientation of human body, we can construct the feature vector based on the skeleton model to orient the human body.

2.2 Extraction Skeleton

Conventional wireless network security model is based on static and open loop control system. This method can't effective response to dynamic network security thread and the low robust of wireless network system. According to high development of wireless network technology, the closed loop control security system is becoming more and more popular for its high dynamic response; its working flow diagram is showed in figure 1.

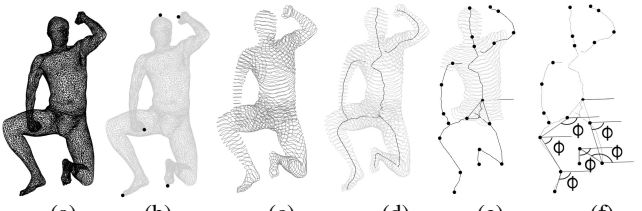


Fig. 2. Computing feature vector. (a) Human body model; (b) Feature points; (c)Isolines; (d) Center Lines; (e) Joint points; (f) Feature vectors.

Before orienting the human body, we have to extract the skeleton from the human body model. A lot of approaches for skeleton extraction have been proposed, and we employ the method similar to [13] which has good generalization performance. First, five feature points at the end points of head and limbs on human body surface are automatically extracted using geodesic distances along the human body surface. Second, the feature point located at the top head can be recognized according to the symmetry of the human body, and the feature points located at the hands or feet can be identified through the fact that the geodesic distance from the top head to one hand is less than the geodesic distance from the top head to one foot, shown in Fig. 2(b). Third, taking the feature points as source points, the geodesic distance isolines on the surface are calculated and the central lines are subsequently constructed, shown in Fig. 2(c) and 2(d). Along the center lines, the joint positions are found relying on the proportions provided by the geodesic model, shown in Fig. 2(e).

2.3 Definition of Human Body Orientation

Human body orientation is the direction that the torso of human body faces to. We take the joints of waist, left and right hip near the pelvis to define the human body orientation because there are no relative motions among these joints.

Definition 1. (Human body orientation) Shown in Fig.1(c), let O, R, L be the joints at the waist, right hip and left hip respectively, $\overrightarrow{OP} = \overrightarrow{OR} \times \overrightarrow{OL} \Big/ \left| \overrightarrow{OR} \times \overrightarrow{OL} \right|$, then the direction of $\overrightarrow{OP}$ is taken as the human body orientation.

$\overrightarrow{OP}$ is the normal vector to the plane determined by the joints at waist and hips, meanwhile $\overrightarrow{OR}$, $\overrightarrow{OL}$ and $\overrightarrow{OP}$ satisfy the right hand rule. After extracting the skeleton from a human body, we usually can not distinguish between the right and the left hip, but we can determine the plane by the joints at waist and hips so that we can obtain two directions perpendicular to the plane. Orienting human body is to judge which one of the two directions is in accordance with the human body orientation. Therefore, orienting human body is converted into a two-class pattern recognition problem, where one class is the direction that the human body faces to, and the other is the direction that the human body opposes to.

2.4 Feature Vector

Although there are some relations between the head orientation and the human body orientation, it is not easy to extract the discriminant feature from the head for orientation. The bone positions of the upper limbs vary drastically with poses, and are symmetric with respect to the plane perpendicular to the human body orientation especially in some poses. Therefore, it is difficult to obtain information about human body orientation from the upper limbs. However, the lower limbs correlate highly with the human body orientation. We take the angles between the human body

orientation and each bone running from hip to knee, from knee to ankle, from ankle to toe, respectively, as the geometry feature to judge the human body orientation. The feature vector $\theta = (\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6)$, $\varphi_i \in [0, \pi]$, is shown in table 1 and Fig. 2(f).

The i^{th} geometry feature function $F_i : \{S\} \to [0, \pi]$, stands for the mapping from the skeleton S to feature ϕ_i, where $i \in \{1, \cdots, 6\}$. Meanwhile, a composite function can be defined: $G_i = F_i \circ D : \{M\} \to [0, \pi]$, which stands for the mapping from the human body to feature φ_i, namely $\varphi_i = G_i(M)$.

Table 1. Feature vector

No.	Define	Sign	Meaning
(1)		ϕ_1	Bend for right thigh
(2)		ϕ_2	Bend for right leg
(3)		ϕ_3	Bend for right foot
(4)		ϕ_4	Bend for left thigh
(5)		ϕ_5	Bend for left leg
(6)		ϕ_6	Bend for left foot

3 Feature Vector Analysis

3.1 Feature Functions Curves

Taking the human body models from the SCAPE model set (with 71 different poses) as samples, we calculate the angles between each bone of lower limbs and the direction that the human body faces to and the angles between each bone of lower limbs and the direction that the human body opposes to for all these models, $G_1\{M\}, G_2\{M\}, \ldots, G_6\{M\}$, and the curves of these feature functions are shown in Fig. 3 (the horizontal axis is for human body models; the vertical axis is for angles).

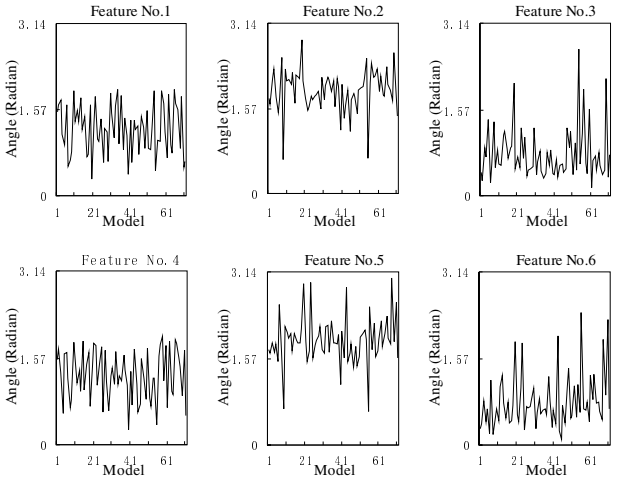


Fig. 3. Curves of feature functions

The above figure illustrates that the values of the feature function vary with poses. To an arbitrary bone of the lower extremities, the sum of the angle between the bone and the human body orientation and the angle between the bone and the opposed human body orientation is π . Therefore, if a feature function curve locates completely above or below $\pi/2$, we can judge the human body orientation only by the feature function. However, Fig.3 shows that none of feature function curves locate completely above or below $\pi/2$. Hence, we take the feature vector containing the six feature functions as the geometry feature to judge the human body orientation.

3.2 Analyzing Feature Vector by PCA

Each of the selected six features contributes to the determination of human body orientation differently, and moreover, there is some correlation among the features, for example, the left leg usually bends when the right leg bends. Therefore, we extract the principal components $\vartheta=(\phi_1,\phi_2,\phi_3,\phi_4,\phi_5,\phi_6)$ from the feature vector $\theta=(\varphi_1,\varphi_2,\varphi_3,\varphi_4,\varphi_5,\varphi_6)$ through PCA, and then construct the discriminant function with the principal components which are highly correlated with the human body orientation.

Taking the feature vector θ as a random vector, we construct the positive sample using the angle between the bone at lower limbs and the direction that the human body faces to and negative sample using the angle between the bone at lower limbs and the direction that the human body opposes to. Because the sum of each pair of angles from the positive and the negative samples is π, the mean of φ_i is $\pi/2$.

$$\hat{\varphi}_i = \frac{1}{T}\sum_{j=1}^{T}\varphi_{ij} = \pi/2 \quad i = 1,\cdots,6$$

Where T is the number of samples. Let the mean of θ be $\hat{\theta} = (\hat{\varphi}_1, \hat{\varphi}_2, \hat{\varphi}_3, \hat{\varphi}_4, \hat{\varphi}_5, \hat{\varphi}_6)$. The process of the principal component analysis includes the following five steps:

Constructing the mean centered feature vector $\overline{\theta} = \theta - \hat{\theta}$

Generating the covariance matrix of samples: $C = (c_{ij})_{6\times 6}$,where $c_{ij} = \frac{1}{T}\sum_{k=1}^{T}\overline{\theta}_{ik}\cdot\overline{\theta}_{jk}$

Calculating the eigenvalues of the covariance matrix C , $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l > 0$, and $U_1, U_2, \cdots U_l$ for their corresponding eigenvectors, where $l = Rank(C) \leq 6$.

Projecting $\overline{\theta}$ onto the eigenvectors to obtain the principal components ϕ_i .

Constructing the principal components feature vector: $\vartheta = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6)$, the components are sorted from high to low according to their corresponding eigenvalues. The component associated with the largest eigenvalue is one that reflects the greatest variance in the feature vector.

We calculate the principal components of the feature vectors for the SCAPE model set. Fig. 4 shows that ϕ_1 correlates highly with the human body orientation, while the rest principal components have no contribution to judging the human body orientation, which are considered as noise. Therefore, the human body orientation can be judged only by $\phi_1 = < U_1, \theta - \overline{\theta} >$.

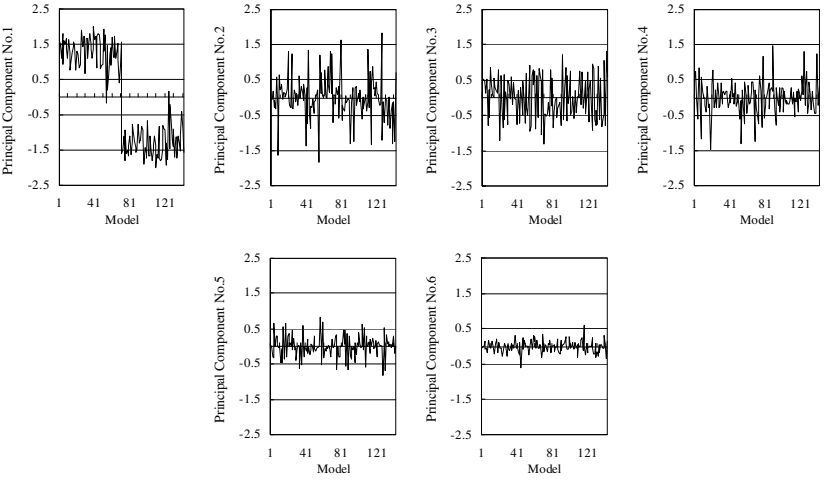


Fig. 4. Curves of principal components of SCAPE

4 Orienting Human Body

Let $< \theta_i, y_i >$ be a sample in the training set, where θ_i represents the feature vector of the human body and y_i represents the class of sample. If $y_i = 1$, θ_i consists of the angles between each bone at lower limbs and the human body orientation and the sample is positive; if $y_i = -1$, θ_i consists of the angles between each bone at lower limbs and the direction that human body opposes to and the sample is negative.

During the training step, we first extract the joints from each human body M_i using the method introduced in section 2, and manually denote the left and right hips to calculate the human body orientation. For each human body model, a pair of samples including a positive one and a negative one is constructed. Then, the mean feature vector $\hat{\theta}$ and the first eigenvector $U_1 = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)$ are obtained from the training samples by PCA, where each component of $\hat{\theta}$ is $\pi / 2$ because the positive and negative samples are paired. Finally, the discriminant function is constructed by combining the six feature functions based on the first eigenvector:

$$g(\theta) = < U_1, \theta - \hat{\theta} > = \sum_{i=1}^{6} \beta_i (\varphi_i - \pi / 2). \; f(\theta) = \operatorname{sgn}\{ g(\theta) \}$$

During the testing step, the joints of the test human body models are extracted using the method introduced in section 2, and then we assign one of two directions which are perpendicular to the plane formed by the connection of joints of waist and hips as the human body orientation. Subsequently, the feature vector of the human body θ is calculated, and then θ is used to compute $f(\theta)$. If $f(\theta) = 1$, then the assigned direction is the human body orientation; otherwise, the assigned direction is the direction that the human body opposes to. Once the human body orientation is determined, the left and right hips are recognized according to the definition of human body orientation in section 2, and the left and right lower limbs are also determined. In the same way, we can judge the left or right upper limbs through the relation between the plane formed by the waist and shoulders and the human body orientation. By this way, we distinguish between the right and the left upper and lower limbs of human body.

5 Implementation Results

To validate the effect of our algorithm, we chose the SCAPE human body models with 71 poses and the DANCE human body models with 201 poses. First, we generated 71 SCAPE sample pairs and 201 DANCE sample pairs, respectively, in which each pair contains one positive and one negative sample, then we took the above sample pairs as the training data and test data to implement cross-over experiment. Fig. 5 shows parts of human body models of SCAPE and DANCE in different poses. All experiments are run on a PC with 3.4GHz CPU and 2G memory,

the programming environment is Windows Visual C++.net 2005. All models in our experiment are made of triangular surfaces, whose faces range from 14000 to 25000, and whose vertices range from 7000 to 12500.

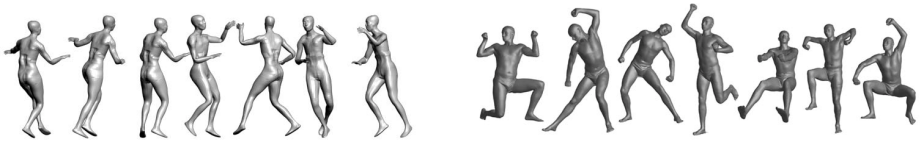
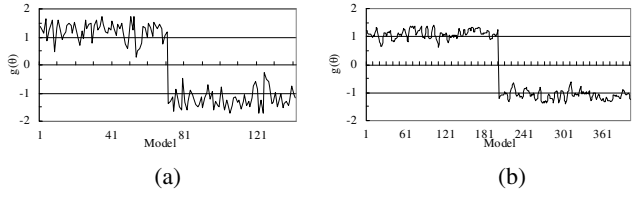


Fig. 5. Human body Models from DANCE (left) and SCAPE (right)

5.1 Classifier Trained with SCAPE

The SCAPE models were taken as training samples and the DANCE models were taken as test samples. Fig. 4 shows the principal components of the training samples, and $\lambda_1 / \sum_{i=1}^{6} \lambda_i = 0.65$. Fig. 6(a) shows that every test sample is correctly judged, and there is $|g(\theta)| > 0.6$ for all test samples.



(a) (b)

Fig. 6. Test results.(a) Classifier trained with SCAPE;(b) Classifier trained with DANCE

5.2 Classifier Trained with DANCE

The DANCE models were taken as training samples and the SCAPE models are taken as test samples. Fig. 7 shows the principal components of the training samples, and $\lambda_1 / \sum_{i=1}^{6} \lambda_i = 0.77$. Fig. 6(b) shows that every test sample is correctly judged. Moreover, there is $|g(\theta)| > 0.5$ for all pairs of samples except one pair of samples, for which there is $|g(\theta)| = 0.26$.

The results of two tests show that the linear separability of the first principal component trained by PCA is perfect (Fig. 7 and Fig. 4). Comparing the results of the two tests, the performance of the classifier trained with SCAPE are slightly better than that trained with DANCE. Although the training set size of SCAPE is less than that of DANCE, the pose in the SCAPE models set varied more obviously so that the SCAPE models are more representative.

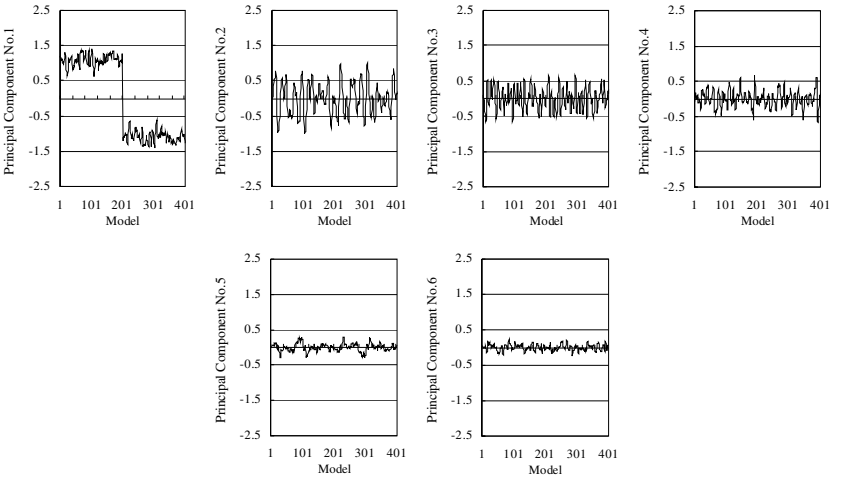


Fig. 7. Curves of principal components of DANCE

Comparing our method to others, ours correctly classified all the models in both training sets, where there were 14 and 12 models in DANCE and SCAPE training set, respectively, were judged incorrectly by the method based on centroid [10]. In addition, since there is no obvious concavity near the waist contours in the above models, we are not able to judge the human body orientation using Yu's method [11].

6 Conclusion

We have presented a general and pose-independent approach to human body orientation. Using this method, the right and the left limbs can be well distinguished, which is the foundation for matching human body to motion data, registering two human body surfaces and retargeting human body motions. Owing to taking the angles between the lower limbs and body orientation as the feature vector, it is very convenient to integrate this method into existing approaches for skeleton extraction. Experiment results show that the linear separability of the feature vector is perfect, the classifier trained by PCA can be used for orienting human body efficiently in a pose-independent manner, and the accuracy of orientation is much better than previous methods. However, due to the limited experimental samples, the human body in some extreme poses may not be correctly oriented by the trained classifier. In future work, we intend to apply this method to human body recognition and automatic animation generation problems.

Acknowledgments. This work is supported by Beijing Municipal Commission of Education of Science and Technology Program (No.KM200910005020) and Program of Training Innovative Talents to Service Beijing (No.03500054R4001). We would like to acknowledge MIT CSAIL Graphics Lab and James Davis for their kindly providing us with the human model data.

References

[1] Au, O.K.C., Tai, C.L., Chu, H.K., Cohen-Or, D., Lee, T.Y.: Skeleton extraction by mesh contraction. ACM Trans. Graph. 27(3) (2008)

[2] Katz, S., Tal, A.: Hierarchical mesh decomposition using fuzzy clustering and cuts. ACM Trans. Graph. 22(3), 954–961 (2003)

[3] Mortara, M., Patan, G.: Affine-invariant skeleton of 3d shapes. In: Shape Modeling International, pp. 245–252 (2002)

[4] Tierny, J., Vandeborre, J.-P., Daoudi, M.: Invariant high level reeb graphs of 3d polygonal meshes. In: IEEE 3DPVT, pp. 105–112 (2006)

[5] Ge, Y., Fitzpatrick, J.M.: On the generation of skeletons from discrete euclidean distance maps. IEEE Trans. Pattern Anal. Mach. Intell. 18(11), 1055–1066 (1996)

[6] Amenta, N., Choi, S., Kolluri, R.K.: The power crust. In: Proceedings of the Sixth ACM Symposium on Solid Modeling and Applications, pp. 249–266 (2001)

[7] Aguiar, E., Theobalt, C., Thrun, S., Seidel, H.-P.: Automatic conversion of mesh animations into skeleton-based animations. In: EUROGRAPHICS 2008, vol. 27(2) (2008)

[8] James, D.L., Twigg, C.D.: Skinning mesh animations. ACM Trans. Graph. 24(3), 399–407 (2005)

[9] Schaefer, S., Yuksel, C.: Example-based skeleton extraction. In: SGP 2007, Aire-la-Ville, Switzerland, pp. 153–162 (2007)

[10] Hu, X.Y., Liang, X.H., Zhao, Q.P.: Match the virtual human model and motion data automatically. Journal of Software 17(10), 2181–2191 (2006) (in Chinese)

[11] Yu, Y.: Research on Individual Modeling of Human Body Based on Examples. PhD thesis, Chinese Academy of Sciences. Institute of Computing Technology (2009) (in Chinese)

[12] Baran, I., Popoví, C.J.: Automatic rigging and animation of 3d characters. ACM Trans. Graph. 26, Article No.:72 (2007)

[13] Wu, W., Hao, A., Yongtao, Z.: Skeleton extraction from 3D dynamic meshes. In: Proceedings of 1st International Conference on Applied Digital Human Modeling, pp. 247–256. CRC Press, Boca Raton (2010)

GPU-Based Large Seismic Data Parallel Compression

Kai Xie, H.Q. Yu, and G.Y. Lu

School of Electronic Information, Yangtze University, Jingzhou 434102, China
pami2009@163.com

Abstract. The motivation of compressing seismic data is to reduce enormous storage space and transmission bandwidth. In this paper, a parallel compression technique is presented for large seismic data compression based on graphics processing unit (GPU). We take advantage of combinations of GPU-based parallel processing and 3D set partitioning in hierarchical tree (3D SPIHT) to accelerate the whole compression process. Experimental results show that our method achieve fast compression for very large seismic data (2.56GB) on standard PC hardware.

Keywords: parallel compression, graphics processing unit, large seismic data, 3D SPIHT.

1 Introduction

The focus of this paper is on lossless seismic data compression, which will reduce enormous storage space and transmission bandwidth. A modern seismic acquisition survey usually can produce up to 100-1000 terabytes of data. This number is still increasing due to 3D or 4D survey instead of the traditional 2D survey. Therefore, large seismic data compression has become necessary to save storage space and access time.

In this paper, we present an efficient volume compress algorithm using graphics processing units (GPUs) for compressing large seismic datasets on standard PC hardware. We take advantage of combinations of GPU-based parallel processing and 3D set partitioning in hierarchical tree (3D SPIHT) to accelerate the whole compression process. As a result, our method achieves fast compression for very large seismic data (2.56GB) on standard PC hardware. The remainder of the paper is structured as follows. In Sec. 2, we briefly review the earlier work on the compression of volume data. Sec 3 gives an overview of our algorithm, and presents each step of the pipeline of our algorithm in detail. In Sec. 4, we analyze the performance of our algorithm, and in Sec. 5 we conclude the paper and also present our future work.

2 Background

We briefly survey earlier work on the compression of volume data and graphics processing unit (GPU)-based parallel processing. Said and Pearlman proposed a efficient coding algorithm using set partitioning in hierarchical tree (SPIHT), and implemented to lossless[1] compression of images. Y. Kim and Pearlman [2] utilized it for volume image compression. In this paper, we use Kim and Pearlman's lossless 3D SPIHT with asymmetric tree structure. In past ten years, the floating-point performance of GPUs has increased dramatically over CPUs (Fig. 1 Source: [3]). This difference is also the reason that significant computational performance can be

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 339–345.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

achieved by using the GPU. Recently, significant work is being conducted in the area of seismic data computing using GPUs. For example, Deschizeaux and Blanc [4] solved the SRMIP migration algorithm and achieved up to 15× speedup over their optimized CPU implementation. Other examples include Shi [5] solving curved-ray, prestack Kirchhoff time migration 16.3×faster than an optimized CPU version, and Wang [6] achieving an impressive 25-40× speedup for a Fourier migration.

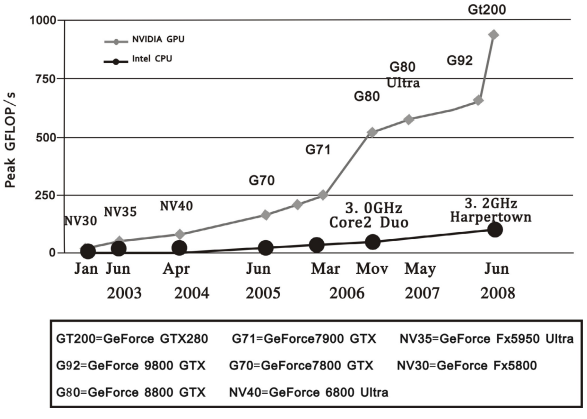


Fig. 1. Floating-point operations per second for the CPU and GPU

3 GPU-Based Parallel Compression Algorithm

3.1 Encoding Algorithm

The pipeline of our encoding algorithm can be subdivided into four major stages (see Fig. 2):

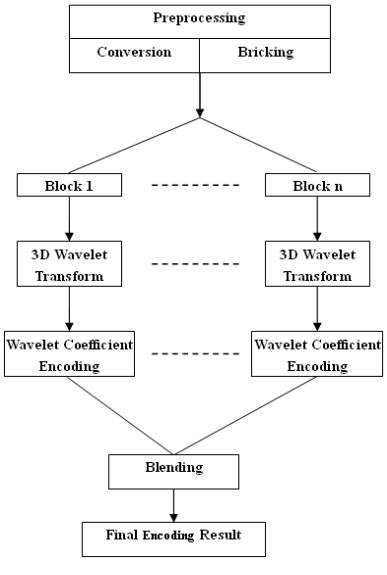


Fig. 2. Encoding pipeline

1. Pre-processing: Our implementation requires a pre-processing step before compression of large seismic data.

(a)Conversion: Prior to the compression, all seismic trace are converted to volume trace (see Fig. 3).

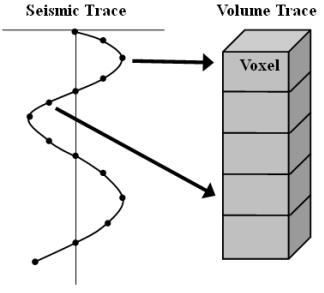


Fig. 3. Convert seismic trace to volume trace

(b)Bricking: the volume data is subdivided into small blocks.

2. 3D Wavelet Transform: To transform signals of higher dimension, we apply the 1d wavelet transform in all three dimensions separately, resulting in a 3d tensor product wavelet transform as seen in Fig. 4. After the first wavelet transform of the whole volume, we apply the 3d wavelet transform to the low sub band recursively. We repeat this step four times to get the sub bands seen in the Fig. 4.

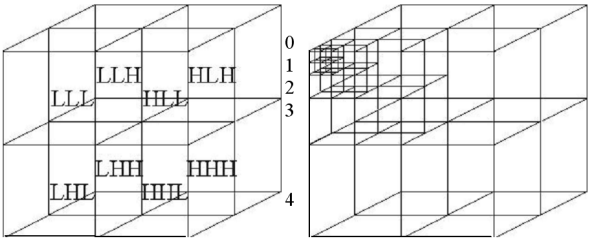


Fig. 4. Single step of three dimensional translation (L=low pass filtered, H=high pass filtered) and complete recursive decomposition

3. Wavelet Coefficient Encoding: To form trees of 2-D SPIHT as shown in Fig. 5 (a) (Source: [3]), groups of 2×2 coordinates were kept together in the lists. On the 3D subband structure in Fig. 5 (b), there are 3D transaxial and axial trees, and their parent-offspring relationships. To apply wavelet packet decomposition, the full axial decomposition precedes the transaxial decompositions, where the node divides in the additionally split subbands. The symmetric tree structure of 3D SPIHT is a straightforward extension from the 2D case to form a node in 3D SPIHT as a block of eight adjacent pixels with two extending to each of the three dimensions, hence forming a node of 2×2×2 pixels. The asymmetric tree structure shown in Fig. 5 (c) in each coefficient frame is exactly the same as the 2-D SPIHT tree structure except that the top-left coefficient of each 2×2 group in the lowest transaxial subband of LLt and LHt bands links to a group in another axial subband at the same transaxial subband location.

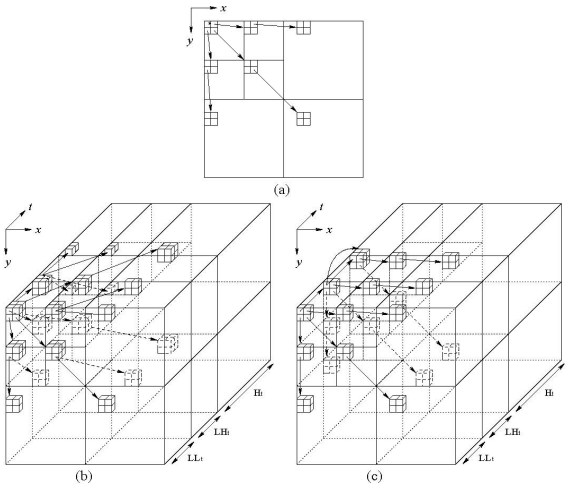


Fig. 5. Comparison of tree structures (a) 2-D original tree structure after 2 level spatial decomposition (b) 3D original tree structure after 2 level wavelet packet decomposition (c) 3D asymmetric tree structure after 2 level wavelet packet decomposition

4. Blending: The encoding results of all blocks are blended to generate final encoding result.

3.2 Decoding Algorithm

The pipeline of our decoding algorithm can be subdivided into three major stages (see Fig. 6):

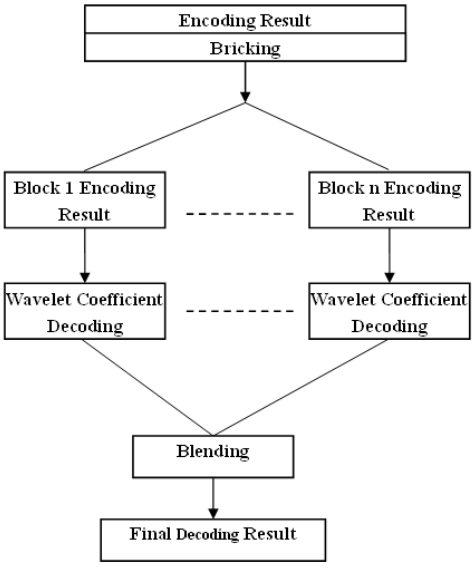


Fig. 6. Decoding pipeline

1. Bricking: the encoding result is subdivided into small blocks.
2. Wavelet Coefficient Decoding.
3. Blending: The decoding results of all blocks are blended to generate final decoding result.

4 Result and Analysis

We have implemented our compression algorithm using Visual C++, and measure its performance on a dual-core 2.4 GHz Pentium4 PC equipped with NVIDIA Quadro FX 4800 (1.5G video memory). The testing data we have used in our experiments are shown in Tab. 1, having data sizes of 429.8 MB and 2623.6 MB, respectively. Figure 7 describes the thread decomposition for a 3D volume (circled in green) that represents the entire grid being computed. Blocks (yellow cube) process regions of the volume using parallel threads (red square) that compute a value for every voxel in the volume. We split the entire volume into textures in powers of two, which is supported by the current GPU architecture. For example, for a volume data of 512×512×512×4 bytes, we subdivide it into 512 unit blocks, each having 64×64×64×4 bytes.

Table 1. Testing 3D seismic data

Name	Data dimensions	Size of Voxel(BYTE)	Data Size(MB)
Case1	626×600×300	4	429.8
Case2	1426×1256×384	4	2623.6

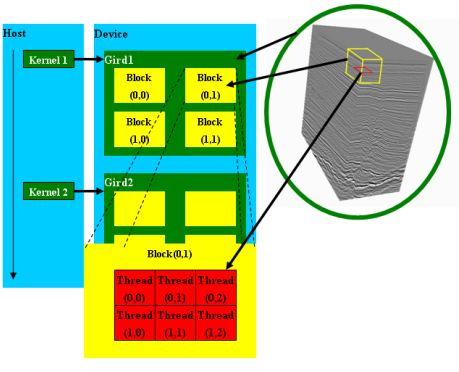


Fig. 7. Thread decomposition on the GPU using NVIDIA's CUDA

We compare the encoding time and the decoding time of our algorithm (3D GSPIHT) with original 3D SPIHT. The results show our algorithm achieve an impressive 5-10× speedup. The average speedup of our algorithm is shown on Tab.2. A 3D seismic data compression and visualization system developed by our lab is shown in Fig. 8.

Table 2. Testing 3D seismic data

Name	# × speedup of encoding time	# × speedup of decoding time
Case1	17.0	17.6
Case2	17.3	19.1

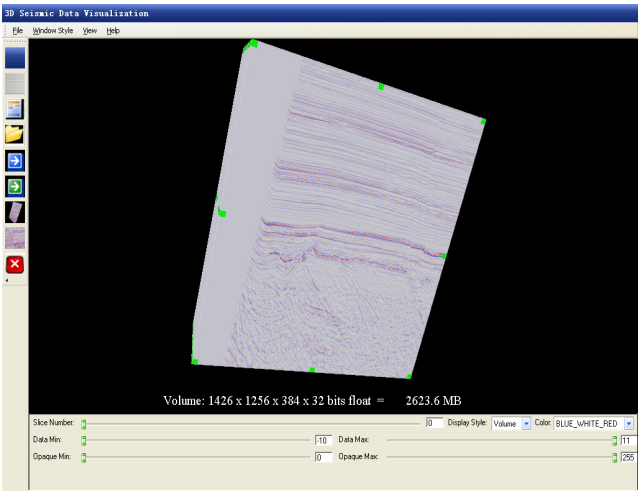


Fig. 8. 3D Seismic Data Compression and Visualization System. The dimensions of the volume datasets is 1426×1256×384×4 Bytes (2.56G)

5 Conclusions

We have presented a parallel compression method for large seismic data compression. By taking advantage of GPU-based parallel processing and 3D SPIHT, we can compress large seismic data quickly on standard PC hardware. In the future, we want to use our parallel compression method for 4D seismic data compression.

Acknowledgements. This work has been partially supported by CNPC Innovation Foundation (2010D-5006-0304), Specialized Research Fund for the Doctoral Program of Higher Education of China (20070532077), Natural Science Foundation of Hubei Province of China (2009CDB308), Educational Fund of Hubei Province of China (Q20091211).

References

[1] Said, A., Pearlman, W.: An image multiresolution representation for lossless and lossy compression. IEEE Transactions on Image Processing 5, 1010–1303 (1996)
[2] Kim, Y., Pearlman, W.: Lossless Volumetric Medical Image Compression. In: Proceedings of SPIE, Applications of Digital Image Processing, pp. 305–312 (1999)

[3] NVIDIA CUDA Programming Guide Version 3.0, p. 2 (2010)
[4] Deschizeaux, Blanc: Imaging Earth's Subsurface Using CUDA. GPU Gems, p. 478 (2007)
[5] Shi, X., Li, C., Wang, X., Li, K.: A practical approach of curved ray prestack Kirchhoff Time Migration on GPGPU. In: I Advanced Parallel Processing Technologies 8th International Symposium, pp. 165–176 (2009)
[6] Wang, S.Q., Zhang, J.H., Yao, Z.X.: Accelerating 3D Fourier migration on graphics processing units. SEG Expanded Abstracts, 3020–3024 (2009)
[7] Sungdae, C., William, A.P.: Error Resilient Video Coding With Improved 3-D SPIHT and Error Concealment. In: Proceedings of the SPIE Image and Video Communications and Processing 2003, vol. 5022, pp. 125–136 (2003)

Transform Optimization Based on Color Image Compression

Kai Xie, H.Q. Yu, and G.Y. Lu

School of Electronic Information, Yangtze University, Jingzhou 434102, China
pami2009@163.com

Abstract. A transform optimization based preprocessing algorithm is presented. The method integrates combination of transforms, color component weighting and CSF filtering. Experiment results showed that proposed algorithm outperforms JPEG2000 lossy and lossless algorithm in both objective and subjective quality in wide range of compression rate.

Keywords: color image compression, color contrast sensitivity Function (CSF), wavelet transform.

1 Introduction

The image compression algorithms improved over the last decades. The latest milestone was set by the standardization of JPEG2000 [1, 2, 3]. However, most of compression algorithms are aiming at improving coding efficiency of single component image compression which differs from color image compression greatly. To improve coding efficiency of color image, in JPEG2000, two preprocessing algorithms are employed, lossy one and lossless one. The lossy algorithm consists of ICT and CDF 9/7 wavelet transform, whereas the lossless algorithm consists of RCT and 5/3 wavelet transform.

In this paper, we will study the preprocessing algorithms in JPEG2000 and propose a preprocessing algorithm used in lossy color image compression. Experiment results proved that proposed algorithm outperforms JPEG2000 lossy algorithm and JPEG2000 lossless algorithm by 1–2 db through the compression range from 0.2 BBP to 6 BBP and can replace JPEG2000 lossy algorithm as preprocessing algorithm for all wavelet-base color image compression.

2 Proposed Algorithm

The main process of proposed algorithm is showed in figure 1. Our algorithm consists of three parts. The first part is combination of RCT and CDF 9/7 wavelet transform and it outperforms both JPEG2000 lossy combination and JPEG2000 lossless combination in improving objective quality. The second part is weighting the components came from the RCT, which improves objective quality further. The last part is CSF filtering and it improves the perceptive quality of reconstruction images of proposed algorithm. In the following part of the paper, we will discuss these parts in detail.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 347–353.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

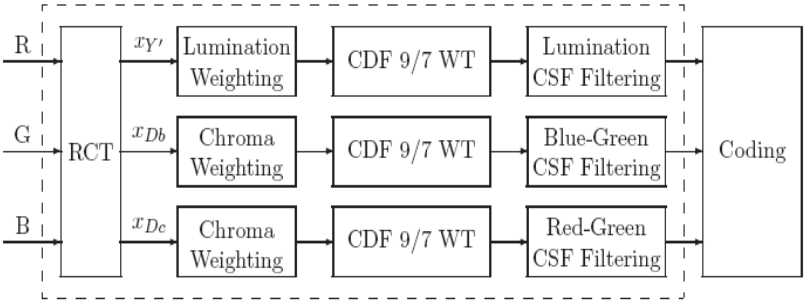


Fig. 1. Diagram of Proposed Algorithm

2.1 Selection of Combination of Wavelet Transform

The correlation of color components in RGB color space is very large and it is possible to reduce this correlation and improve coding efficiency when transformed to other less correlated color spaces. Researches of years has established ICT(irreversible component transform) and RCT(reversible component transform), defined in JPEG2000 still images compression standard, as the most favorite color space transforms used in compression.

As the transforms employed by JPEG2000 are proved to be almost the best choice [4] in their category (lossy and lossless), we experiment on the four combinations of color component transform and wavelet transform. In the figure 2, experiment results on Lena and Hell, which are both standard testing image for compression, is showed. Both results showed that in very wide range of compression rate, combination of RCT and CDF 9/7 wavelet transform has the best performances comparing with other combinations. Experiments on great number of testing images also come to the same results. The wide range of compression rate where combination of RCT and CDF 9/7 wavelet transform proved that this combination is capable to replace JPEG2000 lossy preprocessing algorithm as usual lossy pre-coding algorithm.

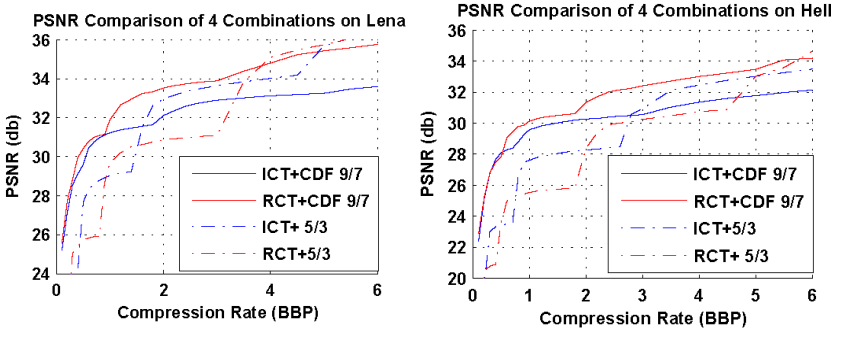


Fig. 2. Coding Efficiency Comparison of Four Different Combinations

2.2 Weighting of Different Component

After color component transform, the standard derivations of each component have been changed which means that information has been partially concentrated to some component. So we should append weights to different color components to distribute bits on different components more reasonably.

In figure 3, weighting effects have been showed. We can see that almost in all range of compression rate, weighting improve the PSNR value strikingly and furthermore the superiority appears greater and greater. Experiments on other standard testing image also get similar results.

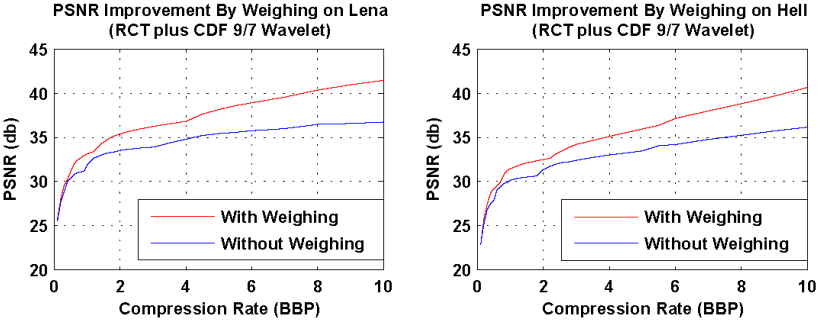


Fig. 3. PSNR Improvement by Weighting

2.3 Perceptive Quality Improvement via Color CSF

Many researchers have proved that objective quality metrics, such as PSNR, MSE, do not exactly reflect the perceptive quality of human eyes. In this paper, we exploit the color CSF to improving coding efficiency.

Human eyes are most sensitive to middle frequency signals, a little bit less sensitive to low frequency signal, and when frequency increases (above certain threshold), the sensitivity decreases exponentially. For chromatic stimulus, the sensitivity just decreases as the frequency of signal increases.

Using the method described in [5], we designed the CSF filters for 5 level wavelet decompositions of three separate color components. The perceptive quality of reconstruction image can be improved by filtering the wavelet coefficients by these filters. We will discuss the result later.

3 Results and Analysis

In our experiments, SPIHT algorithm is employed as coding algorithm. Since our method is a preprocessing algorithm, any other wavelet-based algorithm will get the similar result.

We have made experiments on various standard testing images. Firstly, we compute the PSNR value of reconstruction images to evaluate the performance of proposed scheme. In figure 4, we compare the PSNR curve of proposed algorithm, JPEG2000 lossy algorithm and JPEG2000 lossless algorithm. From the figure, we can

see that in wide range of compression rate proposed algorithm outperforms JPEG2000 lossy algorithm and JPEG2000 lossless algorithm by 1-2 db and objective quality has been improved strikingly.

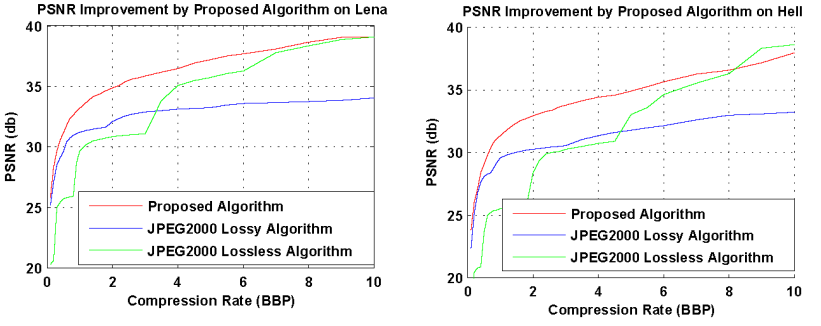


Fig. 4. Objective Quality Evaluation of Different Algorithms

Since objective quality does not coincide with perceptive quality, we employ WPSNR (weighted peak signal-to-noise ratio) as perceptive quality metric and draw WPSNR curves of preprocessing algorithm with CSF filtering and without CSF filtering. In figure 5, the improvement of perceptive quality is conspicuous.

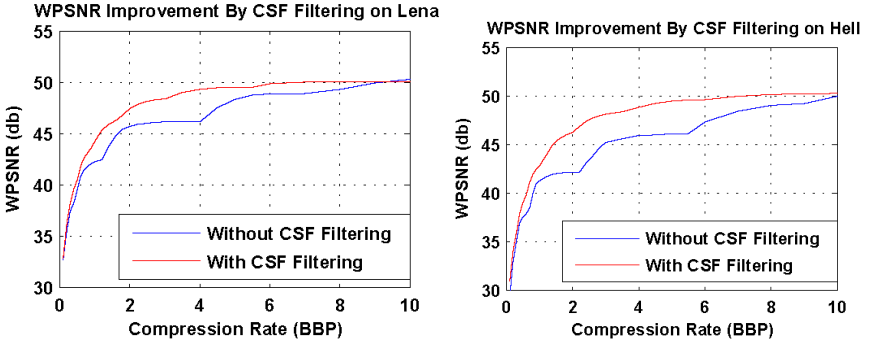


Fig. 5. Objective Quality Evaluation of Different Algorithms

Conspicuously, only comparison of WPSNR value is not enough. We will give another proof. Figure 6(a) is reconstructed from coding scheme without CSF filtering and figure 6(b) is recovered from coding scheme with CSF filtering. The compression rate of these two reconstruction images are both 0.2 BBP. Observing figure 6, we can discover that the reconstruction image with CSF filtering has less distortion than those without CSF filtering.

In figure 7, we will give a comparison of original Lena image, reconstructing image by proposed algorithm, reconstructing image by JPEG2000 lossy algorithm and reconstructing image by JPEG2000 lossless algorithm. All of them are compressed in 0.3 BBP.

(a) (b)

Fig. 6. (a)Without CSF Filtering; (b)With CSF Filtering

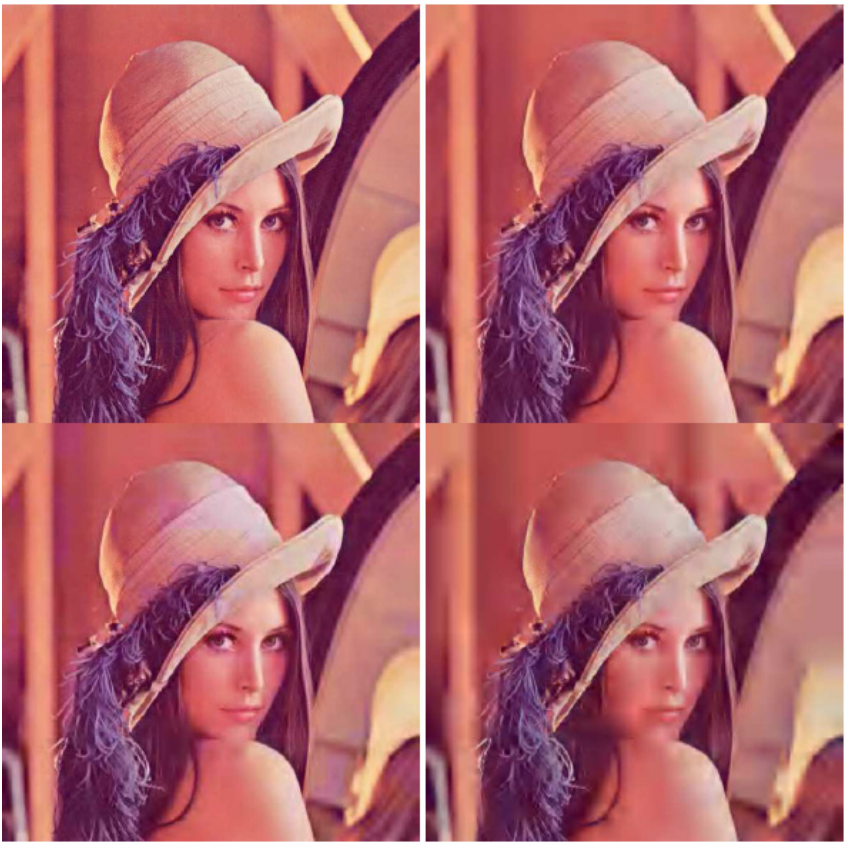


Fig. 7. The top-left image is the original Lena image. The top-right is the reconstruction image by proposed preprocessing algorithm. The bottom-left image is the reconstruction image by JPEG2000 lossy preprocessing algorithm and the bottom-right one by JPEG2000 lossless preprocessing algorithm.

4 Conclusion

The paper researched the preprocessing of wavelet-based color image compression and proposes a new preprocessing algorithm, which includes combination of RCT and CDF 9/7 wavelet transform, weighting components and CSF filtering. Experiment results showed that proposed algorithm improves both objective and subjective quality of reconstruction image and can replace JPEG2000 lossy algorithm as the preprocessing algorithm in wavelet-based color image compression.

In future, we will work on integrating other HVS features, such as masking and so on, into the proposed algorithm and we will also keep on exploiting on improving objective quality of reconstruction images further.

4.1 Structure

4.1.1 Bulleted Lists May Be Included and Should Look Like This

- First point
- Second point
- And so on

Ensure that you return to the 'Els-body-text' style, the style that you will mainly be using for large blocks of text, when you have completed your bulleted list.

4.2 Tables

All tables should be numbered with Arabic numerals. Headings should be placed above tables, left justified. Leave one line space between the heading and the table. Only horizontal lines should be used within a table, to distinguish the column headings from the body of the table, and immediately above and below the table. Tables must be embedded into the text and not supplied separately. Below is an example which authors may find useful.

Table 1. An example of a table

An example of a column heading	Column A (*t*)	Column B (*T*)
And an entry	1	2
And another entry	3	4
And another entry	5	6

Acknowledgement. This work has been partially supported by CNPC Innovation Foundation (2010D-5006-0304), Specialized Research Fund for the Doctoral Program of Higher Education of China (20070532077), Natural Science Foundation of Hubei Province of China (2009CDB308), Educational Fund of Hubei Province of China (Q20091211).

References

[1] Charrier, M., Cruz, D.S., Larsson, M.: JPEG 2000, the next millennium compression standard for still images. In: Proc. of the IEEE International Conference on Multimedia Computing and Systems (ICMCS), pp. 131–132 (1999)
[2] Taubaman, D.S., Marcellin, M.W.: JPEG 2000 image compression fundamentals, standards and practice. Chinese Edition (2004)
[3] ISO/IEC, Ed: JPEG 2000-lossless and lossy compression of continous-tone and bi-level still images: part I minimum decoder, ISO, ISO/IEC (1999)
[4] Adams, M.D.: Reversible integer-to-integer wavelet transforms for image compression: performance evaluation and analysis. IEEE Transactions on Image Processing 9(6), 1010–1024 (2000)
[5] Nadenau, M.J.: Integration of human color vision models into high quality image compression. Ph.D thesis (2000)

The Resilience Framework for Interdependent Infrastructure Systems Using the Dynamic Inoperability Input-Output Model

Wenping Xu[1,2], Zhiwu Wang[3], Mingsheng Hu[4], Liu Hong[1,*], Xueguang Chen[1], and Shuliang Wang[1]

[1] Department of Control Science and Engineering, Institute of System Engineering, Huazhong University of Science and Technology. Wuhan, Hubei 430074, HB, China
[2] Information Engineering Department, Wuhan University of Technology Huaxia College, Wuhan, Hubei 430223, HB, China
[3] Henan Institute of Engineering, Zhengzhou, Henan 451191, HN, China
[4] Institute of Software Science, Zhengzhou Normal University, Zhengzhou, Henan 451191, HN, China
hongliu1978@mail.hust.edu.cn

Abstract. Resilience is increasingly used as an approach for understanding the dynamics of critical interdependent infrastructure system after a natural disaster. In this paper, a probabilistic resilience framework after a disaster is presented for the critical interdependent infrastructure sectors using the inoperability metric of the Dynamic Inoperability Input-output Model (DIIM), the DIIM has been developed to help understand infrastructure interdependencies on deliberate external attacks or natural disasters. It would offer greater potential in application of sector resilience, especially in the process of integrated risk management and dynamic recovery after a disaster or perturbation.

Keywords: Interdependentp, infrastructure, inoperability, resilience, DIIM, dynamic recovery.

1 Introduction

During the past few years, the world has witnessed some unprecedented natural disasters such as Earthquake Wenchuan China, snow disaster, flood and so on, Although some preventative measures have been taken, the reality is that future disasters cannot be prevented due to the likelihood that these events will involve unexpected forms to damage the critical interdependent infrastructure systems, so it is very important to enhance the capacity of the interdependent infrastructure system to resist and recover from the disasters to minimize economic loss after a disaster. The government of China has placed increased emphasis on the resilience and vulnerability of the critical infrastructure system. These systems include electric power system; rail transportation system; water supply system; gas supply system or communication links etc. To understand this property, many studies have focused on

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 355–361.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

conducting consequence analyses that estimate the impacts on these interdependent infrastructures (and on the economy and other systems that rely on them) of some disruptive events [see 3,10 and 11].

Resilience, broadly defined as the capacity to resist and recover from loss [see 5, 9], is an essential concept in the interdependent infrastructures research after a natural disaster or other disruptions and is central to the development of loss reduction at the local, national and international levels. Resilience is increasingly used as an approach for understanding the dynamic of critical interdependent infrastructure system after natural hazards. It would offer greater potential in application of sector resilience, especially in the process of integrated risk management and dynamic recovery after a disaster. The resilience characteristic is highly desirable for assessing the vulnerability of the critical interdependent infrastructure systems. Moreover, the resilience can be described in a probabilistic way, in this paper, we build a probabilistic resilience to measure the economic losses for the critical interdependent infrastructure system after natural hazards.

The inoperability input-output analysis was the most widely used modeling approach to the subject. In this paper, a probabilistic resilience framework after a disaster is presented for the critical interdependent infrastructure sectors using the Dynamic Inoperability Input-output Model (DIIM), it provides an approach to incorporate the dynamic evolution characteristics of the resilience and the probabilistic features of the duration of a disruption events. The probabilistic resilience framework is more applicable when less information is known about a particular infrastructure sector or system, and when we want to evaluate the resilience of the sector or system.

The rest of the paper is organized as follows. Preliminaries are described in section 2. Section 3 gives the probabilistic resilience framework after a disaster for the critical interdependent infrastructure sectors. The paper is concluded in section 4.

2 Preliminaries

This section gives the preliminaries about the Inoperability Input-output Model (IIM) and an extension to the model, denoted as DIIM [see 4].

2.1 Demand-Driven Inoperability Input-Output Model

The demand-reduction IIM is presented in Equation(1). Note that all general input-output equations provided in this paper assume an economy of sectors, resulting in matrices of size and vectors of length, and that all vectors, unless otherwise noted, are column vectors [1, 2]:

$$q = A^* q + c^* \tag{1}$$

If $\det(I - A^*) \neq 0$, we have:

$$q = (I - A^*)^{-1} c^* \tag{2}$$

Where the symbol q is the vector of resulting normalized production loss due to demand reduction; and A^* is the interdependency matrix derived from the economic input-output data; and c^* is the vector of normalized degraded demand resulting from a perturbation.

The Equation (3) is defining inoperability as the degraded normalized production loss of sector i, where q_i is given as:

$$q = diag(\hat{x})^{-1}\left(\hat{x} - \tilde{x}\right) \tag{3}$$

Where the symbol $\hat{x}$ is the value of the nominal production level and $\tilde{x}$ is the value of the degraded production level after perturbation. The symbol $\hat{c}$ is the value of the nominal final demand and $\tilde{c}$ is the value of the degraded final demand after perturbation:

$$A^* = diag(\hat{x})^{-1} A diag(\hat{x}) \tag{4}$$

$$c^* = diag(\hat{x})^{-1}\left(\hat{c} - \tilde{c}\right) \tag{5}$$

The vector q represents the inoperability vector, the elements of which measure the proportion of "unrealized" production per as-planned production resulting from a reduction in demand. The demand perturbation is expressed by vector c^*, whose elements represent the difference in as-planned demand and perturbed demand divided by as-planned production, quantifying the reduced final demand as a proportion of total as-planned output. The matrix A^* is the normalized interdependency matrix describing the extent of economic interdependence between sectors of the economy. More detail on the IIM components, including numerical examples, is provided by Santos and Haimes [see 6, 7].

2.2 Demand-Driven Dynamic Inoperability Input-Output Model

The demand-driven Dynamic Inoperability Input-output Model (DIIM), shown in Equation (6), describes the recovery of industry sectors following a disruption [see 4, 8]:

$$\dot{q}(t) = K[A^* q(t) + c^*(t) - q(t)] \tag{6}$$

Definitions of $q(t)$, A^* and $c^*(t)$ are the same as those for their counterparts in the demand-reduction IIM, except that $q(t)$ and $c^*(t)$ describe those values at a specific time t, and K is a matrix with resilience coefficients $k_1, k_2, \cdots, k_n$ on the diagonals and zeroes elsewhere.

From Equation (6), we can get:

$$q(t) = e^{-K(I-A^*)t} q(0) + \int_0^t K e^{-K(I-A^*)(t-s)} c^*(s) ds \tag{7}$$

If it is assumed that $c^* = 0$ and sector i recovers from some initial inoperability $q_i(0) > 0$, to some inoperability $q_i(T_i) > 0$ at the known time T_i, then k_i can be

defined in Equation (9), where a_{ii}^{*} are the diagonal elements of A^{*}. The calculation of k_i is derived from the continuous time DIIM in Equation (8). The published result in Lian and Haimes to model sector resilience coefficient is as follows [see 4]:

$$k_i = \frac{\ln\left[q_i(0) \Big/ q_i(T_i) \right]}{T_i} \left(\frac{1}{1 - a_{ii}^{*}} \right) \tag{8}$$

The resilience coefficient k_i represents the ability of sector i to recover following a disruption, where the greater values of k_i correspond to a faster response by the sector to a perturbation. And k_i is a characteristic of an appointed economic sector to measure a sector's ability to adjust an imbalanced on the supply and demand.

If the integration operation is applied to Equation (7), the vector of the economic loss during the recovery for each sector can be derived in the following Equation (9):

$$. \left(Q_1, Q_2, \cdots, Q_n \right) = \left[diag\left(\hat{x} \right) \left(I - A^{*} \right)^{-1} K^{-1} q(0) \right] \tag{9}$$

From the relationship of k_i and T, which $T = diag(t_i)$, the vector of the individual economic loss during the recovery for each sector can be directly described as functions of recovery time vector:

$$\left(Q_1, Q_2, \cdots, Q_n \right) = \left[diag\left(\hat{x} \right) \left(I - A^{*} \right)^{-1} diag\left(t_i \right) diag^{-1}\left(\omega_i \right) q(0) \right] \tilde{l}$$

3 The Resilience Framework

Inoperability is defined by the inability of the system to perform its intended function [see 1], which connotes the level of dysfunction, also is analogous to the idea of unreliability from the field of reliability engineering. The measurement of resilience is very important because it enables us to evaluate strategies for reducing economic losses from the disruptive events. In this section, we use the inoperability metric of the DIIM to measure the resilience for the critical interdependent infrastructure system and give a probabilistic resilience framework for the critical interdependent infrastructure systems.

The basic definition of resiliency is "the ability of the system to withstand a major disruption within acceptable degradation parameters and to recover within an acceptable cost and time." [see 6] Based on this definition of resilience, Fig. 1 attempts to give a probabilistic resilience framework after a disaster for the critical interdependent infrastructure systems with the inoperability metric using the DIIM.

Resilience is presented as a way of measuring the ability of one sector or system to reduce the disruptions or perturbations; it reflects the ability of one sector or system to withstand the potential higher disruption or perturbation's impact. Furthermore, the

resilience can also reflect the fact that the sector or system could withstand a larger disruption or perturbation than expected because of the various reasons.

To improve the resilience of the interdependent infrastructure systems or sectors, in other words, to decrease the inoperability in the desired level as soon as possible, we give a probabilistic resilience framework .According to the Fig. 1, the resilience can be realized from the following three aspects: (1) reduce the impact of damage propagation; (2) reduce the response time; (3) decrease the recovery time. And so on.

Fig. 1 gives a probabilistic resilience framework after a disaster for the critical interdependent infrastructure systems.

From Fig 1, the time t_1 represents the time-delayed period of the response of the infrastructure sector or system because of an initial disruption, the time paragraph $t_2 - t_1$ reacts the damage propagation after a disaster due to the interdependent of the infrastructure sectors or systems. $t_3 - t_1$ is the response time of the decision maker for risk management after a disaster. t_4, t_5, t_6 reflect the recovery process, we suppose that the recovery time is a random variable, the recovery time distribution is generated by the domain expert opinions or historical databases. It obeys the triangular distribution: the recovery time for one sector or system is estimated by the domain experts or historical data: the lowest recovery time is $t_4 - t_3$, most likely recovery time is $t_5 - t_3$, the highest recovery time is $t_6 - t_3$.

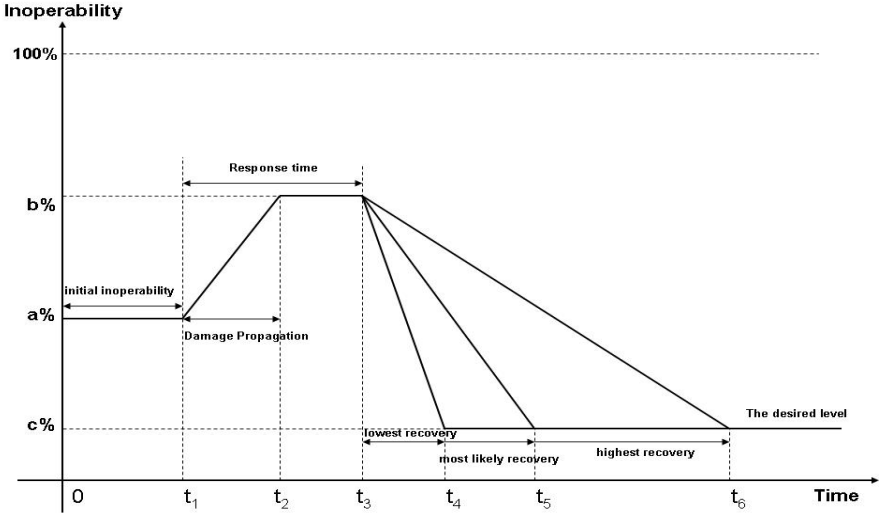


Fig. 1. A probabilistic resilience framework after a disaster for the critical interdependent infrastructure systems

This probabilistic resilience framework measures the efficacy of risk management strategies that promote the entire recovery. From Equation (10), we can draw the conclusion that reducing recovery time can reduce the economic losses.

4 Conclusions

The measurement of resilience is very important and failure to incorporate resilience in loss estimation will result in risk assessments from interruptive events. In this article, although we contribute significant advances to the dynamic inoperability input-output modeling of the economic impacts of natural and other types of disasters, the probabilistic resilience framework after a disaster is presented for the critical interdependent infrastructure sectors or systems using the inoperability metric of the dynamic inoperability input-output model (DIIM). It would offer greater potential in application of sector resilience, especially in the process of integrated risk management and dynamic recovery after a disaster. In the future work, It will be used other metric to measure the resilience for the critical interdependent infrastructure sectors or systems.

Acknowledgements. This study was supported in part by the National Natural Science Foundation of China under Grant No.60903174, No. 61142010 and also supported by the fund of Key Lab for Image Processing and Intelligent control (20093).

References

[1] Haimes, Y.Y., Horowitz, B.M., Lambert, J.H., Santos, J.R., Lian, C., Crowther, K.G.: Inoperability input-output model for interdependent infrastructure sectors. I: Theory and methodology. Journal of Infrastructure Systems 11(2), 67–79 (2005a)

[2] Haimes, Y.Y., Horowitz, B.M., Lambert, J.H., Santos, J.R., Crowther, K.G., Lian, C.: Inoperability input-output model for interdependent infrastructure sectors. II: case studies. Journal of Infrastructure Systems 11(2), 80–92 (2005b)

[3] Santos, J.R., Orsi, M.J., Bond, E.J.: Pandemic Recovery Analysis Using the Dynamic Inoperability Input-Output Model. Risk Analysis 29(12), 1743–1758 (2009)

[4] Lian, C.Y., Haimes, Y.Y.: Managing the risk of terrorism to interdependent infrastructure systems through the dynamic inoperability input-output model. Systems Engineering 9(3), 241–258 (2006)

[5] Greenberg, M.R., Lahr, M., Mantell, N.: Understanding the Economic Costs and Benefits of Catastrophes and Their Aftermath: A Review and Suggestions for the U.S. Federal Government. Risk Analysis 27(1) (2007)

[6] Haimes, Y.Y., Crowther, K., Horowitz, B.M.: Homeland Security Preparedness: Balancing Protection with Resilience in Emergent Systems. Systems Engineering 11(4) (2008)

[7] Santos, J.R., Haimes, Y.Y.: Modeling the demand reduction input-output (I-O) inoperability due to terrorism of interconnected infrastructures. Risk Analysis 24(6), 1437–1451 (2004)

[8] Lian, C.Y., Haimes, Y.Y.: Extreme risk analysis of interdependent economic and infrastructure sectors. Risk Analysis, 2007 27(4), 1053–1064 (2006)

[9] Zhou, H., Wang, J., Wan, J., Jia, H.: Resilience to natural hazards: a geographic perspective. Nat. Hazards 53, 21–41 (2010)

[10] Orsi, M.J., Santos, J.R.: Estimating Workforce-Related Economic Impact of a Pandemic on the Commonwealth of Virginia. IEEE Transactions on Systems, Man and Cybernetics—part A: Systems and Humans 40(2), 301–305 (2010)
[11] Orsi, M.J., Santos, J.R.: Incorporating Time-Varying Perturbations Into the Dynamic Inoperability Input–Output Model. IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans 40(1), 100–106 (2010)

Multi-source Separation Using over Iterative Empirical Mode Decomposition

Yutian Wang, Hui Wang, and Qin Zhang

The Information Engineering School, Communication University of China,
1st Dingfuzhuang East street, Beijing, China 100024

Abstract. In this paper, we propose a novel multi-source fundamental frequency estimation method for music signals which is based on over iterative empirical mode decomposition (EMD). Multi-source fundamental frequency estimation in single channel polyphonic signal has been developed for a long time but still a challenging problem. By means of using over iterative EMD, we can obtain predominant frequency modulation subband signal named Intrinsic Mode Function (IMF) without the interference of amplitude modulation. Then the fundamental frequencies are obtained by a salience estimation method. The experimental results agree that the proposed method can extract the pitch content efficiently.

Keywords: Multiple fundamental frequency estimation, Polyphonic music signal separation, Empirical mode decomposition, Intrinsic mode function.

1 Introduction

Multi-source separation of single-channel audio signal has been one of most challenging problems in audio signal processing and addressed wide range interesting of researchers [1]. Estimation of multiple fundamental frequencies play a key role in polyphonic signal separation problem, but automatically extracting the fundamental frequencies, commonly perceived as pitch, from recorded audio signals is a difficult task.

A lot of multiple fundamental frequency estimation methods use iterative-based approach, in which the fundamental frequency is estimated as the predominant signal and removed before the next iteration until all harmonic sources are extracted[2]. These methods are easy to implement because it allows using single pitch estimation algorithms in each iteration. The joint estimation methods can yield better results with higher complexity [3]. Other kinds of methods using correlogram[4], probabilistic framework with harmonic model[5] and non-negative matrix factorization (NMF)[6] also have their advantage respectively.

Huang *et. al.* raised a novel signal processing method named Hilbert-Huang Transform (HHT) based on EMD (Empirical Mode Decomposition), which is suitable for analyzing nonlinear and non-stationary signal[7]. Different from the traditional signal processing methods, the HHT is an adaptive decomposition method and can yield

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 363–368.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

more physical results. EMD is a complete, approximately orthogonal and self-adaptive method which has the ability to decompose signal by time scale. Some numerical experiments show that EMD behaves as a dyadic filter bank [8][9].

The method proposed in this paper is based on the over iterative EMD which use large iteration times to yield unity envelop IMF. Although over iterative will decrease the physical meaning of IMFs, it can concentrate energy into predominant frequency component which contain pure frequency modulation signal of the fundamental frequency. After this processing, the harmonic sources are allotted into individual IMF components. In each particular IMF, there always have one predominant harmonic source. By means of a salience estimation method, the corresponding fundamental frequencies can be extracted.

In this paper, the EMD method and its filtering property are reviewed firstly. Then the proposed algorithm which is the combination of over iterative EMD and salience estimation is described. Finally, a two instruments mixture signal is used to demonstrate the performance of our method. The experimental results agree that our algorithms can extract fundamental frequency efficiently.

2 Filter Bank Property of EMD

To analysis the nonlinear and non-stationary signal, the definition of instantaneous frequency and energy is needed, which must be functions of time. Empirical Mode decomposition (EMD) was proposed to decompose a complicated signal into a group of IMFs. Then the instantaneous frequency can be obtained by Hilbert Transform. The detail algorithm of EMD was summarized in [7].

The EMD result is highly determined by the sifting stop criterion. Different stop criterions make results vary. Although there are several kinds of stop criterion, Wu *et. al.* suggest fixed sifting time criterion[10]. In separate research, Flandrin and Wu point out that EMD is in fact a dyadic filter bank [8][9]. In the research by Wu *et. al.* [14], the dyadic property is available only shift times are about 10. Too many or too few iteration numbers would decrease the dyadic property. When the EMD shifting iteration number become higher, the filter banks gradual change narrower and their central frequency get closer, which enable the EMD to split the predominant frequency components. The filter-bank property with different iteration sifting number shows in Fig. 1. From the Fig. 1 we can reach the conclusion that with the sifting number increasing from 10 to 100, the filter bands decrease proportionally.

Unfortunately, with increasing number of iterations, the amplitude envelops trend to a straight line which will decrease the physical mean of IMFs. Furthermore, the Fourier transform can be recognized as special case of EMD when the sifting number goes to infinity. G. Rilling[11] analysis the EMD separation ability of two sinusoidal functions with different amplitude and frequency ratio and finds out the interval of parameters which can yield correct results.

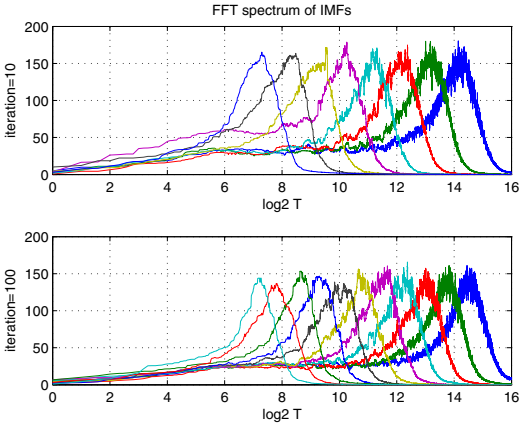


Fig. 1. Filter bank property with different iteration number

3 Fundamental Frequency Estimate Algorithm

For every harmonic source in the polyphonic signal, separating their primal component is the first and the most important step. To detect the pitches in a mixture signal, we use over iterative EMD method with sifting number 1000. From the knowledge introduced previously, we can reach the conclusion that large iteration number will reduce the physical mean of IMF but provide extremely sharp filter ability of the filter-bank. Beside it, due to the ability of decomposition signal by time scale, EMD makes the predominant signal components raise in particular IMFs. It should be noticed that with the increasing of iteration number, the centroid frequencies of the filter-bank are getting closer but the leakage also becoming serious. To overcome this problem, we should pick up available IMFs before estimating the fundamental frequencies. Our method is set a threshold of signal power ratio between the IMF and the whole signal.

After getting the available IMFs, the fundamental frequencies of each IMF can be obtained by:

1. Find the highest spectral peak $f_{\max}$ from the FFT spectrum of IMF.
2. Calculate $f_{ca} = f_{\max} / n$, in which the initial value of n is 10. If the amplitude of f_{ca} is higher than a given threshold μ, the f_{ca} is considered as a fundamental frequencies. Then terminate the process, otherwise go to step 3.
3. Set $n = n - 1$, go to step 2.

After this process, each IMF will produce a individual fundamental frequency. To find out the essential harmonic source in this candidate value set, a judgment method is designed as two steps. Firstly, if some of the value in the candidate set have the relationship of approximate multiple times, we can assert that these values belong to one harmonic source. If candidate values $f_j > f_i$, and

$$f_j = mf_i \pm \sigma, \quad m = 2,3,\ldots,M \ , \tag{1}$$

these two value are categorized into one harmonic source set H_i. σ denotes the frequency tolerance. Then the whole set is divided into several subsets $H_1, H_2, \ldots, H_n$. To determine which value is the real fundamental frequency in each subset H_i, candidates are ordered decreasingly by the sum of their harmonic amplitudes. It should be noticed that the candidates value must be in the range $[f_{\min}, f_{\max}]$, which is used to avoid the high and low order harmonics influence the result. Then all the possible candidate combinations are evaluated and the combination with the best salience is chosen. The salience is a function of the harmonic amplitudes in each IMF function.

$$S(i) = \sum_{j=1}^{C} A_{i,j} \tag{2}$$

,where C is the number of IMFs, $S(i)$ is the salience of i^{th} harmonic source H_i and $A_{i,j}$ denotes the corresponding harmonic peak amplitude of each harmonic source in every IMF.

To verify our algorithm, an mixture signal of two violin samples with pitches of A5 and D4 was evaluated. The samples are from the MUMS of McGill University. The mixture signal spectrum is shown in Fig. 2 in which the amplitude is drawn in dB scale. In the diagram the harmonics distribute in a complex way which is hard to identify and extract the fundamental frequency.

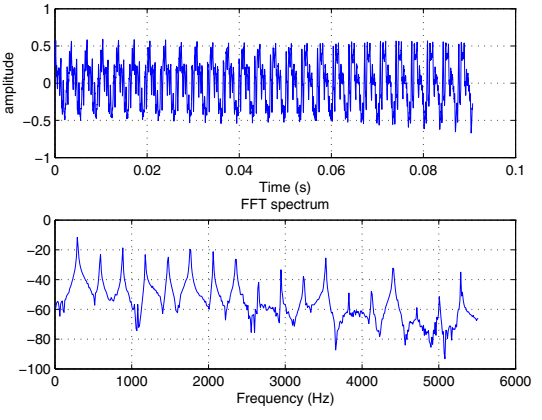


Fig. 2. The FFT spectrum of mixture signal of A5 and D4 violin

Then the over iterative EMD with sifting number 1000 is used to yield a set of IMF function, of which the power lower than 5% of total power are abandoned. After that, they are transformed into FFT spectrum. The final results are shown in Fig. 3.

From the Fig. 3 we can clear see that all the harmonic peaks are separated into different IMF and distribute in sparse and regular way. Consequently, to estimate the fundamental frequency from these spectrums become easier than directly from the original signal spectrum. By using the algorithm introduced before, we can obtain the fundamental frequency at 882Hz and 297Hz.

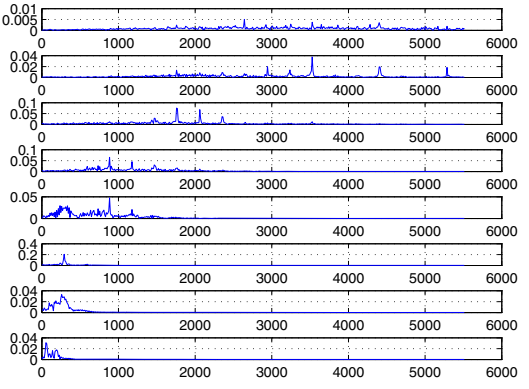


Fig. 3. The FFT spectrum of the available IMFs

4 Conclusions

In this paper we propose a single channel polyphonic music signal multiple fundamental frequency estimation method based on over iterative EMD. By use over iterative sifting number, the filter-bank property change sharper which can yield narrower band of energy. The combination of over iterative EMD and salience estimation method can obtain good results especially for multiple times fundamental frequencies mixture signal. The experiment agrees that the proposed algorithm is efficient in both same and different instruments mixture.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (Grant No. 61071149) , the State Administration of Radio Film and Television of China (Grant No. 2011-11) and Program for New Century Excellent Talents of Ministry of Education of the People's Republic of China (Grant No. NCET-10-0749).

References

1. de Cheveigne, A.: Multiple F0 Estimation (Auditory Scene Analysis, Algorithms and Applications (2006)
2. de Cheveigne, A.: J. Acoust. Soc. Am. 93(6), 3271 (1993)
3. Maher, R.C., Beauchamp, J.W.: J. Acoust. Soc. Am. 95(4), 2254 (1994)
4. Wu, M., Wang, D., Brown, G.: IEEE Transactions on Speech and Audio Processing 11(3), 229 (2003), doi:10.1109/TSA.2003.811539
5. Christensen, M.G., Stoica, P., Jakobsson, A., Jensen, S.H.: In: Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing ICASSP 2007, vol. 3 (2007), doi:10.1109/ICASSP.2007.367063
6. Smaragdis, P., Brown, J.C.: In: Proc. IEEE Workshop. Applications of Signal Processing to Audio and Acoustics, pp. 177–180 (2003), doi:10.1109/ASPAA.2003.1285860
7. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: In: Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, vol. 454, p. 903 (1998), doi:10.1098/rspa.1998.0193
8. Wu, Z., Huang, N.E.: In: Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, vol. 460, p. 1597 (2004), doi:10.1098/rspa.2003.1221
9. Flandrin, P., Rilling, G., Goncalves, P.: IEEE, Signal Processing Letters 11(2), 112 (2004), doi:10.1109/LSP.2003.821662
10. Wu, Z., Huang, N.E.: Adv. Adaptive Data Anal. 2(4), 397 (2010)
11. Rilling, G., Flandrin, P.: IEEE Transactions on Signal Processing 56(1), 85 (2008), doi:10.1109/TSP.2007.906771

Paraphrase Based Similar Expression Generation

Wei He, Zhanyi Liu, and Ting Liu

Research Center for Social Computing and Information Retrieval,
Harbin Institute of Technology
{whe,tliu}@ir.hit.edu.cn, Baiduliuzhanyi@baidu.com

Abstract. In this paper we propose a novel paraphrase based method to improve the performance of the similar expressions generation that has been widely used in writing assistant system. The users' queries are paraphrased to multiple expressions which cover abundant example sentences in the system. Three paraphrase methods are presented to satisfy the users' different intentions. First of all, we employ a statistical collocation model to generate collocations for two-word queries. Then we use a phrase based paraphrase model and a dependency based paraphrase model to substitute contiguous phrases and long-distance collocations in multi-word queries (> 2 words). The experimental results indicate that the proposed method can effectively help the users to correct their inappropriate expressions.

Keywords: Paraphrase, Similar Expression, Sentence Retrieval.

1 Introduction

Example sentence retrieval is proved helpful for computer-assisted writing in foreign languages (Liu et al. 2000). Users can find correct expressions by searching for the example sentences and learning from them. However, when the user has no idea or only fuzzy ideas about the target expression, they cannot find suitable example sentences, since the sentence retrieval system retrieves example sentences according to the users' queries. To solve this problem, (Wu et al. 2010) employed contextual information to automate suggestions for verb-noun lexical collocation. Stein et al. (2010) introduced a web service that indexes billions of n-grams, along with their occurrence frequencies on the web, and provide statistically relevant suggestions by retrieving users' queries.

Paraphrases are alternative ways to convey the same meaning. In this paper, we propose a paraphrase based method that generates similar expressions for the user to correct their inappropriate expressions. In contrast to previous researches, we employ a uniform maximum entropy framework for the suggestion generation. We perform different strategies to satisfy the users' different searching intentions. For short queries, users tend to search for the collocation of certain expressions, such as verb-noun groups. For long queries which are similar to short sentences, the intentions are usually to verify the correctness of the queries. The experiment confirms that the similar expressions generated by our method prompt writers with better expressions.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 369–374.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Suggestion Generation

2.1 Collocation Suggestion

Collocation is generally defined as a group of words that occur together more often than by chance (McKeown and Radev, 2000). In this paper, a collocation is composed of two words occurring as either a consecutive word sequence or an interrupted word sequence in sentences, such as "by accident" or "take ... advice". We use the monolingual word alignment (MWA) method (Liu et al., 2009) for collocation extraction. The MWA method adapts the bilingual word alignment algorithm to monolingual scenario to extract collocations only from monolingual corpora.

Given a two-word query <u, v>, the suggestion candidates are first constructed by replacing one word with its substitutions. Then, the optimal suggestion is detected by calculating the score. Here, we use two features to score the quality of the suggestion candidate. One is the collocation probability of the two words in the suggestion candidate. The other is similar possibility of the word and its substitution. In order to obtain a high-quality suggestion, only one word in a two-word query can be replaced by its substitutions. Thus, we can obtain two kinds of suggestion candidates, as follows:

$$< w^*, v >= \arg\max_{w \in \Re(u)} \{\log p_c(w, v) + \log p_s(w \mid u)\}$$

$$< u, w^* >= \arg\max_{w \in \Re(v)} \{\log p_c(u, w) + \log p_s(w \mid v)\}$$

Here, $p_c(u, v)$ denotes the collocation probability and $p_s(w|v)$ denotes the similarity between two words.

Finally, we combine the two kinds of suggestion candidates in a descending order due to their scores. And the n-best candidates are selected as the suggestions.

2.2 Phrasal Paraphrasing

To generate searching suggestions for long queries, we use an application-driven paraphrase generating framework (Zhao et al. 2009). The framework contains three models, a paraphrase model, a language model, and an authentic model, which control the adequacy, fluency, and the authentic degree of the paraphrases.

The procedure of paraphrase generation is described formally as follows: The input query **s** is first segmented into a sequence of *I* units $\vec{s}_1^I$, which are then paraphrased to a sequence of units $\vec{t}_1^I$. Let ($\vec{s}_1^I$,) be a pair of paraphrase units, their paraphrase likelihood is computed using a score function $\phi_{pm}(\vec{s}_1^I, \vec{t}_1^I)$. Thus the paraphrase score $p_{pm}(\vec{s}_1^I, \vec{t}_1^I)$ between **s** and **t** is decomposed into:

where λ_{pm} is the weight of the paraphrase model.

A tri-gram language model based score for the paraphrase **t** is computed as:

where *J* is the length of **t**, t_j is the j-th word of **t**, and λ_{lm} is the weight for the language model.

We use the hit number given by the sentence retrieval system to evaluate the authentic degree in the authentic model. Given an expression, a large hits number means that the expression is authentic, yet a small number indicates a bad expression.

$$p_{am}(t) = \mathrm{I}(t)^{\lambda_{am}}$$

Finally, the paraphrase model, language model and authentic model are combined on a log-linear framework:

$$\mathbf{p(t|s)} = \lambda_{pm} \sum_{i} \log \varphi_{pm}(\bar{s}_i, \bar{t}_i) + \lambda_{lm} \sum_{j=1}^{J} \log p(t_j | t_{j-2} t_{j-1}) + \lambda_{am} \log h(t)$$

We make a limitation that the hit number of the generated similar expressions must be bigger than that of the original queries. With this limitation, the paraphrasing procedure cannot cause a negative influence.

The phrasal paraphrasing algorithm includes three steps:

1. Paraphrase enumeration. The system enumerates all the paraphrase units. In this step, every continuous phrase (up to 7 words) in the query is considered.
2. Paraphrase filtration. The paraphrase units that cannot achieve a positive result of target function () are filtered out. (where q denotes the user queries)
3. Paraphrase generation. A decoding process that searches for the best paraphrasing result.

2.3 Dependency Based Paraphrasing

Non-continuous phrases usually play important roles in users' queries. For example, in the query *enhance the system's performance*, it is more conventional to say *improve performance*, but the phrasal paraphrasing methods cannot process this non-continuous situation. We propose a dependency based paraphrase method to solve this problem. The users' queries are first parsed into a dependency tree. Every arc (The head and its modifier) in the dependency tree is considered as a candidate for paraphrasing. Then the procedure of generating collocation suggestion can be performed on these word pairs. For example, the parsing result of the query *enhance the system's performance* is shown in Figure 1.

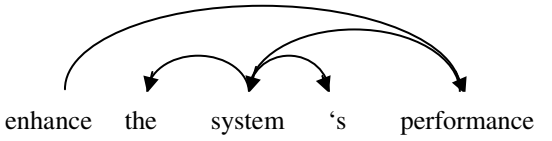


Fig. 1. Dependency tree for the query *enhance the system's performance*

In the dependency tree of Figure 1, the following arcs are candidates of paraphrase: (enhance, performance), (system, the), (performance, system), (system, 's). With the method same as generating collocation suggestion, these word pairs can be paraphrased: *enhance performance* ←→ *improve performance*, *the system* ←→ *a system*.

Note that it is safe to change the word of *enhance*, i.e. *enhance performance* → [X] *performance*. However, changing the word of *performance* (*enhance performance* → *enhance* [X]) is risky. Because the word *performance* has more than one linked arcs, changing of this word may damage other dependency nodes. Therefore, we make a limitation that only the word with no more than one linked arc can be substituted.

3 Experiments

3.1 Resources

To build a reliable monolingual corpus for example sentence retrieval, we collected 8,127 PDF files from ACL Anthology[1] which is a digital archive of research papers in the area of computational linguistics. After a series of processing steps, we extracted 1,716,418 high quality sentences.

The paraphrases were extracted from 129,938 bilingual sentence pairs which were collected from CNKI[2]. The language model used in paraphrase generation is a 4-gram model trained from BNC corpus. The dependency parser used for the method described in Section 2.3 is the Stanford Parser[3].

3.2 Results

Evaluation of Collocation Suggestion

We evaluate the quality of collocation suggestions by calculating the hits numbers in sentence retrieval system. For example, the phrase *enhance performance* gets 14 hits in the sentence retrieval system, while *improve performance* gets 539 hits. From this result we can see the second expression is more authentic than the first one.

From the query log of the system, we randomly select 400 two-word-length queries. The average hits number of users' queries (baseline) and collocation suggestions given by the system are listed in Table 1.

[1] http://www.aclweb.org/anthology/
[2] http://www.global.cnki.net/grid20/index.htm
[3] http://nlp.stanford.edu/software/lex-parser.shtml

Table 1. Average hits number of collocation suggestions

	Aver. Hits Num.
baseline	348.1
1st suggestion	623.1
all suggestions	201.4

Evaluation of Paraphrase Suggestion

We first manually evaluated the accuracy and coverage of the two paraphrase methods on 165 user queries. The results are shown in Table 2.

Table 2.. Accuracy and coverage rates of the phrasal method and the dependency based method

	Accuracy	Coverage
Phrasal method	76.5%	49.0%
Dep. Based method	81.2%	15.8%

The main reason of low-coverage problem for the dependency based method is because the dependency parser was trained on a corpus of different domain. The paraphrase substitution was conducted on a dependency structure, but the paraphrase resource used in the experiment was linearly extracted. Therefore, most long distance collocations obtained on the dependency structure cannot find matched paraphrases.

We use the standard IR measure, p@n to evaluate whether the paraphrase suggestion is more conventional than the original query. We collected 42 pairs of queries and the generated paraphrased suggestions as a test set. All the paraphrase suggestions in the test set are correct paraphrases judged by human. The dependency based method is not evaluated here because we could not collect enough paraphrase suggestions for the low coverage rate.

All the queries and paraphrase suggestions were sent to the sentence retrieval system. The human annotators counted the number of matched sentences in the top 1 (p@1) and top 5 (p@5) results. (Small morphological mismatch is allowed.) Table 3 gives the evaluation results of P@1 and P@5 for the queries (baseline) and their paraphrase suggestions.

Table 3. P@1 and P@5 evaluating results

	P@1	P@5
Baseline	0.3571	0.2810
Phrasal paraphrase	0.5952	0.3524

From Table 3 we can see that the precisions of the first and top 5 retrieval results for the suggestions generated by phrasal paraphrasing method are much higher than the baseline, indicating the generated similar expressions are more conventional than the users' queries.

4 Conclusion and Future Work

In this paper, we proposed a paraphrase based similar expression generation method. Using a statistical collocation model, a phrased based model and a dependency based model, paraphrases of high quality were generated for two-word queries, contiguous phrases and long-distance collocations, respectively. The experimental results showed that (1) the similar expressions given by the collocation method matched more than 79% example sentences in the retrieval system; (2) the improvement of phrasal paraphrase in P@1 and P@5 scores are 0.28 and 0.07.

Acknowledgement. This work was supported by National Natural Science Foundation of China (NSFC) (61073126, 61133012), 863 High Technology Program 2011AA01A207.

References

[1] Liu, Z., Wang, H., Wu, H., Li, S.: Collocation Extraction Using Monolingual Word Alignment Method. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, pp. 487–495 (2009)
[2] Wu, J.-C., Chang, Y.-C., Mitamura, T., Chang, J.S.: Automatic Collocation Suggestion in Academic Writing. In: Proceedings of the ACL Conference Short Papers, pp. 115–119 (2010)
[3] Stein, B., Potthast, M., Trenkmann, M.: Retrieving Customary Web Language to Assist Writers. In: Proceedings of the 32nd European Conference on Information Retrieval, pp. 631–635 (2010)
[4] Liu, T., Zhou, M., Gao, J., Xun, E., Huang, C.: PENS: A Machine-aided English Writing System for Chinese Users. In: Proceedings of ACL, pp. 529–536 (2000)
[5] Zhao, S., Lan, X., Liu, T., Li, S.: Application-driven Statistical Paraphrase Generation. In: Proceedings of ACL, pp. 834–842 (2009)
[6] McKeown, K.R., Radev, D.R.: Collocations. In: Dale, R., Moisl, H., Somers, H. (eds.) A Handbook of Natural Language Processing, pp. 507–523 (2000)

Parameters Allocation of Genetic Algorithm Based on RSSI of DSRC Communication

Chun-Liang Chen

Department of Electronic Engineering, Nan Kai University of Technology, Taiwan

Abstract. Dedicated Short Range Communication (DSRC) is a key enabling technology for the next generation of communication-based safety applications. Distance measurement based on RSSI, featuring low communication overhead and low complexity, is widely applied in the range-based localization of DSRC communication. However, the uncertainty factor and range error of shadowing model vary under different circumstances. Furthermore, the drawback of positioning error will affect its priority in geocasting or multi-hop of DSRC signals. This paper presents an operational scheme of parameters contribution design by utilizing Taguchi method, and the experiment inference is redesigned to compensate circumstance factors for shadowing model using genetic algorithm in the proving ground. The genetic algorithm provides an advanced solution to the circumstance factor and measured uncertainty using minimal absolute errors. These results may serve as benchmarks of parameters design for future DSRC channel communication.

Keywords: DSRC, Taguchi, delay filters, genetic algorithm.

1 Introduction

Since the inception in LAN standard, the wireless communications market has been growing rapidly, which provides the location-based services. In the low cost positioning research and development, the common approach is to establish small scale indoor environments positioning network [1]. Positioning methods mainly are divided into two categories: distance-based location and the algorithm without. Algorithm is based on the distance between nodes by measuring the distance and relative angle. The use of trilateral relation or triangulation is the maximum likelihood estimation algorithm. Ranging technology is used to receive signal strength indication (RSSI), time of arrival (TOA) and angle of arrival (AOA) [2]. The RSSI-based is the most popular method that requires less overhead and low implementation complexity in recent years.

However, the situation is more various in the radio propagation conditions. Typically the signal is transmitted through the multipath propagation to reach the receiver so that the multipath effect will make the received signal envelope and phase change in non-direct and direct signal components composed of a combination in terms of the relative emission signal distortion. As the environment becomes complex, it reflects a lot of signals in any selected communication path. The path loss is correspondingly different and leads to great difficulties. In order to overcome the

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 375–380.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

aforementioned problems, Taguchi's method used in the experiment design is proposed to find parameters effects.In this paper, the DSRC signal transmission parameters are optimized to meet specific environmental parameters using Taguchi analysis. The compensated model uses a genetic algorithm to provide optimal parameters.

2 System Architecture

2.1 System Platform

The research platform is built on both vehicles, and the test campus is located in general performance test track of ARTC proving ground as shown in Figure 1. This example is simulated in a metropolitan area by DSRC signals broadcasting. The test track is divided into several segments of 10 meters. To transmit DSRC signals with different packets speed, packet sizes and packet length, the commander plays a role and is provided with an embedded channel switch, packet slicing and power control programs. The receiver is used to receive packets from transmitter; and furthermore, the horn antenna is the key research tool for measuring environment power spectrum density. The measurement result is shown in antenna analyzer located in receiver vehicle. The DSRC receiver is used to check transmitting data and packet availability.

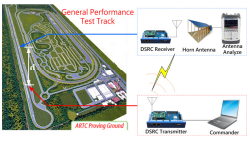


Fig. 1. Proving ground of system operation platform in ARTC

The DSRC device is the receiver or transmitter device located in vehicle. In addition to the hardware satisfying industry standard, the software can control packet TX/RX operations in different channels for different applications.

In DSRC communication, the priority usually is based on distance likelihood estimation. The DSRC device can be used to receive signal strength indication (RSSI) to mark a location in the RF signal strength [3]. In general, the RSSI will decay when the distance increases. This intensity decay, as the signal propagation loss, is related to the environment reflection. According to the received signal strength, by using the theoretical and empirical propagation loss model, transmission loss can be converted into the distance. Equation (1) is a logarithmic normal distribution of propagation loss model, named as shadowing model. The d0 is specified distance, usually 1m, and the first item of right part is receiver RSSI whose distance is located at d0. The left item is receiver RSSI whose distance is located at d. The n is path loss constant, and its value is dependent on environment and building. XdBm is masking factor, and the mean value is 0 and has low variance whose quality is a Gaussian distribution.

Under different circumstances, the values of XdBm and n are varying. The masking factor usually affects its power decay and measuring errors, and n has larger impact on path loss propagation. The following discussion will be focused on DSRC parameters configuration, and Taguchi method is the solution to find out crucial ratio distribution. The result will give a benchmark for parameters configuration of DSRC communication. From orthogonal array, the purpose of this paper provides optimized parameters learning using a genetic algorithm. The experimental plan is to determine all of the available cases under present parameters configuration. The demonstrated distance is divided into ten segments of 10 meters.

$$[\overline{p_r(d)}]_{dbm} = [\overline{p_r(d_0)}]_{dbm} - 10n\log(\frac{d}{d_0}) + X_{dbm} \tag{1}$$

2.2 Analysis of TAGUCHI Method

To clarify parameters distribution and meet better efficiency, Taguchi method is adopted in this study. The parameters configuration of DSRC communication has three factors including bit rates, power density, packet sizes and number of packets. Each factor also has three levels; these selected values are common applied and limited ability in DSRC application as shown in Table 1. As bit rate, the typical bit rate is least 3 Mbits. The standard of packet sizes is according to the maximum length of OSI network layer under 1024 bytes.

Due to four factors and three levels, the Latin squares uses L9(34) orthogonal array. The number of experiments is nine times with selected parameters design. In normal operation, the number of experiments should be done at least 81 (i.e. 3×3×3×3) times. But Taguchi method make large reduction with statistical distribution. Taguchi uses a number of progressive trials. To meet an optimal and quantitative design with signal to noise ratio, the orthogonal array is operated under a fixed distance. The orthogonal array L9(34) satisfies the selected factors and levels with low interaction. An initial experiment is often used to examine many factors in order to identify those that have a major effect. These 'control factors' are then used to predict a combination that will lead to optimal performance. If these results are satisfactory, further experiments are unnecessary.

Fig. 2 shows the hardware setup in system operation, and each transmitter or receiver uses two antennas to transmit or receive packets. The middle of Fig. 2 uses a horn antenna to do objective measurement, and the right side device is a spectrum analyzer used to measure the power density of RSSI. The measured result is shown in spectrum screen, and output data is processed for mean value with linear scaling integration method. A figure is presented as disorder level power spectrum in the right part of Fig. 2. A quality loss or mean square deviation (MSD) function is used to calculate the deviation between the experimental value and the desired value using equations (1)-(2). The type of MSD is nominal-the-best where yi represents the normalized quality loss using shadowing model. The analysis of variance (ANOVA) for DSRC packets is shown in Table 2, where it consists of freedom (f), sum of squares (S), sum of mean squares (V), partial sum of squares (S') and contribution (ρ). The analysis of experiments refers to the calculations that are done with results after experiments are carried out.

From Table 3 result, the main impacts are packet sizes and packet length under available transmitting condition. The error (e) is the total S/N values and subtracts of four factors. The integration error ((e)) is the summation of error and low SN. The optimal condition would be applied to parameters learning, and the combination is bit rate (12 Mbits), power density (15 dBm), packet size (1024 bytes) and number of packets (1000), where it is the integration result from ANOVA and S/N.

$$MSD = \sigma^2 = (\frac{1}{n})\sum_{i=1}^{n}(y_i - m) \tag{2}$$

Fig. 2. System allocation, horn antenna with analyzer and spectrum

Table 1. Factors and Its Levels

Control Factors	Levels		
	Level 1	Level 2	Level 3
A. Bit Rate(Mbits)	3	12	27
B. Power Density(dBm)	12	15	18
C. Packet Size(bytes)	64	256	1024
D. Number of Packets	1000	6000	10000

Table 2. Analysis of Variance

factors	f	S	V	S'	ρ%
A	2	29.16*	--	--	--
B	2	1.54*	--	--	--
C	2	165.58	82.79	153.28	58.2
D	2	66.99	33.49	54.69	20.8
e	1	0.03*	--	--	--
(e)	5	30.73	6.14	55.314	21
T	7	263.3	--	263.3	100

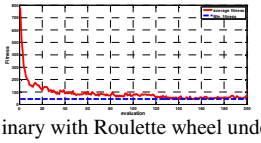
3 Verification of GA

From Taguchi method, these experiments are used to do parameters learning and test patterns. With parameters change, these test patterns are presented over all area cases with worse cases pattern. The following step is focused on parameter learning, and the parameters are calculated by minimal square error (MSE).The equation (1) is calculated by GA algorithm, and furthermore, the masking factor and path loss constant are learned from Taguchi experiments. The test result is shown in Table 3, and the table includes population size (Nps), crossover rate (Pcr), mutation rate (Np), size of parent (Pmr), path loss constant (n), masking factor (XdBm) and mean square error (MSE). The GA parameters are assigned by operator, and this index presents its performance and learning error.

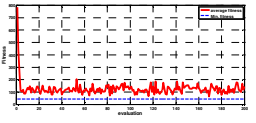
Table 3. Results of GA calculation

	Run	\ N_{ps}	P_{cr}	P_{mr}	N_p	Shadowing model n	X_{dBm}	MSE
Binary	1	10	0.8	0.01	100	1.9180	0.15	41.87
	2	5	0.8	0.01	200	1.9303	0.0	41.67
	3	10	0.9	0.05	100	1.9287	-0.06	41.71
	4	5	0.9	0.05	200	1.8870	0.70	42.36
Real	5	10	0.8	0.01	100	1.9297	0.01	41.68
	6	5	0.8	0.01	200	1.9302	-0.00	41.67
	7	10	0.9	0.05	100	1.9292	0.04	41.99
	8	5	0.9	0.05	200	1.9023	0.52	41.87

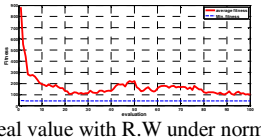
Fig. 3 shows MSE in different conditions.The GA and Taguchi method are applied separately in many investigations, but a combination of them applied in DSRC communication design is a new effort. The main considerations of the GA are choosing the proper population size, crossover rate, mutation rate, and the remaining size of the parent to search for optimal solutions most effectively. The MSI includes about thirty experiments, and the average error is only about 1 dBm. The accuracy of estimating distance is less than 4.6 meters. The experiment result shows the applicability of compensated modeling, and the result can calibrate model parameters.



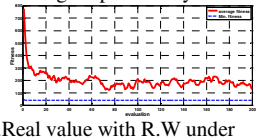
(a).Binary with Roulette wheel under normal probability



(b).Binary with Tournament under higher probability



(c).Real value with R.W under normal probability



(d).Real value with R.W under higher probability

Fig. 3. MSE summation under binary/real value calculation

4 Conclusion

Taguchi method based on the use of orthogonal arrays was used successfully to determine the optimal conditions necessary for DSRC communication, greatly increasing the availability with optimal parameters in transmitting and receiving packets. Using conventional strategies would have required an experiment of 81 separate reactions. However, the Taguchi method optimization was used to achieve the experiment of just nine reactions. In addition to optimizing product yield, the Taguchi method was used to examine the effects of specific components on data transmitting.

The GA search results may deviate greatly from the optimal solution. The power of the Taguchi experiments help locate the best GA parameters that give a search result which is the most close to the optimal solution and the least insensitive to the randomness of the search in this paper. These results will serve as benchmarks of parameters design for future DSRC channel communication.

References

[1] Benkic, K., Malajner, M., Planinsic, P., Cucej, Z.: Using RSSI Value for Distance Estimation in Wireless Sensor Networks Based on ZigBee. In: Proceedings of 15th International Conference on Systems, IWSSIP (2008)
[2] Fang, Z.: RSSI Variability Characterization and Calibration Method in Wireless Sensor Network. In: International Conference on Information and Automation (2010)
[3] Arai, M., Kawamura, H., Suzuki, K.: Estimation of ZigBee's RSSI fluctuated by Crowd Behavior in Indoor Space. In: Proceedings of SICE Annual Conference (2010)

Identification of FOG Multivariable Model Based on Fuzzy Relation

Zhao Xijing, Wang Lixin, He Zhikun, and Li Rui

Xi'an Research Institute of High-Tech, Xi'an, Shaanxi 710025, P.R. China
zhaojingjing05@sina.com

Abstract. In view of the complicated nonlinear characteristic of fiber optic gyroscope (FOG) temperature drift, a new fuzzy approach is proposed to build the input-output model. Based on fuzzy logic theory, fuzzy regulation is formed according to the input and output information of FOG. Then the model of FOG is identified by corresponding fuzzy inference. The simulation result shows that the model presented is effective and the FOG temperature drift is reduced greatly. The identification method may provide a new approach to the temperature drift compensation of FOG on line.

Keywords: FOG, Fuzzy Relation, Nonlinearity Analysis, Multivariable Model, Model Identification, Temperature Compensation.

1 Introduction

Based on Sagnac effect, FOG is an inertial instrument, which made by fiber can be used to measure the rotation angular velocity of objects relative to inertial space. As FOG has a lot of advantage such as small volume, low cost, wide dynamic range and quick start-up, it has been widely used in vehicle navigation and guidance, satellite positioning and earth measurement. The drift of FOG caused by environmental temperature is the main factor affecting its measure precision. The FOG technology has developed rapidly and the research has made great progress in recent years[1-3]. So research in the temperature characteristic of FOG has profound significance for improving the measure accuracy through modeling and compensation.

Up to the present, there are many traditionary methods in modeling FOG drift, such as multinomial fitting, neural network, wavelet theory, fuzzy identification etc[4-6]. Fuzzy identification theory is adapt to model for the complex system due to the abundant experience and direct expression. The method for the FOG temperature drift model based on fuzzy relation is analyzed and designed.

In this paper, a new approach for FOG input-output model based on fuzzy relation is proposed. The rest of this paper is organized as follows. In section 2, researching purpose is provided. Fuzzy reference theory is introduced in section 3. Then, the input-output model of FOG is identified based on the method of fuzzy inference in section 4. Finally, the significative conclusion is given in section 5.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 381–387.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Description of the Problem

The fuzzy theory as the foundation of fuzzy identification adapt to resolve almost all the problems in nonlinear modeling. Fuzzy identification based on fuzzy set theory may describe the characteristic of complex system by modeling with the input and output information of the system. It is a logic method by language express, so it have robustness and stability in accurate mathematic modeling[7-9]. The fuzzy regulation was formed according to fuzzy information, and the character was recognized by fuzzy inference.

The standard of evaluating the model quality based on fuzzy relation include:

(1) Fuzzy modeling form regulation that can describe the system character according to input and output measure information. The amount of regulations is the symbol of complex degree of fuzzy algorithm. The regulation is less, the algorithm is easier; the regulation is more, then the algorithm is more complex.

(2) Choose the average variance p that is the difference between measure figure $y(t)$ and predicted output figure $\hat{y}(t)$ as the accurate index of fuzzy model.

$$p = \frac{1}{L}\sum_{i=1}^{L}[y(t) - \hat{y}(t)]^2 \tag{1}$$

where L is the measure times.

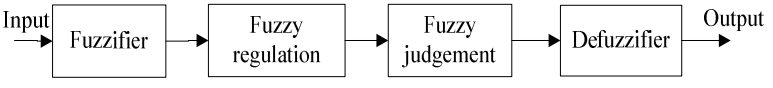


Fig. 1. Block diagram of fuzzy inference

The diagram of fuzzy system theory is shown in Fig.1. Firstly, the accurate measurements turn into fuzzy measurements by fuzzification. The fuzzy regulations are saved into computer after mathematic operation. Then, the computer makes decision by fuzzy inference according to the fuzzy regulations. At last, the fuzzy set of output turn into accurate measurements by defuzzification to drive the executive machine.

3 Model Identification for FOG

The main output drift of FOG include the error caused by the structure state, environment and so on. The relationship between input and output of FOG is nonlinear and complicated. Design the model with two-input and one-output, choose the temperature and FOG rotational speed as the input and choose the output of FOG as the output of the fuzzy system. Measurement grade should choose according to the model precision because of the relativity[10-12].

If x is the accurate variable among $[a,b]$, y is the fuzzy variable among $[-n,+n]$,where n is a positive integer. That sketch map is shown in Fig.2.

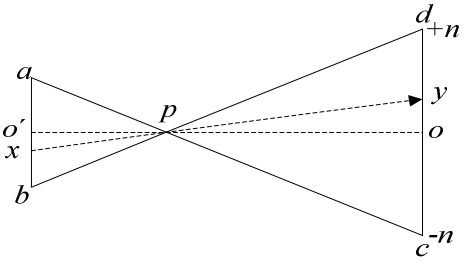


Fig. 2. Diagram of accurate variable to fuzzy variable

As $\Delta xo'p \sim \Delta yop$, $\Delta abp \sim \Delta cdp$, so the conversion relation is

$$y = 2n[x-(a+b)/2]/(b-a) \tag{2}$$

When fuzzy variable region isn't symmetrical, the conversion relation is

$$y = (m+n)[x-(a+b)/2]/(b-a) \tag{3}$$

The membership function $\mu(x)$ of fuzzy sets is found by

$$\mu(x) = \exp[-\frac{(x-\tau)^2}{\sigma^2}] \tag{4}$$

Where τ and σ may figure out by least square method.

Choose the identification regulation as "*if* A *and* B *then* C", then fuzzy relation is

$$R = (A \times C) \cdot (B \times C) \tag{5}$$

Conversion is

$$R = A \times B \times C \tag{6}$$

Where A denotes fuzzy temperature input, B denotes fuzzy rotational speed, C denotes the identification output. R denotes the relation of fuzzy model. According to illation,

$$C = (A \times B) \circ R \tag{7}$$

Design the temperature model based on fuzzy inference. Let the temperature inside the FOG and angular rate of carrier as the system input, and the output voltage of the FOG as the system output. The structure of the model is shown in Fig.3.

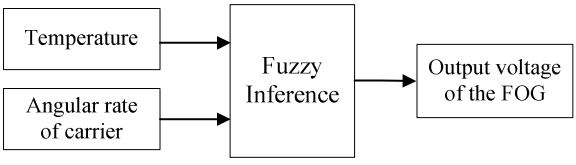


Fig. 3. Structure of temperature model based on fuzzy inference

Choose the region of temperature is -30～+60℃ and the region of temperature variety rate is -80～+80°/s. According to formula (5) and (6), the fuzzy output C can be identified. So, the actual output can be received. Change the fuzzy set into accurate figure by weighted illation. Translate the conclusion of $if \;\; A \;\; and \; B \;\; then \; C$ into

$$if \quad A_i \quad and \quad B_i \quad then \quad C_i \rightarrow w_i / z_i$$

Where $i \in [1, n]$, w_i / z_i denotes a element of the fuzzy set, w_i denotes of fuzzy regulation. The aptness between $x_0 \quad and \quad y_0$ and $A_i \quad and \quad B_i$ is received according to

$$h_i = \mu_{A_i}(x_0) \cdot \mu_{\mathbf{B}_i}(y_0) \tag{8}$$

Thus, the conclusion is

$$z_0 = \frac{h_1 w_1 z_1 + h_2 w_2 z_2 + \cdots + h_n w_n z_n}{h_1 w_1 + h_2 w_2 + \cdots + h_n w_n} \tag{9}$$

4 Simulation Experiments

The certain type FOG was sampled and investigated in order to verify the effectiveness of the proposed Fuzzy method to model the FOG drift. The date can be obtained by the experiment at different temperature and in different rotational speed. The process of operation is shown in the following:

(1) The FOG was placed in a temperature control unit on a platform and adjusted temperature according the requirement.
(2) Adjusted the rotational speed as the required order and collect the output date.

Considering the large deviation occurred in the drift, the data for analysis should be collected in stable working condition. The sampling time at every temperature or rotational speed point was two minutes. Let the temperature grade is -30℃, -20℃, -10℃, -5℃, 0℃, 5℃, 10℃, 20℃, 40℃, 60℃, Let the rotational speed grade is -80, -40, -20, -10, -5, -2, -1, 1, 2, 5, 10, 20, 40, 80°/h. The original output date curve of the FOG is shown in Fig. 4. The output of identical rotational speed and different temperature is also shown in Fig. 5. And the output of identical temperature and different rotational speed is also shown in Fig. 6.

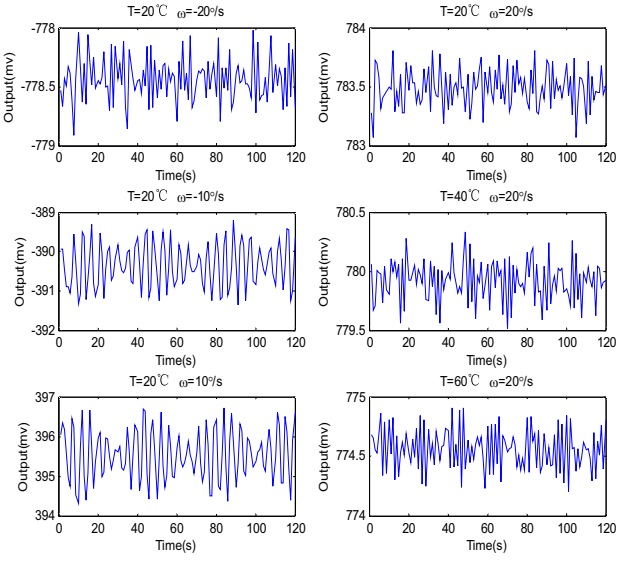


Fig. 4. The part of original data curve

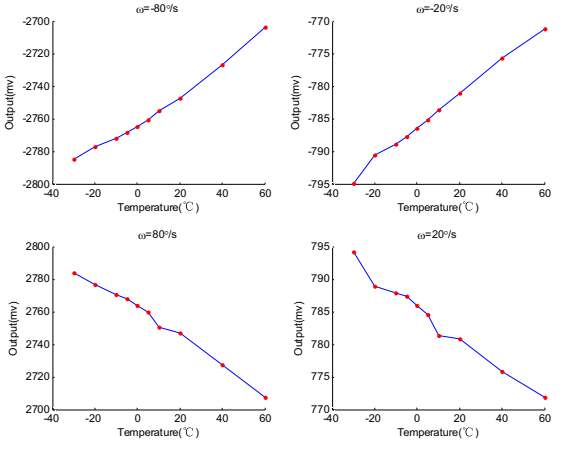


Fig. 5. The identical rotational speed output at different temperature

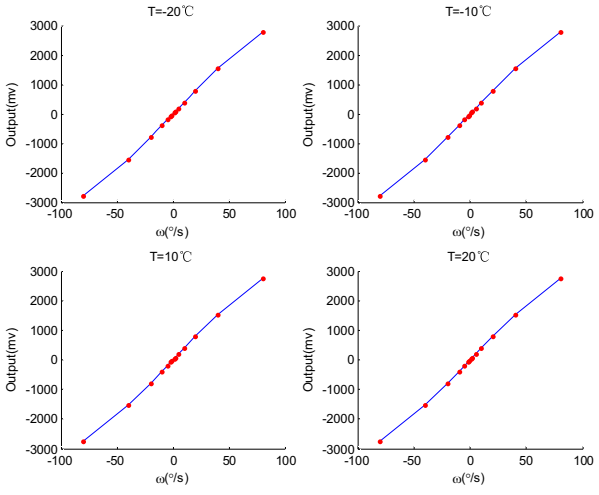


Fig. 6. The identical temperature output at different rotational speed

From Fig. 4 to Fig. 6, it is shown that the nonlinear characteristic of FOG drift against temperature is obvious. In this paper, a new fuzzy method is investigated to the input-output model. The identification result is shown in Fig. 7.

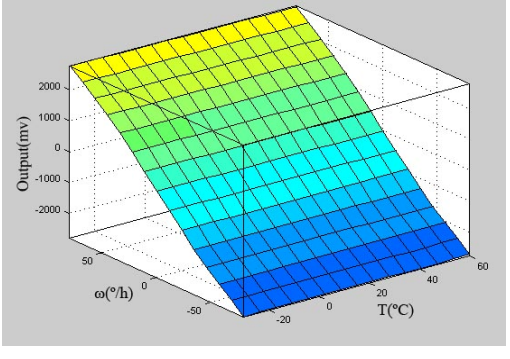


Fig. 7. Output of fuzzy inference system

5 Conclusion

In this paper, the proposed identification method had been tested according to the model of FOG. The result of fuzzy reference method coincides with theoretical analysis, which indicates the method is effectively to build input-output model of FOG. It is obviously that the research can promote the measure accuracy of FOG. Furthermore, the model can be used to compensate for the drift of FOG. The fuzzy algorithm can meet the identification requirement for FOG. In further studies, expanding experiment to high precision FOG will be investigated.

References

[1] Savage, P.G.: Strapdown inertial navigation integration algorithm design. Part 2. velocity and position algorithms. Journal of Guidance, Control and Dynamics 21(2), 208–221 (1998)
[2] Li, M., Zhang, C.X., Song, N.F., et al.: Online RWE monitoring based on cross-correlation method for IFOGs in space applications. Chinese Journal of Sensors and Actuators 23(3), 383–387 (2010)
[3] Meng, Z.K., Cui, J.T., Zhang, B., et al.: Experiment research of the high precision optical fiber gyroscope temperature. Journal of Astronautics 28(3), 580–583 (2007)
[4] Wang, X.L., Ma, S.: Applicability compensation method for random drift of fiber optic gyroscopes. Journal of Beijing University of Aeronautics and Astronautics 34(6), 681–685 (2008)
[5] Han, B., Lin, Y.R., Deng, Z.L.: Overview on modeling and compensation of FOG temperature drift. Journal of Chinese Inertial Technology 17(2), 218–224 (2009)
[6] Li, X.Y., Zhang, N., Gao, W., et al.: Application of wavelet analysis in testing of dynamic characteristics of fiber optic gyroscope. In: Symposium Gyro Technology 2009, Karlsruhe, Germany (2009)
[7] Huang, C.P.: Stability analysis of discrete singular fuzzy systems. Fuzzy Set and Systems 151, 155–165 (2005)
[8] Gao, H.J., Chen, T.W.: Stabilization of nonlinear systems under variable sampling: a fuzzy control approach. IEEE Trans. Fuzzy Systems 15(5), 972–983 (2007)
[9] Li, S.Y.: Fuzzy control neurocontrol and intelligent cybernetics. Harbin institute of technology press, Harbin (1998)
[10] Li, T.H.S., Tsai, S.H.: T-S fuzzy bilinear model and fuzzy controller design for a class of nonlinear systems. IEEE Trans. Fuzzy Systems 15(3), 494–505 (2007)
[11] Fang, J.C., Zhang, X., Li, J.L.: A compensation method for MEMS gyro scale factor error. Acta Aeronautica et Astronautica Sinica 31(2), 350–355 (2010)
[12] Jin, J., Li, M., Zhang, Z.G., et al.: Analysis of temperature errors in digital closed-loop fiber optic gyroscope. Infrared and Laser Engineering 37(3), 521–524 (2008)

An Improved Ant Colony Algorithm Combined with Genetic Algorithm and Its Application in Image Segmentation

Zhou Haifeng

The City Vocational College of Jiangsu

Abstract. This article applies the improved ant colony algorithm to the fuzzy c-means clustering, which overcomes sensitivity to initialization of fuzzy clustering method(FCM) .This article improves the shortcomings which the traditional genetic algorithm and the ant colony algorithm work step-by-step, makes the mix algorithm work in the entire cluster's process, simultaneously, puts the a swarm degree function in the ant colony algorithm, enhanced the ant algorithm search of the overall situation, increase the algorithm traversal the optimization capacity.

Keywords: Ant Colony Algorithm, Genetic Algorithm, Image Segmentation.

1 Introduction

The fuzzy C-means clustering algorithm is on the premise of the deterring initial cluster centers, causes the sum distance from various patterns to the assess category center is smallest. The algorithm for the classification result is influenced by the numbers and cluster centers of the initial location; therefore the result is only local optimal [1].

In this paper, the algorithm integrates the genetic algorithm with the ant colony algorithm, simultaneously, puts the swarm degree function into the ant colony algorithm, and increases the ant colony algorithm traversal optimization ability. Unify this mix algorithm and the FCM algorithm carry on the fuzzy clustering. On the one hand, the ant colony algorithm's robustness can effectively overcome the FCM algorithm to the sensitive of the initialize; On the other hand, the parallel distributed computing can accelerate the convergence and improve the efficiency of clustering. The most important thing is the algorithm's adaptive and intelligent search features can be achieved global optimal. Use this mix algorithm solves the dynamic cluster centers and the numbers, then applies this mix algorithm in the image segmentation, the result shows this mix algorithm can accurately image segmentation.

2 FCM Based on the Improved Ant-Genttic Algorithm

This algorithm as follows:

2.1 Basic Thought

Consider the data will be clustered as ants with different attributes, each ant has N - dimensional vector, the cluster center as the "food source", the definite cluster center's

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 389–393.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

process is the ant with different characteristic search food source from the ant nest, when searching, different ants choose some data elements is mutually independent.

2.2 Algorithm Step

(1) Suppose each data eigenvector is:

$$X = \{X_k \mid k = 1, 2, \cdots, n\}$$
$$X_k = \{x_{k1}, x_{k2}, \cdots, x_{km}\}$$

With n input samples, each sample has m characteristics. N denote iterative numbers, r denote cluster radius, ε_0 denote statistical error, each pheromone $\tau_{ki}(0) = 0$, V denote the initial center; $d_{ki} \|x_k - v_i\|$ denote the distance between kth sample and ith class. The path's information is given by equation below:

$$\begin{aligned} &if && d_{ij} < r, \tau_{ki}(t) = 1; \\ &otherwise && d_{ij} > r, \tau_{ki}(t) = 0. \end{aligned}$$

(2) $p_{ki}(t)$ denote the probability of ant k transfer from x_k to v_i at the t moment

$$p_{ki}(t) = \frac{\tau_{ki}^{\alpha}(t)\eta_{ki}^{\beta}(t)}{\sum_{s \in S} \tau_{ki}^{\alpha}(t)\eta_{ki}^{\beta}(t)}$$

$$S = \{X_s \mid d_{ks} \geq r, s = 1, 2, \cdots, i, i+1, \cdots, N\}$$

And, S denote a sample set which ant k chooses in next step.

(3) Compute the path crowd degree $q_{ki}(t)$

$$q_{ki}(t) = 2\tau_{ki}(t) \Big/ \sum_{i=k} \tau_{ki}(t)$$

In the early, the crowded threshold value is closed to zero, so the majority of ants can be independently chosen path at random, if a certain path with high pheromone degree, but also have many ants on this path, then the next ant will not choose this path, but choose at random, so this can avoid the ant premature concentrating in some local minimum.

(4) Updating local pheromone

$$\tau_{ki}(t+1) = \rho\tau_{ki}(t) + \Delta\tau_{ki}(t, t+1)$$

(5) Intercross

The intercross operation is carries between the most optimal ants of iterative and other random selection ants, or by probability group after finish a cycle, the two randomly

selected among the ants. You can also use probability p_c (cross-probability) from the entire ant colony random selection of a certain number of ants and the most optimal ants in iterative.

(6) Updating global pheromone

With the movement of ants, the pheromone changes, after one cycle, each path pheromone adjustment according to the equation below:

$$\tau_{ki}(t) = \rho.\tau_{ki}(t)(t-\Delta t) + \Delta\tau_{ki}$$
$$\Delta\tau_{ki} = Q/d_{ki}$$

Q is a constant, denote the sum pheromone of ants released on search path . Then calculates the new cluster center and the distance between sample point to this new cluster center.

(7) Path decision-making

If $p_{ki}(t) \geq p_0$,then x_k belongs to v_i

$$m_i = \left\{ X_j \middle| d_{ji} \leq r, j = 1, 2, \cdots, I \right\}$$

And, m_i m denote the data set belong to v_i ; Otherwise, compute the next sample weighted Euclidean distance .

$$m_i = \frac{1}{I}\sum_{j=1}^{I} X_j, X_j \in m_i$$

(8) End condition

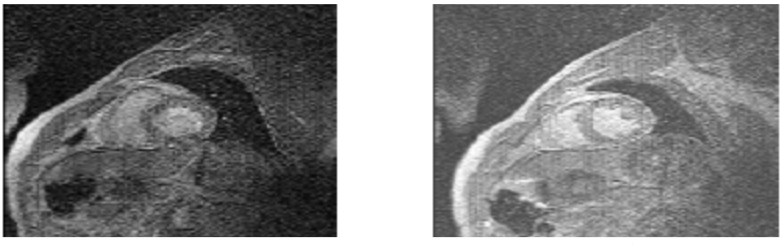
Compute $D_i = \sum_{j=1}^{I}\left[\sum_{k=1}^{m}\left(x_{jk} - v_{ik}\right)^2\right]^{1/2}$

Compute the bias error and overall error of cluster i , calculate the overall error $\varepsilon = \sum_{I=1}^{j} D_i$,

if $\varepsilon \leq \varepsilon_0$, output the cluster number c ,otherwise, keep iterativing.

3 Experimental Result

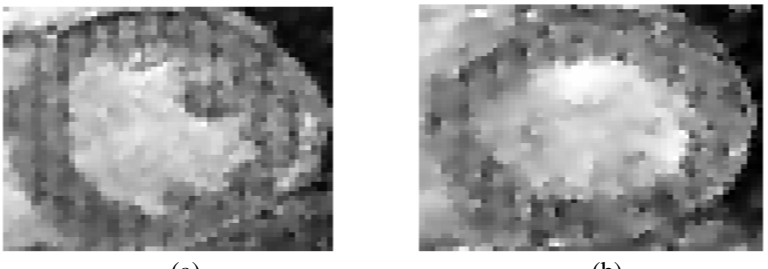
In order to verify this mix algorithm's effective and practical, take the left ventricle in dilatation and constringency image as an example, were using traditional FCM and the mixed method of this article proposed to study the image segmentation. Figure 1 (a) is the left ventricle is in dilatation, (b) is the left ventricle is in constringency.



(a) (b)

Fig. 1. (a) Left Ventricle is in dilatation, (b) Left Ventricle in constringency

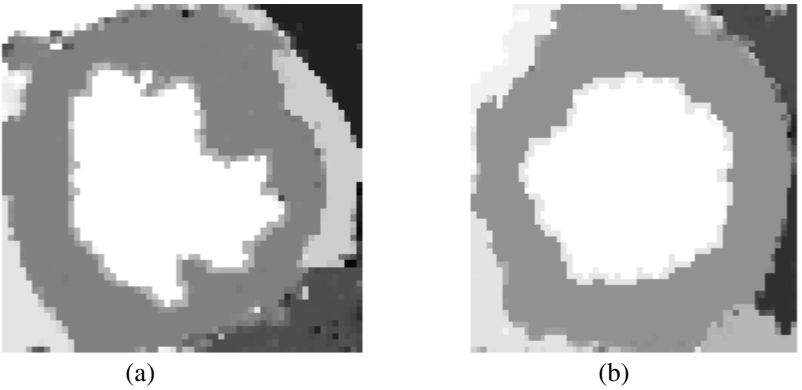
Figure 2 shows the result using the tradition FCM, we can see in the divided border region rather ambiguous, it is the local FCM clustering algorithm arising from the lack of convergence.



(a) (b)

Fig. 2. (a) Left Ventricle segmentation is in dilatation, (b) Left Ventricle segmentation in constringency

Figure 3 shows the result use this article's mix algorithm; we can see the division image contour which obtains using this article mix algorithm to be quite clear, the segmentation result is very obvious.



(a) (b)

Fig. 3. (a) Left Ventricle segmentation is in dilatation, (b) Left Ventricle segmentation in constringency

4 Conclusion

This algorithm can increase the ant colony algorithm's feedback mechanism, to avoid premature, can improve the accuracy of clustering. Application of the algorithm proposed to image segmentation and comparative experiments show that the algorithm has great ability of detection the fuzzy edge and exiguous edge.

References

[1] Bian, Z.: Pattern Recognition. Tsinghua University Press, Beijing (2000)
[2] Wu, Q., Wang, L.: Intelligent ant algorithm and application. Shanghai Science and Technology Education Press, Shanghai (2004)
[3] Huang, G., Wang, X., Cao, X.: Ant colony optimization algorithm based on directional pheromone diffusion. Chinese Journal of Electronics 15(3), 447–450 (2006)
[4] Yang, L., Zhao, L., Wu, X.: Medical image segmentation of fuzzy c-means clustering based on the ant colony algorithm. Shandong University Journal (technology version) 37(3), 51–54 (2007)
[5] Kamel, S.M.: New algorithms for solving the fuzzy C-means clustering problem. Pattern Recognition 27, 421 (1994)
[6] Wu, L., Yang, D.: Portrait background segmentation based on improved fuzzy C-Means Clulstering. Computer Application 26(2), 424–428 (2006)
[7] Ren, C., Zhang, J.: Robot path planning based on improved ant colony optimization. Computer Engine 34(15), 30–35 (2008)
[8] Xiu, C., Zhang, Y.: Hybrid optimization algorithm based on ant colony and fishi school. Computer Engine 34(14), 206–207 (2008)

Paper's Indirect Strain Control Bases on Inverter

Tianli Li

College of Electrical and Information Engineering, Shaanxi University of Science and Technology, China

Abstract. Because of the direct tension control have disadvantages of high investment, a indirect tensile control based on the inverter sliding compensation is putted forward, the inverter's fine-tuning is changed by detected motor torque.It is the micro adjusted out of speed hain ,so as to achieve the purpose of adjust tension.The principle of indirect tension control is analysed,its realize means,control arithmetic and applied effect are presented.

Keywords: inverter, slip compensation, indirect strain control, control algolithm.

1 Introduction

Most paper machine have coating on line,paper machine is divided into two parts: it is former drying before coating machiner, after drying after coating machine. Generally speaking, paper after drying by defining the moisture content is close to finished product, it is so smoothly paper moisture to coat.This kind of dry degree adjustable rate has been very small with paper, it has lost paper more flexible characteristics. In this case,it will cause the great changes paper tension,if the tiny error appear of speed ratio between the last group advance drying cylinder and coating machine[1] ,it is extremely easy to make paper produce broken page.How to solve this problem,The most common method is to use direct tension control:the tension sensor is installed between the last a cylinder that turbulence department after the ministry of turbulence and the last one cylinder leading paper roll paper,this is direct tension control,But the investment of tension sensor is very big, and it increases the difficulty of maintenance[2].We adopt indirect tension control, do not adopt direct tension control:the mechanical properties will become soft,Use of the converter sliding compensation,the paper elasticity can be compensated through electrical elasticity,it get better result[3].

2 The Principle of Sliding Compensation

2.1 The Tension Analysis of Adjacent Two Points

Paper tension is stretched force between two paper points, namely the elastic force[4]. Setting A, B two sheets of paper path length is L, natural newspeak L_0, by drawing in

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 395–399.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

length for L = L_0 + Δ L,setting: inside unit time t , the speed of the paper through A point is V_A, B is V_B.

According to the definition of tension:

$$F_{AB} = \frac{\Delta L \times K}{L_0} = K \frac{V_B \times t - V_A \times t}{V_A \times t} = K \frac{V_B - V_A}{V_A} = K(K_1 - 1) \quad (1)$$

In the formula:K—constant due to material relevant ; K_1—B、A two speed ratio $V_B = K_1 V_A$; The above formula instructions,In the running process of the paper,the tension through the two points is proportional with speed rate, isnot relevant with speed.When K_1=1, ension F_{AB}=0, each point should maintain this ratio:When K1 > 1 ,tension F_{AB}>0 , paper is tauted,then later a transmission point load increases, the output current increase; When K1 < 1 ,tension F_{AB}<0, paper is stacked, then later a transmission point load decreases, the output current decrease[5].

2.2 The Principle of Sliding Compensation

Setting the rated speed of asynchronous motors is n_e, Synchronous speed is n_0 ,then rated slip is:

$$s_e = \frac{n_0 - n_e}{n_0} \quad (2)$$

If the setting parameters ange of turning difference compensation for a certain brand of the converter is percentiler ,compensation frequency value may be estimated under its rated load[6]. As for 100%, a compensation parameters, when the load is 100%,compensation frequency is 50s_e the output of this inverter automatically jump to 50+50s_e from 50H_Z;similarly, As for -100%,the output of this inverter automatically jump to 50-50$s_e$from 50;Its mechanical characteristics shown in figure 1,piont B is not compensation, point A of compensation parameters is 100%,point C is compensation parameters for - 100%.

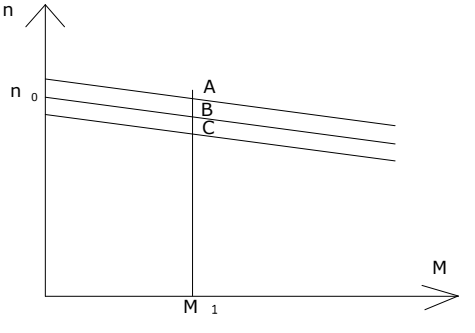


Fig. 1. Mechanical properties of sliding compensation

From the above analysis ,we know inverter with turn difference compensation is calculated for a given torque changes frequency fine-tuning, it is a kind of speed chain outside of a given frequency micro adjust, so in fact affected, and after the level of tension adjust speed ratio of purpose.

3 Indirect Tension Control of Sliding Compensation

3.1 The Apply of Sliding Compensation for Paper Tension Control

It is knowable through 2.1 analysis, if tension of sizing machine has changed, it canaffect motor load of sizing machine. WhenF_{AB}<0 ,it should belong to speed chain problem to be solved,When F_{AB}>0, paper is tauted, sizing machine transmission point motor load increases, the current increases, At this time, the speed of glue machine should be kept as constant,thus it can fall to change the two ratio, to prevent from relief tension.known from 2.2 analysis, sliding compensationshould be set into negative compensation[7].

3.2 Design of Control Algorithm

3.2.1 The Actual Factors of Affect Compensation Effect

① Theoretically,slip compensation parameters according to the motor determines the converter made loads an analog loop automatic control,Although it make the external mechanical characteristic motor changes.But essentially it is a kind of utomatic compensation results of invertor,when motor speed or torque changes, inverter will participate adjust and control[8].Practice because of data on current compensation calculated parameters and motor,Such as resistors and inductance, etc,and these parameters especially resistance changing with temperature resistance changes greatly.Such sliding compensation parameters should be adjusted dynamically according to the actual working condition , should not be a fixed value.

② Seeing from drive requirements for paper machine,。 Speed steady precision should be within 0.01%.Set a asynchronous motor parameters as follows:: n_e=1450r.p.m,p=2,For ±0.01H_Z frequency changes,Speed change for approximate is:Δn=0.01f×60/p=15r.p.m,This value has reached the rated speed is about 1%,Therefore sliding compensation value in this scope shall be compensated, otherwise it will cause, have had the opposite effect.The rated slip for taking 1/4 or less lcan satisfy,according to the theory and practical situation.

3.2.2 Control Algorithm

The control algorithm can be applied by the above analysisand designed: when the equipment is normal (FAB = 0), the output current of the converter of sizing machine is A_r, and the actual operation current is A_f , sliding gain is S_G:then[9]:

① IF $A_f \geq K_{11} A_r$ THEN $S_G = -K_{21} \times S_e\%$

② IF $K_{11}A_r < A_f \leq K_{12} A_r$ THEN $S_G = -K_{22} \times S_e\%$

③ IF $K_{12}\ A_r < A_f \leq K_{13}\ A_r$ THEN $S_G = -K_{23} \times S_e\%$

④ IF $A_f > K_{13}\ A_r$ THEN $S_G = -K_{24} \times S_e\%$

Among them:

S_e---Rated slip;Because of some of converters sliding compensation value are given by frequency value directly, so in the equation of the S_G have two kinds of forms.

K_{11}、K_{12}、K_{13}、K_{14}---there are decided byo the converter brand, motor parameters, factors such as paper characteristics , in 0.050 - 0.095 commonly;

K_{21}、K_{22}、K_{23}、K_{24}---there are decided byo the characteristics of converter brand, paper factors, generally at 0.25 - 0.98 value;When the compensation value is given in the form of frequency directly, it isconsidered ompensation frequency range in coefficient calculation and setting

3.3 Realization of Sliding Compensation Indirect Tension Control

3.3.1 Control Principle

V_r ---the speed from speed chain;

A_r ---motor current namely reference current;

A_f --- motor practical running current. It is compared with actual current and reference current the difference, the inverter sliding compensation value will be determined.

3.3.2 How to Get Reference Current and Actual Current

It is known from motor theory,after the confirmation for the synchronous speed, motor's torque is the definite relationshipactual with motor's actual speed .And torque and current is so certain relations,So the relation of the speed and the current motor practical corresponding can determine.If $F_{AB}=0$,speed of the given value brought by the speed chain is transformed into its current according to above analysis ,this value is a reference current.We can obtain multiple speed under the given value chain multiple reference current , according to this method, these value will exist in programmable controller.In practice, how to calculate reference current is more complex[10],This calculation can influence real-timeof control ,So, method of getting reference current rely on calculation and commissioning , is not a single rely on calculation.Motor actual current is obtained by communications PLC and inverter.

3.3.3 Realization of Indirect Tensile Control

Indirect tension control is finished by the paper core device - PLC ,output current of the converter of sizing machine collected by PLC,the error range is determined by compare to this value and normal operation, the current A_r ,sliding compensation parameters setting value is findied out according to above control algorithm and is sended to invertor by communication between the PLC and inverter,Thus reach the purpose of tension control.

4 Actual Usage

1940/450multi-cylinder paper machine of long nets manufactures advanced culture paper,Electrical drive adopts EMERSONTD3000 inverter, sizing machine adopts 11KW motor drive.Begining,Don't use inverter sliding compensation, F403=0, F404=0, Two days after the initial runing (24 hours a day),The continuous production to emerge seven paper break between six groups of cylinder and glue ,Average every eight hours 1 times,these bring bigger loss for factory.Later we use slip compensation of invertor,Specific parameters such as:When A host frequency is40.01H_Z. for as 16.14A, now A_r=16.14A,then K_{11}=0.050, K_{12}=0.065, K_{13}=0.075, K_{14}=0.090, K_{21}=0.31, K_{22}=0.31, K_{23}=0.60, K_{24}=0.90, S_G=0.00H_Z~0.03H_Z。 After that the operation of for 312days only appear once paper break,this paper machines is running in good condition, the manufacturer satisfaction.

5 Conclusion

The indirect tensile control method based on sliding compensation has been used in several paper factory,get better result,to look from running situation,there arenot disconnec for paper caused by tension problem .

References

[1] Chuang, C.-W., Chiang, C.-L., Kao, C.-C., Fung, R.: Applying Modified MIMO DIVSC to synchronism and Tesion Control of Dual Motor System. Journal of Dynamic System Measurement and Control 131, 054501 (2009)
[2] Zhao, H.-Y., Rahn, C.D.: Iterative Learning Velocity and Tension Control for Single Span Axially Moving Materials. Journal of Dynamic System Measurement and Control 130, 051003 (2008)
[3] Li, M., Gao, D.-B., Ge, B., Yang, T.: Dual-servo warper of constant tension control. Machinery Design and Manufanture, 127–129 (December 2009)
[4] Martin, J.-D., Laurendeau, D., Gosselin, C.: Extremum Seeking Tuning for Reel Tension Control in Haptic Application. In: AIP Conf. Proc., vol. 1107, p. 275 (2009)
[5] Zhao, Y.-G., Qin, L.-G.: Application of siemens MM440convertor in Reel packaging machine. China Pulp and Paper 29(8), 53–55 (2010)
[6] Zhang, C., Yang, L.: Realization of tension control of paper machine based on PLCand VFD. China Pulp and Paper 29(9), 53–55 (2010)
[7] Shi, Y.-Y., Tang, H., Yu, Q.: Key technology of the NC tape winding machine. Acta Aeronanticaet Astronautica Sinica 29(1), 233–239 (2008)
[8] Shi, Y.-Y., Yan, L., He, X.-D.: Research on control technology for discontionous tape winding. Acta Aeronanticaet Astronautica Sinica 31(6), 1294–1298 (2010)
[9] An, L.-X., Liu, Y.-G., Cui, L.-J., Ma, H.-M.: Margincs new method of uncoier torque control. Electric Deive 40(3), 66–68 (2010)
[10] Zhang, D.-Z., Chen, Y.-L., Jiing, F.-W., Lin, S., Su, W.-J., Cao, P.-L.: Hot strip mill coil tension control system based on siemens T400 technology madule. Iron and Steel 44(3), 56–61 (2009)

Study on the Online Clustering Algorithm Based on Grid Structure

Wu Peng, Jia Jun, and Zhu Tianxiang

Shenyang University of Technology School of Software, Shenyang 110023, China

Abstract. Clustering is an important task in data mining and knowledge discovery, which groups objects into meaningful subclasses. As the most existing stream clustering, algorithms can not generate online clustering results in real-time, an online data stream clustering algorithm is proposed by using sliding windows and density-based grid storage structure. The algorithm achieves a rapid speed for online clustering data stream and it can provide users with real-time clustering results and reflect the dynamic evolution of data streams. Experimental results show that the algorithm proposed has a good capacity of dealing with rapid evolutional data stream and have a good clustering quality.

Keywords: data mining, data stream, online clustering.

1 Introduction

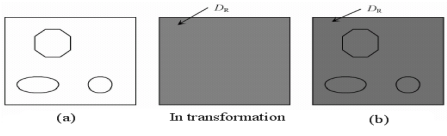
In the field of data stream system, the cluster varies spontaneously at different time[1]. Thus, how to manage the data stream and obtain valuable information effectively, becomes to be of great significance for related application nowadays. Recently, more and more research have put focus on the subject of data stream clustering analysis[2][3]. For instance, Aggarwal et al suggest the clustream concept, meaning solve the varied issues on data stream clustering; Guha et apply the “seperation theme” into related fields, all of which play an important role in the clustering analysis. However, the approaches are insufficient on the cluster with random shapes. Besides, a common problem in all the above algorithm is that the produced grid is static, its dynamic behavior can not change with time variation. Therefore, a static grid is not adequate for providing accurate predictions throughout a long time horizon. The theory suggests the D-Stream method based on the online-offline clustering frame, solving effectively the current problems metioned above, in which believes that three fundamental requirements are essential: compressed description; effective treatment on novel data points; fast judgment on special data points, respectively.

In the paper, an overview of the proposed adaptive grid structure is given first, followed by a discussion on the clustering algorithm of the input space and the determination of the neighbor grid. Then the paper concludes by highlighting the advantages of the proposed approach. Finally, the paper gives out a novel approach on data stream analysis, which is called effective online clustering algorithm.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 401–405.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 The Density of Grid Structure

Clusters are dense regions of objects in the data space and are separated by regions of low density. The method is robust against outliers since an outlier affects clustering only in the neighborhood of the data point. It can handle outliers and discover clusters of arbitrary shape. Density based clustering has a complexity of the same order as hierarchical clustering[9]. Given a competitive clustering method, learning is first conducted to adjust the algorithmic parameters; after the learning phase is completed, the work is ready for generalization. As is shown in Figure 1, the internal meaning of symbols in the clustering process of two-dimensional grid are as follows: if $X=\{x_1, x_2, \ldots, x_i, \ldots\}$, then each d-dimensional element data $xi=\{x_{i1}, x_{i2}, \ldots, x_{id}\}$. Every element is d-dimensional, the space of D_R is divided into d dimension, the results indicate that each data will be applied into the d-dimensional space[10].



(a) Original state of data distribution; (b) Clustering results

Fig. 1. Illustration of online clustering

3 Online Clustering Algorithm of Data Stream

When an initial data stream is in a region with few patterns, this results in a large cluster. The disadvantage can be remedied by the modified approaches. The clustering starts from one cluster, and it splits the cluster with the largest intracluster distance into two. After each splitting, the result is applied until the existing clusters are convergent.

The Clustering Process

Part 1: The pseudo code of the content mentioned in part 1 in the article above describes as follows:

0 Input: data stream $X=\{x_1, x_2, \ldots x_i, \ldots\}$
1 c=0, initialization on density grid structure
2 while x_i reaches do
3 apply the data x_i into sliding window c + +
4 if c=N, clustering initialization
5 if c>N, clustering modification
6 end while

Part 2: The pseudo code of the clustering initialization can be described as follows:

0 Input: the element of data in the present window{ x_1, x_2, …x_n, …}
1 Mapping data x_i→window $g(x_i)$, refresh the density D(g)
2 Mark the grid type: dense/variation/sparse, set dense grid as the groups center g
3 For each groups center g
4 Merge all the neighbor grid
5 Merge the connectivity groups
6 End for

Part 3: The pseudo code of advanced approaches on clustering description is:

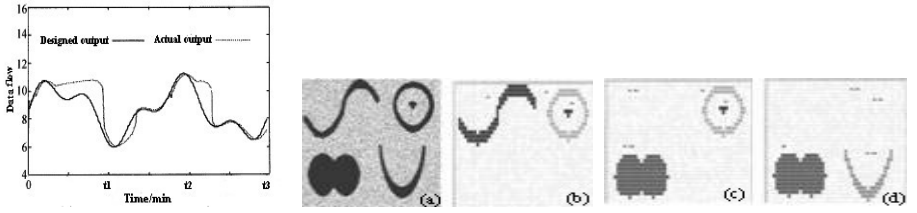
0 Input: the element data appearing for windows sliding are x_p and x_q
1 Remove x_p, and refresh the clustering density g_p
2 Mapping x_q into corresponging clustering, refresh the density of g_q
3 If g_p is parse, then remove it
4 If the original groups splits after removal, then novel groups appears
5 If g_p is the groups center core, then check the moving state
6 If the core moves, then choose the novel center
7 If g_p is dense, then apply it into the neighbor grip
8 If the density of g_p is higher than that of the core, then replace it
9 If x_q=x_p, then do nothing
10 If the connectivity groups exist, then merge the groups

4 Results and Analysis

Many clustering algorithms are based on the concept of algorithm extends the grid guoups by using the equal volume. The algorithm takes typically n times as long as the grid to complete cluster formation.

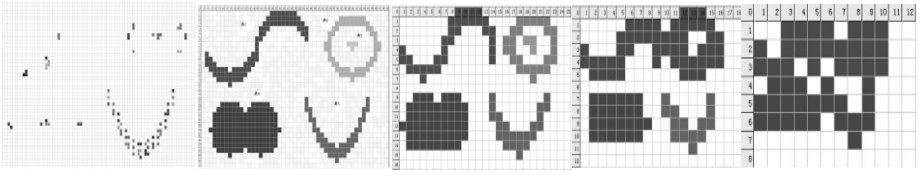
Experimental Design
The investigation is studied on the computer platform with Windows XP operation system, and the calculation approach iss attained with Microsoft Visual C + + 6.0. The first group is to check the dynamic change of data stream by two-dimensional database. The second parts show the effect of L values on clustering. The third group describes the clustering purity variation at different time.

Experimental Results
No 1 group: The investigation contains two-dimensional point set of 30 kB data volume with 5 clusters, and put 3 kB noisy point into the group. Related parameters are N=10000, L=0.02, C_{max}=5.0 and C_{min}=0.8. In the data stream, these 5 clusters appears sequently, as is shown in Figure 2(a), 5 intensive regions and plenty of abnormal points could be found. The clustering achievements at different time are indicated from Figure 2(b) to Figure 2(d). The results shows: the clustering points could be found automatically without clustering quantity k being given, the abnormal points appears at the same time, indicating the excellent clustering quality.



(a) Original data; (b) Clustering results at t_1; (c) Clustering results at t_2; (d) Clustering results at t_3;

Fig. 2. Online clustering process



(a) L=0.0125 (b) L=0.0200 (c) L=0.0400 (d) L=0.0600 (e) L=0.1000

Fig. 3. Online clustering on different L

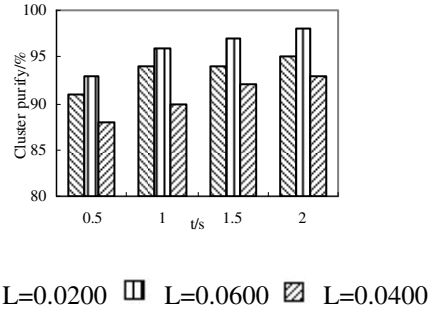


Fig. 4. Clustering purity on synthetic data

Table 1. Online clustering efficiency comparation

Time/s	20	40	60	80	100	120
Online clustering efficiency	0.2	1.3	1.7	2.1	2.3	2.5
Data stream	2.7	3.8	5.8	6.7	7.3	8.2

No 2 group: Choose different region length to reveal the effect of L on clustering results. The results in Figure 3 show: the cluster purity is optimal at L=0.0200, the forming quantity of cluster equals to the original ones; if L=0.0600, 3 clusters appears with low purity; while L=0.1000, most of the data are collected into the same grid

with 2 clusters. In the paper, the value of L is manual setting, different L influents cluster quantity obviously. The bigger the value of L, the lower purity attained; besides, much smaller L results in the failure of essential clusters forming. In conclusion, the value of L should be selected properly.

No 3 group: The purity of cluster refers to the ratio of primary type to the special type in data points groups. The paper believes the online clustering algorithm owes higher efficiency than that of D-stream method. At the beginning of the clusteing process, essential pretreatment is done on neighbor grid.

5 Conclusion

Clustering is one of the most important data analysis methods, and it has become an important tool for data mining, also known as knowledge discovery in grid structure, which emerges as a rapidly growing area. Data mining is to automatically search large stores of data for consistent patterns and relationships between variables so as to predict future behavior. Grid structured databases have well defined features and data mining can easily succeed with good results.

The paper proposed the online clustering algorithm, based on grid structure, could effectively enhance the clustering achievements. The approaches use the benefits of grid clustering, and play an important role in the treatment process on data stream, which might provide essential online clustering information in real-time.

References

[1] Read, N.K., Ray, W.H.: J. J. Process Contr. 8, 1 (1998)
[2] Sarimveis, H., Alexandridis, A., Tsekouras, G., et al.: Engi. Chem. Res. 41, 751 (2002)
[3] Zhu, Q., Cai, Y., Liu, L.: Neural Networks 12, 527 (1999)

The Measurement of Low-Level ^{220}Rn by Software Delayed Coincidence Method

Yan Yongjun, Lai Wei, and Zhou Jianliang

School of Nuclear Science and Technology University of South China
Hengyang, China
yan_jason@163.com

Abstract. This paper introduces a new software delayed coincidence counting technology,which requires less storage space. An experiment system was established for measurement of lowlevel ^{220}Rn .The methods were validated experimentally by intercomparison measurements. The results show that ^{220}Rn of decay rate of between 3dpm～20dpm can be measured by the experiment system with error of ±2%, when the background count rate is of 2dpm～4dpm.

Keywords: software delayed coincidence, low-level ^{220}Rn, and comparative experiments.

1 Introduction

The inhalation of ^{222}Rn/^{220}Rn in the indoor and outdoor environments is responsible for the greatest single contribution to the radiation dose to the human population. The decay products from ^{222}Rn are the dominant reported source, partly due to the difficulty of measuring ^{220}Rn, especially at low concentrations and in the presence of ^{222}Rn. But the doses from the decay products of ^{220}Rn can not be neglected, which makes the accurate measurement that discriminate between ^{222}Rn and ^{220}Rn in the environment an important subject in radon metrics and radiation protection.

A kind of delayed coincidence method has been successfully applied to measurement of low concentration of ^{220}Rn in air and radium isotopes indirectly in seawater [1]-[6]. Inspired by these successful applications, a new improved system designed for measurement of ^{220}Rn concentration has been constructed, which is based on a new software coincidence counting (SCC). The method is different from conventional coincidence counting (CCC) mainly in the method how the coincidence count rate is obtained. The software coincidence counting (SCC) is based on recording the time and the amplitude of individual pulses during a measurement. In this case, only the time information of pulse trains is recorded, coincidence counts are consequently obtained by software processing of the recorded data, which is stored in a format that requires less space.

The SCC can also be used for liquid scintillation count-ing and nuclides spectrum analysis. It can use continuous sampling of all detector signals and records both shape

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 407–413.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

and time-interval information of the pulses. The main advantage of this method is the possibility to adjust various parameterssuch as resolving time and energy windows, after the measurement and to treat the same data records using different calculation method. SCC can replace the CCC in activity measurement.

2 Software Delayed Coincidence Theory for Measurement of ^{220}Rn

Delayed coincidence technique is based on the coincidence of alpha particles from the decay of ^{220}Rn and its relatively short-lived decay component ^{216}Po [1]. Owing to its short half-life (150ms), the alpha particle emitted by the decay of a ^{220}Rn atom will, within a fraction of a second, be followed by an alpha particle from the decay of ^{216}Po. This method separates the fraction of alpha events originating from ^{216}Po from all the other alpha events.

The time and time-interval information pertaining to pulse trains is recorded by using radiation detection system. If the time interval between successive pulses is less than coincident time set by software, the pulse is used to calculate the activity of ^{220}Rn. The coincidence count is used in statistical analysis to ascertain that the real ^{220}Rn-^{216}Po pair counts are correctly identified. Poisson distribution probability is introduced to correct for accident coincidence pulses.

In the situation when ^{220}Rn and ^{216}Po coexist with additional background activity without short-lived daughters, the time intervals will be distributed as:

$$P(t)dt = \left(A_P \eta_P \eta_d e^{-\lambda_d t} + B_g \right) dt \tag{2.1}$$

Where,

- P(t)=probabitity that the time interval appears between t and t+dt
- A_p =activity of the paratent nuclide
- η_p =counting efficiency for alpha particles from parent nuclide
- η_d =counting efficiency for alpha particles from daughter nuclide
- λ_d =decay constant for daughter nuclide
- B_g =random coincidence distribution (background and a part of parent and daughter activity if η_p and η_d is less than 1.0).

According to the probability of time interval and the decay scheme of ^{220}Rn, this method uses the short decay half-life of 150 ms of ^{216}Po in order to distinguish between ^{220}Rn-^{216}Po pair counts and accident coincidence. This method is effective and convenient because it can apply to the flow-through type scintillation cell, which can be constructed comparatively simple and requires little maintenance. Here coincidence time is set at 600ms equaling to four times half life of ^{216}Po to ensure that all pulses from decay of ^{216}Po are recorded.

^{224}Ra------->^{220}Rn------->^{216}P$_0$------->^{212}Pb
$T_{1/2}$ =3.6d 55s 150ms

Fig. 1. Decay scheme of ^{220}Rn

3 Software Delayed Coincidence System and Measurement

3.1 Experimental System Described

A schematic diagram of the measurement system is shown in Fig.2. Software delayed coincidence system consists of FD-125 Analyzer of Radon and Thorium, flow-through scintillation cell facing a photomultiplier tube, amplification and shaping circuit, EL-ARM-830 test system. The air then passes through a flow mter followed by a special scintillation cell. The air is driven by a pump through a buffer bottle which is used for stabilization of the flow rate.

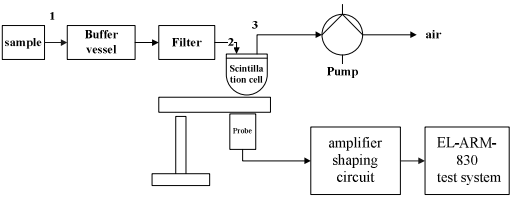


Fig. 2. Sketch of the apparatus used for the measurement

Software delayed coincidence system uses EL-ARM-830 test system as the control center. It records time information of pulse sequences from amplifier sha-ing circuit.According to the above theory of measurement of ^{220}Rn, the software comprises three main sections: pulse-train recording, pulse processing and display of results. The signals from thedetectors can also be viewed in a manner of GUI. The length of pulse trains to be stored can be preset according to either the couing time or the amount of Flash-ROM memory space to be filled.

3.2 Measurement Methods and Formula

3.2.1 Methods and Formula of ^{220}Rn Decay Rate in Scintillation Cell

As shown in Fig.2, the detector used is a commerci-ally available radon monitor using scintillator cell as detector working in flow-through mode. Air is sucked through a filter to remove any airborne particles and decay products from ^{220}Rn and ^{222}Rn before entering the sensitive volume.

Formula of ^{220}Rn decay rate in scintillation cell.

$$A_2 = A \times f_{decay} = A \times f_1 \times f_2 \tag{3.1}$$

Where: A is the initial activity of ^{220}Rn, f_1 is factor for the ^{220}Rn atoms flowing through the sample to the scintillation cell (from 1 to 2 in figure 2), f_2 is factor for the ^{220}Rn atoms decay through scintillation cell (from 2 to 3 in figure 2).

$$f_1 = e^{-\lambda \times \frac{V_1}{\omega}}, \quad f_2 = (1 - e^{-\lambda \times \frac{V_2}{\omega}}) \tag{3.2}$$

Where V1 is the volume from sample to scintillation cell inlet, V_2 is the volume of scintillation cell,ω is air sampling pump flow rate (L / min), λ is the decay constant of ^{220}Rn.

3.2.2 Detection Efficiency Formula

Detection efficiency of ^{220}Rn channel (E_{220}) includes two aspects: *a)* detection efficiency of scintillation cell, f_D*; b)* the fraction of detected pairs for which the ^{216}Po event falls within coincidence time T_g, f_E.

Therefore,

$$E_{220} = f_D^2 \times f_E \tag{3.3}$$

Coincident time to 600ms, then $f_E = 1 - e^{-\frac{\ln 2}{150} \times 600} = 0.9374$

The detection efficiency is about 0.86, so theoretical value of detection effici-ency of ^{220}Rn is 0.693 according to the formula 4.

3.2.3 The Chance Coincidence Count Rate

Giffin et al [1] derived expressions for calculating the fraction of accident coincidence events which were expected to occur in ^{220}Rn channel. The correction factor is :

$$y = \frac{(b.c - a.c)^2 \times Tg}{1 - (b.c - a.c) \times Tg} \tag{3.4}$$

Where y is chance coincidence count rate, b.c is the total count rate, a.c is the after coincidence count rate in ^{220}Rn channel, and Tg is the coincidence time.

Net count rate =a.c- y.

3.2.4 Formula of ^{220}Rn Activity

$$^{220}\text{Rn activity= Net count rate}/(E_{220} \times f_{decay}) \tag{3.5}$$

4 Measurement Result of ^{220}Rn Source and Verification

4.1 Measurement Result of ^{220}Rn Source

^{220}Rn source is measured using experiment system of delayed coincidence measure-ment. The results are shown in Table 1.

Table 1. Result of ^{220}Rn source using delayed coincedence system

Flow rate (L/min)	*Background of scintillation (cpm)*	*Total count (cpm)*	*Coincidence count (cpm)*	*Chance coincidence count (cpm)*	*Net count (cpm)*
0.32	3.72	9.78	2.93±0.06	0.51±0.02	2.42±0.07
0.40	1.72	20.44	7.90±0.10	1.81±0.06	6.09±0.12
0.52	4.00	41.01	19.00±0.16	6.22±0.18	12.78±0.24

Annotate: result is average of five times measurement.

4.2 Comparative Experiments

Experimental system of delayed coincidence measure-ment is compared with RAD7 continuous radon monitor to verify the accuracy of the results.

Measurement results of ^{220}Rn sources using RAD7 are shown in Table 2. The correction factor of measurement ^{220}Rn is 1.43 [8].

Table 2. Result of ^{220}Rn source using RAD7

Five measurement results of 220Rn sources using RAD7					*average*
12400±2000	12400±2000	12800±2030	12300±2000	11700±1950	12320±396

Conversion formula between concentration and activity of ^{220}Rn is as follows:

$$1.43 \times C(Bq/m^3) = \frac{A(Bq) \times \lambda(/s) \times 60}{\omega(L/\min) \times 10^{-3}} \tag{4.1}$$

Internal pump flow rate of RAD7 is 0.65L/min. Activi-ty of ^{220}Rn is calculated according to the formula 7.

$$A = \frac{1.43 \times C \times \omega \times 10^{-3}}{\lambda \times 60} = \frac{1.43 \times 12320 \times 0.65 \times 10^{-3}}{(0.693/55) \times 60} = 15.147 Bq \tag{4.2}$$

Comparison between experimental system of delayed coinceidence measurement and RAD7 are shown in Table 3.

Table 3. Comparison of results by using two method

Result by experimental system of delayed coincidence measurement			*Activity by RAD7 (Bq)*	*Relative error (%)*
Net count(cpm)	*fdecay*	*Activity (Bq)*		
2.42±0.07	0.003832	15.20	15.15	0.33
6.09±0.12	0.00972	15.07		-0.52
12.78±0.24	0.02057	14.94		-1.39

Annotate: V1=2.2L, V2=0.5L.

As shown in Table 3, Relative error between experimental system of delayed coincidence measurement and RAD7 is ± 2% or less, showing the measurement accuracy using this method for ^{220}Rn source.

5 Conclusion

As shown in table 4.1 and table 4.3, ^{220}Rn of decay rate of between 3dpm～20dpm can be measured by the experiment system within error of ±2%, when the background count rate is of 2dpm～4dpm. Experimental results show that this study reached the target of resolved measurement of radon isotopes. the source of Thoron (^{220}Rn) and radon (^{222}Rn), thorium and Uranium, is always present together in soil gas of outdoor, in building material indoor. The pulse-coincidence monitor will be used to investigate time-varying ^{220}Rn/^{222}Rn concentration in tis situation.This research could lead to a broader knowledge of the source of ^{220}Rn/ ^{222}Rn.

The sensitivity of the monitor is estimated to be particularly useful for a low-level concentration. Furthermore, One of the main benefits of the software delayed coincidence method is that the same pulse trains may be processed multimes repeatedly by defferent methods. In this way, methods such as selective sampling and anti-coincidence counting may be directly compared offers advantages over conventional coincidence counting (CCC) in terms of its reliability, flexibility, cost and convenience. The method of soft-ware coincidence counting (SCC) has proven its feasibility and is useful for absolute activity measurements.

Acknowledgments. This study was supported Supported by the National Natural Science Foundation of China (Grant No. 11175082).

References

[1] Giffin, C., Kaufman, A., Broecker, W.: Delayed Coincidence Counter For The Assay Of Actinon and Thoron. Journal of Geophysical Research 68, 1749–1757 (1963)
[2] Moore, W.S., Arnold, R.: Measurement of ^{223}Ra and ^{224}R-a in coastal waters using a delayed coincidence counter. Journal of Geophysical Research 101, 1321–1329 (1996)
[3] Moore, W.S.: Fifteen years experience in measuring ^{224}Ra and ^{223}Ra by delayed-coincidence. Counting. Marine Chemistry 109, 1–10 (2007)
[4] Bogen, P., Sappok, M., Schumann, G.: Measurement of Atmosphereic Radon-220 Using Delayed α-particle coincidence. Arch. Met. Geophys. Biolk. Ser. 21, 171–182 (1972)
[5] Falk, R., More, H., Nyblom, L.: Messurement of ^{220}Rn in air using a flow-through lucas cell and multiple time analysis of recorded pulse events. Nuclear Technology Publishing 45, 111–113 (1992)
[6] Limoto, T., Kurosawa, R.: A pulse-coincidence ^{220}Rn monitor with three time-gates. Environment International 22, 1139–1145 (1996)
[7] Qiu, S.: ^{220}Rn measurement instruments calibration problem. Nuclear Technology 26, 561–563 (2003)
[8] Qiu, S., Xiao, D., Liu, L.: First Intercomparison of Th-Rn Discrimination Measurement in China. Radiation Protection Dosimetry 128, 124–127 (2008)

An Optimization of Intelligent Fire Alarm System for High-Rise Building Based on ANASYS

Liu Keding

Hunan Urban Construction College, Xiangtan, Hunan Province, 411101, P.R. China

Abstract. High rise building has become the main building type in urbanized china. The Fire Alarm System (FAS) in an intelligent high rise building is the key of reducing casualty and property loss in an unpredictable fire. In this paper, a model of typical FAS is established, and with the analysis of fire evolution in region with ANASYS, the parameters of the FAS are suggested to optimize the effect of FAS. Case study shows that when response time of FAS reduces to less than 5 seconds, FAS could help the evacuation of a high rise building effectively.

Keywords: ANSYS, High-rise Building, fire alarm.

1 Introduction

With the development of economy, the process of urbanization goes faster and deeper. More and more people rush into city. And also because of the improvement of the facilities service capabilities of critical infrastructure system, the city can accommodate more people in the case of a certain limited area. With the invention of the elevator (lift) and cheaper abundant building materials, high-rise buildings became possible. High-rise, which is also known as apartment tower, office tower, apartment block, or block of flats, is a tall building or structure used as a residential and/or office building. The building is with an architectural height between 35 and 100 meters, and is a relatively good choice for high-density commercial and residential area for big city, even for smaller city to concentrate its limited superior resources.

The quantity increase speed of high-rise buildings is extraordinary in recent years. In Hong Kong, there are 2354 buildings taller than 100 meters, but most of citizens live and work in the thousands of high rise buildings. These buildings also rise fast in china when urban construction vigorously develops, and high-rise buildings gradually become one of the symbols in many major cities. For example, shanghai has built more than six thousands of high rise in 2010 and the number is still rising.

Fire always is the biggest threaten to human beings. Because of the high building volume rate, a high rise building accommodates more households, at the same time it also accommodate more fire hazards. In recent years, several fire accidents taken place in china[1,2], including the fire of CCTV new site North Side Building in Beijing, fire of “11.15" Teacher's apartment in Shanghai, fire of Dynasty Wanxin in Shenyang, fire of "8.9" in Chongqing, fire of "4.22" in Qingdao, and fire of Mong Kok in Hong Kong, cause not only great loss of life and property, but also impact

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 415–421.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

greatly people's understanding of high-rise building fire safety, and request a higher demand for building fire alarm system and fireproof system.

In this paper, an optimization of fire alarm system is discussed basing on the analysis of fire process in high rise building with ANASYS software. Literature reviews are given in section 2, then the model of FAS in a high rise building are introduced in section 3, the fire evolution simulation based on ANASYS are emphasized in section 4, then a typical case study of high rise building is shown in section 5.

2 Literature Review

High rise buildings are taking benefit from new material science. In China, during the past years, fire safety design of a building must comply with prescriptive codes, such as National Standards for Fire-Safety Design of Buildings (GBJ16-87) [3]. New fire proof material can effectively delay the expansion of fire in buildings and changes the result of fire protection design [4]. Because of the development of information science and technology, a building can be equipped with multifunctional sensors, modern energy control system, and by integral the building, technology, and energy systems to form an intelligent building. These technologies are widely used in high rise building to improve comfort and energy efficiency, and to reduce the response time for fire alarm in building [5]. Many research work [2,3,6] have been done to improve the fire alarm system in an intelligent buildings.

In high-rise building once fire breaks out, it would rapidly expand to adjacent area and cause damage to building structure. To reduce casualties, evacuation should be carried out at the time of the fire is detected and the alarm is triggered. Numerous scholars carried on the discussions of numerical simulation on many aspects after the high-rise building fire [1, 7-8]. ANSYS is a powerful tool of finite element method, for simulation methods to solve the most challenging engineering problems, especially for the analysis of the character in fire district, and the micro process of fire expansion.

These researches have provided effective support to understand of the fire disaster and to improve the design of fire alarm system (FAS). In this paper, research will be to optimize the character value of FAS and Fire Fighting System (FFS) with ANSYS simulations.

3 Fire Alarm System of High Rise Building

In a high rise building, goal of the fire alarm system is to alarm on time, so that unpredictable fire situation can be put out at the very beginning, or people in building has enough time to evacuation, or help the fire department to put out uncontrollable fire soon.By the international standards, an intelligent building is composed with by the Building Automation System (BAS), Office Automation System (OAS), Communication Automation System (CAS), Structured Cabling System (SCS), Automatic Fire Alarm System and Security Automation System (SAS). FAS plays a vital security role in the modern intelligent building. It belongs to a subsystem of the

intelligent building systems, but runs in total isolation from other systems or networks to complete its mission of anti-disaster and fire-fighting with absolute priority. The broad FAS also includes the fight fighting devices, such as automatic sprinklers, fire hydrant controls, automatic fire door and fire shutter controls, ventilation systems, so that FAS is able to take action at the first time.

The FAS of an intelligent high rise building usually consists of three parts: the fire detectors (FDs), manual alarm buttons (MABs), regional fire alarm centers (RFACs), a central fire alarm unit (CFAU) and alarm display and sound devices (ADSDs), and fire-fighting devices (FFDs), shown in Figure 1.

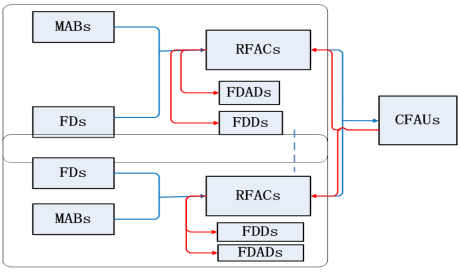


Fig. 1. Structure of FAS

Blue lines refer to fire information flow, and red lines refer to command flow.

Usually to a high rise building which has about 30 floors, there are regions of every limited number of floors, and each region has numbers of various fire detectors, manual alarm buttons on each floor, alarm display/sound devices on each unit or room, and fire-fighting devices. The situation of the building region, including the reading of thermal, smoke and light detectors, are transmitted to regional fire alarm centers, which will be analyzed by regional distributed program to prevent false alarm. If real alarm is to be triggered, the situation information will be also uploaded to CFAU. The CFAU will report fire alarm to local fire department, and at the same time CFAU begins to coordinate all the FDDs in building to fight against fire, including turning on the automatic sprays, normally opened doors, fire resisting shutters and so on.

In a distributed FAS system, the RFACs are the most important units. The program in one RFAC is customized for the environment of the specified region, to collect and analyze the real-time date of all connected fire detectors in order to form an overall fire hazards. By model analysis and simulation, the trend caused by small changes of detector readings are listed, and the risk are assessed so that proper fire-fighting action will be taken before the irreversible result happened. By crossing check of redundancy data from multiple detectors and multiple dimensions (including the sound, thermal and light dimensions), the false alarm will be identified and avoid precisely.

So the setup of the fire evolution model for the specified high rise building is the base of correct setup of the control program in RFACs and CFAU. The fire model in high rise building represents a complex nonlinear system without simple answers.

4 Optimize the FAS Parameters with ANASYS

Finite Element Method (FEM) provides an effective solution to the fire model, including model of high rise building.

Fire reaction starts from the burning of combustible materials, then the release of heat by convection and radiation on surface, heat conduction from outside to inside of the structure, then non-uniform temperature field inside causes the thermal response of the structure. The heating up lids to a series of physical and chemical changes, then mechanical properties change, then material's structure strength, stiffness and deformation capacity change. In the role of fire, the building structure may be rapidly destroyed, even lose support capabilities and collapse.

The first step of the structural fire response analysis is fire analysis which can be resolved with CFX software from ANSYS Inc. Usually a series of abstract assumptions, such as the fire was controlled within a limited range, and uniform temperature distribution in the space, are introduced to simplify the formula of temperature-time. The second step, thermal analysis, is precisely to be done with ANSYS software. According to the temperature distribution, the building structure, the analysis is to find the temperature field within building components. The initiative action of FAS should be included in the analysis, such as fire sprinklers. The timings to perform the action are the key factors to consider.

The FAS works in a typical procedure: when the real time environmental data provided by the FDs changes through the preset critical values, RFACs will assume that a fire is taken place at somewhere, then at the same time action commands are sent from CFAU and RFACs to FDDs around the fire position to fight fire, and alarm the householders to evacuation.

5 Example Practice

In this paper, a 30-floor, steel-structure, high rise building is taken as example. Each floor there are 3 apartments with support of one lift and one stairwell.

The FAS contains 15 RFACs and 1 CAFU, and 3 kinds of detectors are all equipped in the hall and in each room of the three apartments. Each RFAC covers three floors, and to every pair of closet RFACs there are one floor covered by them both. All detectors connect to RFACs by hardwire, and RFACs connect to CAFU by both cable and wireless links.

Use ANASYS to simulate the fire model in the target building. Take each 3 floors as a model unit, and assume that the fire take place at the bottom floor. Using the standard fire model, the temperature in room are expressed as

$$T = T_0 + 2451g(8t + 1) \tag{1}$$

Here T is the real-time temperature, $T_0 = 20$ is the environmental temperature, and t is relative time from the fire start.

Because the fire compartment (FC) can prevent spread of fire and Isolate heat transfer, the heating process of structural components will be significantly only in FC. In order to the accurately simulate the non-uniform temperature distribution on the beams and columns in FC and to capture the local buckling phenomena, the beams and columns are modeled as shell elements, and the others are modeled as beam elements. In the joint point, constraint equation is used to ensure deformation compatibility. The Model is shown in figure 2.

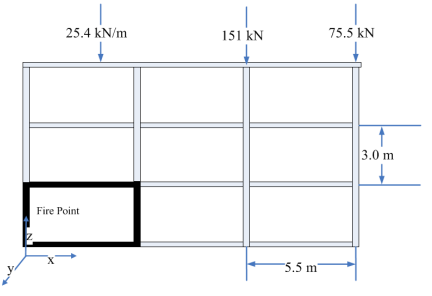


Fig. 2. Frame model of building part covered by one RFAC

Thermal analysis is taken in ANSYS based on the principle of conservation of energy, with consideration of five heat loads including temperature, heat flow rate, convection, heat flow density and Health generation rate. In this case of transient thermal analysis, ANSYS uses formula (2),

$$[C]\{T'\} + [K]\{T\} = \{Q\} \tag{2}$$

Here $[K]$ is transmission matrix including thermal conductivity rate, heat transfer coefficient, emissivity and shape factor, $[C]$ is matrix of specific heat, $[T]$ is temperature vector and $\{T'\} = \dfrac{d\{T\}}{dt}$, $\{Q\}$ is vector of heat flow rate at node including thermal generation.

The calculation of ANASYS gives the temperature distribution of the iron frame. In figure 3 it shows the temperature change at left end. From the ($T \mid t$) curve, an conclusion can be drawn: (1)the temperature of steel frame rise rapidly, and after 180s the maximum temperature of the steel framework reaches of 243(2) the FDDs' actions can delay the temperature rising effectively. If the FAS get alarm and takes FDDs into action in five seconds, the fire will be suppressed for 100 seconds before

temperature rising fast. Even the FDDs taken actions after 50 seconds, it will still buy about 50 seconds time to evacuation people from the fire point room. (3) The coordination of RFACs can delay the fire expansion from one region to the other by sending commands to FDDs in region border.

6 Conclusions

High rise building has already become the main type of building in big city. With the assistant of intelligent building concept and technology, the fire alarm system in high rise building is effective to delay the expansion of fire and reduce the casualty and damage. ANASYS and FEM provide a great help to analyze the fire model and improve the efficiency and response speed of the FAS.

In this paper, the fire models in ANSYS are still abstract, without detail concern on the distribution of detectors, and the coordination of RFACs are not shown in simulation. Our future work is to improve the detail level of model, and study the coordination between RFACs.

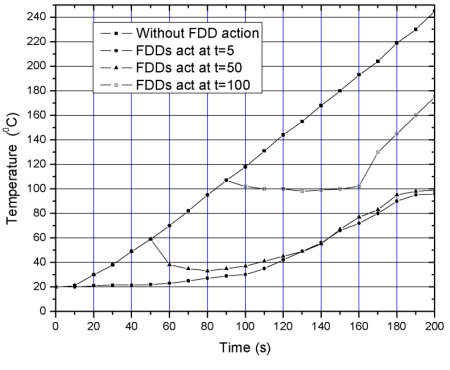


Fig. 3. Temperature at left end of the steel frame

The actions of FDDs affect the temperature rising obviously.

References

[1] Zhu, J., Huo, R., Fu, Y.: Smoke control in super tall buildings. Fire Science and Technology (1), 59–62 (2007)

[2] Jing, C., He, J., Zhou, R., Xu, X.: Numerical Simulation of Fire Smoke Control Models in High-Rise Buildings. Building Science 25(7), 18–22 (2009)

[3] Ministry of Construction, Ministry of Public Security. Building Construction Fire Safety Engineering Assessment Clause No. 30. Amendment to GBJ16-87. Planning Press, Beijing (1999)

[4] Guo, T.-N., Fu, Z.-M.: The fire situation and progress in fire safety science and technology in China. Fire Safety Journal 42(3), 171–182 (2007)
[5] Sinopoli, J.: Smart Buildings Systems for Architects. Owners and Builders (2010)
[6] Song, Z.: Design of fire alarm systems in the intelligent buildings. Journal of Chongqing University of Posts and Telecommunications 16(4), 98–100 (2004)
[7] Zhu, J., Huo, R., Fu, Y.: Smoke control in super tall buildings. Fire Science and Technology 26(1), 54–57 (2007)
[8] Xiao, G., Liao, G.: Study on Human Evacuation Style in Building Fire. China Safety Science Journal 16(2), 26–29 (2006)

Coal Mine Safety Monitoring and Evaluation System Based on .Net

Zhang Chunzhi, Hu Guangzhou, Li Haoru, and Wang Kejia

Harbin Institute of Petroleum, Harbin, Heilongjiang Prov 150027

Abstract. With the advent of information age in 21th century, it is becoming an inevitable trend of development and survival for coal enterprises to use information means to ensure safety production. However, coal enterprises possess a lot of special characteristics, such as complex geographical environment and large amount of information. Certain functions including traditional underground monitoring system and dispatching telephones act on their own ways which leads to a bad coordination. This paper takes full consideration of personnel transferring, material transferring and information transferring and use .net technology to construct coal mine safety monitoring system. We use human-machine-integration method to evaluate working safety and provide well-developed security for production safety of coal enterprises.

Keywords: net, information transferring, coal safety, comprehensive.

1 Introduction

It is becoming a trend of 21th century informatization for coal enterprises use information method to assure safety production. However, coal enterprises are usually located in remote area which has a bad environment. And underground work has a complex working conditions and extreme production environment. Moreover coal production needs complicated technologies and equipment which blocks the construction of management information system [1]. Traditional coal dispatching telephone, underground monitoring system and real-time monitoring system provide security for coal safety production in different aspects. But they act in their own ways which results in a bad coordination. So they cannot overall improve the safety production level.Coal mine safety monitoring and evaluation system constructed in this paper is developed with .net of Microsoft cooperation. We fully analyze the coal working procedure and establish a developed work regulation management system, disaster response system and information safety system to monitor personnel, material and information flows. In addition man-machine-integration method is adopted for intelligent evaluation on safety which provides a well-fledged security for coal safety production.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 423–429.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Key Technology in System

.net is an internet-oriented and user-support development platform developed by Microsoft cooperation. This platform is famous for its powerful function and full object-oriented. It offers abundant service widgets and brings information, equipment and human together in an individualized and unified method. Its functions include generating desktop applications, mobile applications, web applications and XML Web services. Sometimes it integrates some development languages such as Visual Basic .NET, Visual C++ .NET, Visual C# .NET and Visual J# .NET, allowing them to share tools. This contributes to creating mix language to solve problems to some degrees [2]. C/S network model is a structure based on server and user. It fully considers advantages of both sides, assigns tasks reasonably and lowers the communication expense between server and user. C/S model can be divided into two types: 3 layers and 2 layers. Two layers type of C/S model puts data logic layer and storage layer in server-side, presentation layer and transaction layers in user-side. Because most business logic and interface display need user-side to implement, user-side need to assume heavier pressure. This framework is a model of fat-user-thin-server. Three layers C/S model added a middle application server where transaction layer is. Only presentation layer is in user-side. Users only need to run under low pressure. This is a thin-user model [4]. B/S network model is a structure based on browser and server. It has following advantage. First, maintenance and upgrade work is performed in server-side which leads to a low development and maintenance cost. Second handling with key affairs such as complicated calculation and data access can relieve user's stress. Third it adopts open specification to ensure generality and cross- platform. Browser installed in user-side has a better portability. Fourth clients cannot direct access database because of Web server, effectively preventing illegal invasion. Finally all layers are independent, improving system maintenance [4].

3 Analysis on Characters of Coal Enterprise Working Procedures

The core task of coal enterprise is coal mining and transportation. Its main work place is underground mines which have a complex working condition, a severe environment and complicated technological equipment. According to characters of coal enterprise working procedure, we should consider safety monitoring from three parts: personnel flow, material flow and safe information flow.

Guarantee workers' safety is the first step. The personnel flow means the working conditions of dispatching workers, monitoring workers, mining workers, transporting workers and emergency response workers. Each kind of works acts its own function.

Coal mining and transportation is the core task of coal industry. The material flow includes ore extraction, transportation, supply material and emergency material transportation. It must be orderly developed in strict accordance with related regulations and rules. Safe information flow is the primary measure of safety

production. It consists of real-time data obtained by sensors such as ventilation, drainage, electricity, transportation, extraction and monitoring. It provides data basis for safety evaluation of production environment and references for management decisions.Therefore coal mine safety monitoring and evaluation system should possess following functions:

1) Construct scientific enterprise information management, refine working system and rewards and penalties system, improve information consciousness of managers and workers and secure the overall work safety.
2) Comprehensively control the personnel flow, position workers anytime and make sure all workers in position under safer environment.
3) Comprehensively control the material flow, check real-time of coal mining and transportation anytime, and ensure smooth progress, supply materials and open transportation route.
4) Use monitoring safe information flow to find hidden danger in mines, carry out quantitative evaluation and provide safety measures or advice for policymakers.

4 Fuzzy Algorithm Design of Safety Evaluation Fuzzy Algorithm

According to the features of coal enterprises working flow, we use human-machine-integration method in this paper. First we organize experts to ensure evaluation factors and acquire data through questionnaire, on-the-spot investigation and data collection. Second we rate the evaluation factors by Delphi method. In the end, we comprehensive evaluate the factors to gain quantitative assessed value. The algorithm process is as follows.

1) Establish evaluation factor set S={S1,S2…Sn}. n is the amount of factors.
2) Establish weight sets and the weight of each factors in set:

$$A = \{a_1, a_2, \ldots, a_n\}, \sum_{i=1}^{n} a_i = 1$$

3) Establish evaluation factor comment set and assign values: $V = \{v_1, v_2, \ldots, v_m\}$. m means the amount of comment.
4) Establish membership matrix of evaluation factor. Each expert uses Delphi method to grade each factor according to comments. The interval is from 0 to 1 while the score should satisfy equation: $\sum V = 1$. Taking these scores as corresponding membership lead to a matrix: $R = (r_{i,j})$.

   $i = 1,2,\ldots,n; j = 1,2,\ldots,s$. The $r_{i,j}$ stands for single evaluation factor s_i's membership to comment v_j.
5) Conduct fuzzy integrated assessment and obtain the safety.

Calculate the comprehensive evaluation vector of comment set: $B = A \bullet R = (b_1, b_2, ..., b_m)$ Obtain the safety result: $G = V \bullet B$

5 System Design

5.1 Network Structural Design

Most modern coal enterprises adopt distribution structure. All the working system is made up of site control stations and control center station. Site control stations can be independent from center and automatically implement local monitoring and control functions. Control center is responsible for sensor data collection, data storage, data and video display, alarm, safety assessment, report printing, information safety maintenance. According to analysis on advantages and disadvantages of C/S and B/S network models, site control station should adopt C/S model because of high instantaneity while control center should take B/S-C/S mixed model because of large number of business and low instantaneity. The whole network structure is shown in figure 1. It can adopt two modes integration method.

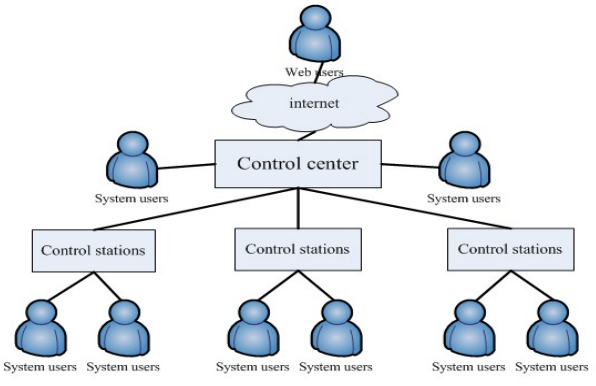


Fig. 1. Network structure

According to the analysis on working procedures and content of coal enterprise, this system is divided into three layers: physics layer, information layer and application layer as shown in Figure 2.

System physical layer consists of ventilation sensor, ventilation controllor, drainage sensor, drainage controllor, electronic sensor, electronic controllor, transportation sensor, monitoring camera, RFID equipment and warning device. Physical layer provides data of personnel, material and information flow for various sensors. On the other side, sensors accept and execute control command from application layer.

Information layer includes data of personnel, material, information flows and related information acquired from RFID equipment. It can obtain material flow

information from transportation sensor, extraction sensor, monitoring equipment and RFID device. Meanwhile, it obtains material flow data from ventilation sensor, drainage sensor and electronic sensor. Application layer has functions such as mechanism management, staff scheduling, material scheduling, safety assessment, disaster response, alarm management, video monitoring and information safety. Because mechanism management, safety assessment and information safety do not need real-time collected data, the system integrates them into the B/S framework. Mechanism management provides access for management layer. Administrators login the system through WEB and lay down various regulations and rules. Safety assessment function gives quantitative safety assessment in comprehensive consideration of analysis on experts' comment and collected data which provide policy support for administrators. Information safety function includes setting hardware firewall in the port between system and internet, arranging intrusion detection system and antivirus programs and storing collected data by RAID storage arrays which secure the safety of system operation and information. Functions such as staff scheduling, material scheduling, disaster response, video monitoring and alarm have a high demand on real-time data. The system integrates them into C/S framework. Staff scheduling function can acquire personnel flow information through information layer and allocate the staff according to site condition. Material scheduling function can monitor material flow information through information layer and manage material flow information scientifically according transportation and mining working conditions. Disaster response and alarm functions can find disasters and send an alarm in time and conduct message warning through warning device and GSM system. Video monitoring function can reflect underground working video on the terminals of control stations and control center.

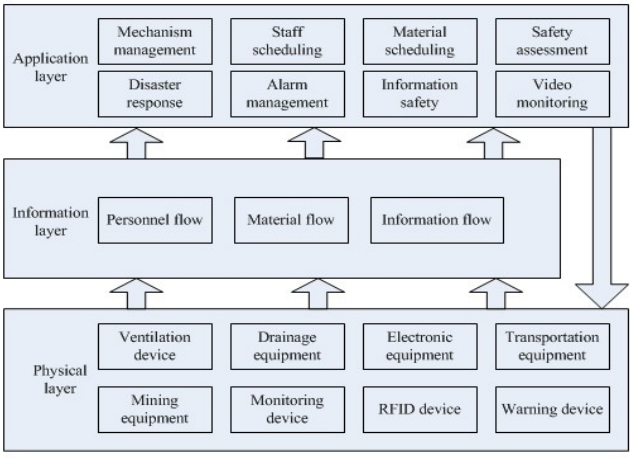


Fig. 2. Overall design

5.2 Function Module Design

To improve the system's maintainability, we adopt .net component technology according to the above design on the network structure of evaluation system. Each component is a relative independent function module. The entire system is divided into 7 main modules including system management module, personnel flow management module, material flow management module, information flow management module, safety assessment module, disaster response module and information safety module as shown in Figure 3.

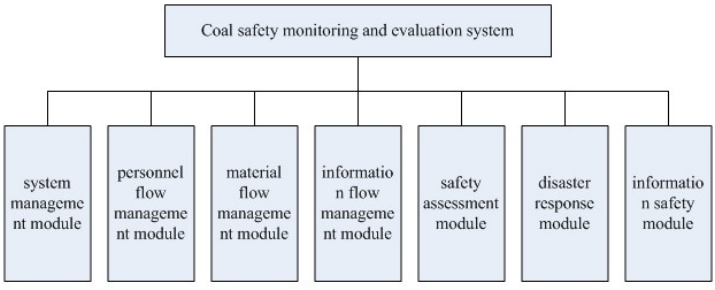


Fig. 3. System function module

6 Conclusion

This paper is aimed to improve the current situation of bad coordination among safety means such as dispatching telephones, underground monitoring system and real-time monitoring system which evaluation system which fully analyze the coal working procedures and establish a developed work regulation management system, disaster response system and information safety system to monitor personnel, material and information flows. Meantime man-machine-integration method is adopted for intelligent evaluation on safety which provides a well-fledged security for coal safety production.

Acknowledgement. The work was supported by Scientific and Technological Project of Heilongjiang Province Education Department under Grant(NO.11553018).

References

[1] Haibin, R.: The thought on coal enterprise information constructure. Mining Research and Development 24(2), 52–54 (2004)
[2] Xiaoqi, P.: Distributed control and test system based on. net platform–software design and implementation. Graduate dissertation. Northwestern Polytechnical University, Xi'an (2004)

[3] Zhang, Z., Huanghai: Information system based on integration mode with B/S and C/S. Administration of economy management department in Nanjing science and technology Univesity, Nanjing
[4] Jin, H.: Implementation of agricultural consult template–development method of WEB database. Journal of Northeast Agricultural University 31(3), 280–282 (2000)

Research on Web Information Filter Model

Liu Yiqin

Yunnan Radio & TV University, Kunming 650223, China

Abstract. According to high development of internet technology, web information has become richer and richer on a signal pages, traditional methods based on HTML technologies can't filter information effectively. Therefore, this paper designed a new information filter model based on DOM (Document Object Model) and theme analysis in partition. This method divided a web page into several parts and analysis them in a DOM tree with multiple theme key words, and stored analysis results in data base for search engine. Finally, after experiment, the result shows this method performs better in recycle rate and correct rate.

Keywords: Web Information Filter, Content Extract, DOM tree.

1 Introduction

In the recent decades, according to rapid development of computer and communication technologies, and the popular of personal computer, internet has becoming a global network and become part of people's life. It's very important in life, working, learning, and business [1]. Enterprise can publish their products and information on internet, individuals can get information what they need on internet. World Wide Web has become a new information hold which used to be showed in books, magazines, television. People's life has been deeply revolute by internet. It's very convenient, a laptop with wireless net card or a PC with a net cable is enough for you to surf on the net, reading latest news, reading books, publish personal information, get answers and knowledge what we desire to know. We can nearly get any information we want on internet. Internet is shared and rich content, which highly accelerate information exchanging, knowledge sharing [2]. There is rich content information on internet, from math to physical, from literature to music; we can even get personal thoughts on internet, which provides a dramatic world for us. What's different with traditional information holders only can get or put information in a signal way, we can communicate on internet by questions and answers, or we can play a game on internet [3]. It comes true, information period has come, and internet is becoming more and more important in people daily life. Since conventional internet web pages content searching applications are based on a signal pages which is not be used to divide. According to the high development of web pages, a signal page may contain many kinds of information, which even haven't much relativity [4]. Besides, a web page often contains some less attractive

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 431–437.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

information, such as navigation bars, web pages user interface pictures, user interactive areas and contactor information copyright, which delay the searching accuracy and speed. Therefore, divided the web pages into several semantic modules and extract the web pages theme content can highly improve the information engine's performance, as the same, extract the correct information is very important for web pages classification.

2 HTML Content Structure

2.1 HTML Structure

Content in web pages is described as HTML document, which is named html or htm as its extend name. HTML (Hyper Text Markup Language) is used to build web pages, which contain a set of tags to mark the content on internet. It contains two kinds of information; one is the content text itself and the other is tags, which is used to present page's element, structure, and hyperlink. The first edition of HTML is mainly based on text and images sometimes, which is made by HTML110, the second edition of HTML added some features on table process, and users can interact with web pages, it can provides very beautiful background images and form the text as tables. The third edition of HTML added many new important tags which can be used to create page layout, color control, fast image process and font viewer. Developers can create webpages to interact with users, and it became more and more colorful. The fourth edition of HTML is more successful than before, it has been used for more than 10 years, texts and images are treated as object in HTML4 by the control of script language. It is very powerful to dynamic layout and present web pages content. At the same time, with the help of common gateway interface, web pages are more and more excellent in interactive [5]. A common HTML document is showed as follows:

```
<!DOCTYPE html>
<html>
    <head>
        <title>Hello World </title>
    </head>
    <body>
        <p>Hello World!</p>
    </body>
</html>
```

2.2 XML/XHTML Pages

XML is named extensible markup language, which used to markup information and computer can read the marked tags and process the information. XML is designed to transfer information not for present or show the information. We can know what the

data by XML is just like we can see the data by HTML. The first common usage of XML is rich document, which can make its content more colorful. The second usage of XML is metadata, which can describe other data and web information [6].

XHTML is extensible hypertext markup language; it is similar with HTML but stricter on grammar. XHTML is considered to describe HTML by XML, therefore, all XHTML web pages need to declare XML version at the head of file.

3 Researches on Web Information Filter Model

3.1 General Architecture

Since it's important to get and filter web pages information and make it easy for users to get which kinds of information they want. This paper designed web information filter model is based on HTML technology. The design principle is: set the web information filter as a proxy server, which can match the web application needs. We set the system as a proxy server when users surf on pages and then all content of pages can show to users after filter. It's also need to provide a set of settings button for user to customize their needs. The working flow of this system is showed in figure 1.we use HTML parser to transfer HTML pages into structured DOM tree and then depth first traversal the tree for filter.

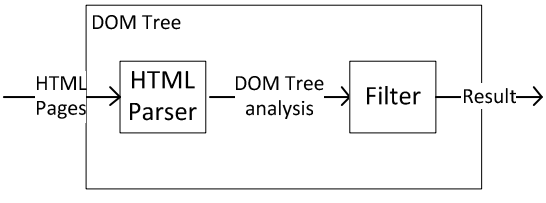


Fig. 1. Working flow of web information filter model

The detail of web information filter model architecture is showed in figure 2.

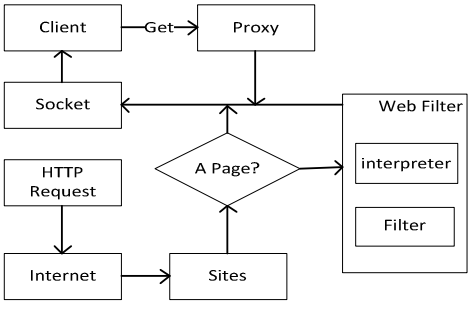


Fig. 2. The Self-adjusting module architecture

The whole system is divided into three main modules:

1. Web proxy server, which is used to handle requests from client and create sockets for web servers, and process results send from server are also transferred to client by web proxy server.
2. Filter module, which can process the structured DOM tree and filter useful information for users. This module will create a new DOM tree for users.
3. Graphics user interface, which is used for customize settings according to different user experience.

3.2 DOM Tree Analysis

DOM tree structure is an important element on researching web pages layout, how to transfer HTML web pages into DOM tree is important in pages content analysis.

DOM is an application program interface for HTML and XML web pages, software engineer can construct, add, edit and delete web pages element by DOM model. All resources present in HTML can be stored, edit by DOM tree. A typical DOM tree is consisted of documents, nodes, elements, text nodes, attributes, and N-level trees.

1. Document: Document in DOM is consisted of layer node objects, which is part of HTML web pages. A document is a node which means this node only contains one element, document interface represent the whole web page, and this node is the root of document tree.
2. Node: node is a common type which used to mark all elements in web pages.
3. Element: element is the most parts in web pages, it represent all resources in web pages except text. Element is related to node, and it contains a attribute.
4. Text node: text node is used to process text in webpages.
5. Attribute: attribute is the property of element.

DOM is language independent, and all DOM interface are under the protocol of interface definition language. DOM aims to bring object oriented concept to HTML/XML web pages.

All web pages can be analyzed by DOM tree. A common web page which is extract from internet is showed as below:

```
<html>
    <head>
        <title>This is a title </title>
    </head>
    <body>
        <h1>This is a document</h1>
          <p>Document content</p>
    </body>
</html>
```

After DOM analysis, the following DOM tree can be realized, which is showed in figure 3.

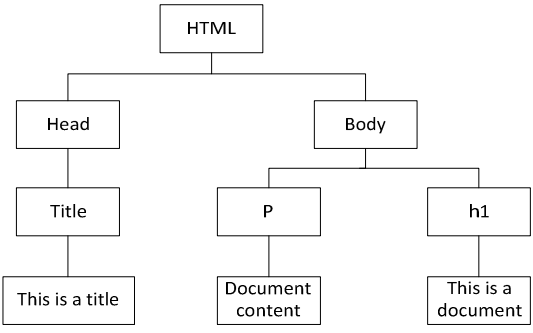


Fig. 3. DOM tree structure

3.3 Content Filter

Content filter module is designed to continue process DOM tree. Each node in DOM tree can be seemed as a partition in visual. In some rich content nodes, such as <table>, <p>, which is large enough and not fit for a signal partition, therefore, it's highly needed to divide large node into small further. And then replace them with the smaller one. We can divide them according to the color, tag, size, shape. Main parts of algorithm of web pages filter are showed as below:

```
Filter_DOM_Tree (pNode, nLevel){
    IF (Filter(pNode, nLevel)==TRUE){
        FOR EACH child in pNode
        Filter_DOM_Tree(child, nLevel+1);
    }
    ELSE{
        Get and Filter the subTree pNode ;
    }
}
```

After DOM analysis and filter process, a common web page would be described as the below filter tree.

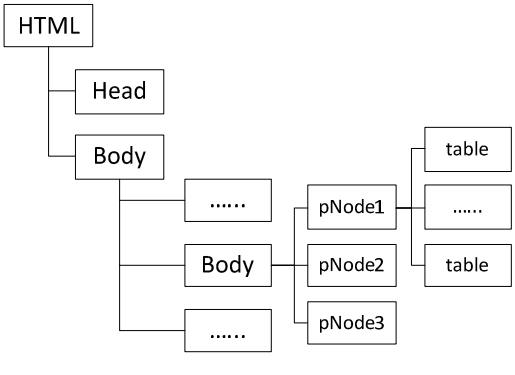


Fig. 4. DOM tree structure

3.4 Experiment Validation

We test the web information filter model in real internet environment. 1028 web pages are collect randomly for test. The content in web pages are divided into email address, hyperlink, text, title, images, media files, navigation.

Recycle rate(Rr) is the percentage of correct results which are analyzed by information filter model system compare to the whole correct results. Correct rate(Cr) means the the percentage of correct results which are analyzed by information filter model system compare to the whole results. Currect rate is smaller than recycle rate normally. F parameter is intruduced to accuracy measure the system. F is calculated by:

$$F = \frac{\left(\lambda^2 + 1.0\right) * Rr * Cr}{\lambda^2 * Cr + Rr} \tag{1}$$

λ is the priority parameter between recycle rate and correct rate. If $\lambda = 1$, these two rates are under the same importance. And if $\lambda > 1$, it means recycle rate is more important.

All sites are divided into six categories, e-commerce, education, news, BBS (Bulletin Board System), SNS (Social Network Service) and entertainment. Test results are showed in table 1.

Table 1. The stimulate result in second test

Pages Category	Recycl e No.	Correct No.	Recycle Rate	Correct Rare	F%
E-commerce	97	94	0.88	0.85	87
Education	78	54	0.87	0.60	76
News	76	73	0.95	0.91	94
BBS	88	84	0.98	0.93	97
SNS	93	79	0.93	0.79	89
Entertainment	45	34	0.50	0.38	45

From the above table, BBS web pages and new web pages are performing better in this paper designed filter model, e-commerce and SNS are also be filtered by the web information filter model quite well. Main parts of pages can get more than 80% F parameter in the filter model.

4 Conclusion

Internet is becoming more and more important in people daily life. According to rapid technology development, web information has become richer and richer. A signal page may contain many kinds of information, which even haven't much relativity. Besides, a web page often contains some less attractive information. Therefore, this

paper designed a new information filter model based on DOM analysis and content filter algorithm. Its general architecture and working flow are presented in this paper, and provide content filter algorithm to filter web pages content effectively. Finally, designed a set of experiment to evaluate the performance of this paper designed filter model, the result shows it works well in main parts of web pages.

References

[1] Hobbs, J., Appelt, D., Bear, J.: A Cascaded Finite-State Transducer for Extracting Information from Natural Language Text. In: Finite State Devices for Natural Language Processing. MLT Press, Cambridge (2002)

[2] Gupta, S., Kaiser, G., Neistadt, D.: DOM-based content extraction of HTML documents. In: Proc. of 12th Int. World Wide Web Conference, pp. 207–214. ACM Press, New York (2003)

[3] McKeown, K.R., Barzilay, R., Evans, D., Hatzivassiloglou, V., et al.: Columbia Multi-document Summarization: Approach and Evaluation. In: Document Understanding Conf., pp. 156–172 (2006)

[4] Kaasinen, E., Aaltonen, M., Kolari, J., Melakoski, S., et al.: Two Approaches to Bringing Internet Services to WAP Devices. In: Proc. of 9th Int. World-Wide Web Conf., pp. 342–348 (2004)

[5] Li, X., Shi, Z.: Innovating web page classification through reducing noise. Journal of Computer Science & Technology 17(1), 9–17 (2007)

[6] Chakrabarti, S., Joshi, M.M., Tawde, V.B.: Enhanced topic distillation using text, markup tags and hyperlinks. In: SIGIR, pp. 59–66 (2008)

Power Flow and Power Flow Calculation Research in Power System

Li Jun

Chongqing Water Resources and Electric Engineering College, Chongqing, China

Abstract. In power system, Power flow and power flow calculation research problem is important with the rapid development of modern power grid, this paper research power flow and power flow calculation principle, and then put forward mathematical model: PQ node, PV node, balanced node. To control and calculate power-flow, we have three solutions equations: Gauss elimination, Jacobi and Gauss-Seidel, Newton-Raphson, in these solutions equations, Gauss-Seidel and Newton-Raphson applied extensively. In the power system stability of economy, has an important significance.

Keywords: power system, Power flow, calculation equations and research.

1 Introduction

With the rapid development of modern power grid, equipment and the corresponding auxiliary system of power generation, power transmission, power transformation, power distribution, power utilization the enlargement of interconnected electric power system and the increasingly complex electric power system structure[1], power flow and power flow calculation are significant in power system. The load-flow included the representation of components to extensions itself such as high-voltage direct-current (HVDC) transmission lines, if we want a better methods for power-flow in loss calculation, must solve the optimum power flow (OPF) and state estimation problems , the power-flow continues , and determines the constrains-pluses , of course we have a better ways to visualize and present the load-flow results, It called power flow calculation [2].

2 Power Flow and Power Flow Calculation

From generation to load, power flow through a transmission network that can create many questions, so it is essential to computes the real and reactive power-flow on every transformation or transmission line, along with associated buses voltages (in other words, the voltages at each node). Since networks containing tens of thousands of buses and branches, power-flow calculations need a lot of numbers.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 439–445.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2.1 Power Flow

In power engineering, power flow refers to the power grid voltage everywhere (including amplitude and phase Angle), active power, reactive power steady state distribution, so to operate scheduling units and maintain department we should know the distribution of power flow in power grid. The energy is injected into the power network system, and to access the load in getting (match) power station, in the mean time. [2].

2.2 Power Flow Calculation

Power flow calculation system mathematical model mainly includes: (1) PQ node. For this kind of node of the PQ node active power P and reactive power Q is given, node voltage (V,δ) is to stay for quantity [3]. Some power plants in the system that power must be sent out in fixed time, the plant also as a bus PQ node. Most of the power systems under this type. (2) PV node. This kind of node of the operation parameters are given for the active power P and voltage amplitude V, stay for volume is accepts the reactive power Q and voltage vector Angle θ. Usually choose to have certain reactive power supply power plant as the bus PV nodes, when substation has reactive compensation equipment, also can be used as a PV nodes processing. (3) Balanced node. In load-flow distribution work out previously, the network power loss is unknown, so the network at least one of the active power P nodes can not be given, the nodes bear the active power balance of the system. In addition, must select a node, appoint the voltage phase is zero, as a calculation of node voltage phase reference, this node called benchmark voltage, its amplitude is given. The numerical methods of solving calculation of flow calculation are basic requirements: fast-calculation speed, take up less storage capacity, good convergence, method is simple.

3 Direct and Iterative Solutions Equations in Power Flow

3.1 Direct Solutions to Linear Algorithm Equations : Gauss Remove Method

In power system, There are many unknown parameters in power flow's question. We can solve these unknown parameters by solving the following set in linear algebraic equations in matrix format:

$$\begin{pmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \quad (1)$$ and $Ax=y$ (2). There x and y are N vectors and is an $N \times N$ order matrix. These parts of x, y and A may be real or complex. A and y are given x is an unknown parameters that we want to solve. If det(A) is nonzero parameters, we can get a unique solution to (1) equation. And also we

x can easily obtained the solution when A is an upper triangular matrix and then nonzero diagonal elements was existed, we can obtain

$$\begin{pmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \quad (3),$$

because x_N was involved only in the last equation (3), $x_N = \frac{y_N}{A_{NN}}$. If x_N was computed, we can solve the next-to-last equation $x_{N\text{-}1}$, lastly, when x_N, $x_{N\text{-}1}$,…, x_{K+1} already was computed, we also can obtain the K_{th} equation . Generally, This process for solving (3) is named back substitution. If there is not an upper triangular for A, by using an upper triangular matrix we can transform to an same value equation. This procedure, can be called Gauss remove method, we can solve it by the following steps. During steps, the first equation can be used in (1) to elimination x_1 from the other equations. So be it, A_{n1} / A_{11} can multiplied equation 1 and subtracted from the last one equation. When we finish this step, we have the following form $A^{(k-1)}x = y^{(k-1)}$. Parameters k of these equations was triangularized and left unchanged. Then, $A_{nk}{}^{(k-1)} / A_{kk}{}^{(k-1)}$ can multiplied Parameters k and then subtracted the last equation, at last we obtain the equivalent $A^{(N-1)}x = y^{(N-1)}$, and then $A^{(N-1)}$ is upper triangular.

3.2 Iterative Solutions to Linear Algebraic Equations: Jacobi and G-S

For solving these many unknown parameters in power flow of power system. These iterative solution to (1) proceeds can be like this. Firstly, we can select an initial value $x(0)$. Then by using $x(i+1) = f[x(i)]$ (4), i =0,1,2,… and $x(i)$ is the ith value and f is an N vector functions that explain the solving method. By solving the process until we obtain the following stopping condition :

$$\left| \frac{x_k(i+1) - x_k(i)}{x_k(i)} \right| < \zeta \quad (5),$$

all k=1,2,3…,N, all these parameters are for two specific explain methods: Jacobi and G-S. We can obtained the Jacobi method from considering the kth equation of (3.1.1), like this equation:

$y_m = A_{m1}x_1 + A_{m2}x_2 + ... + A_{mm}x_m + ... + A_{mN}x_N$ (6),m=0,1… by obtaining x_m. They can use the $x(i)$ "old" values at i on the right side to get the "new" value $x_k(i+1)$ on the left side. That is

$x_k(i+1)=\frac{1}{A_{kk}}[y_k-\sum_{n=1}^{k-1}A_{kn}x_n(i)-\sum_{n=k+1}^{N}A_{kn}x_n(i)]$ (7), k=1,2,3…N, The Jacobi type given by (7) can write into this matrix format: $x(i+1)=Mx(i)+J^{-1}y$ (8),where $M=J^{-1}(J-A)$ (9). For Jacobi, J contains the matrix A diagonal elements. G-S and Jacobi is not similar except that during each repetitiveness, the "new" values, $x_n(i+1)$, for n<k the right side values of (7)are used on to get a "new" value $x_k(i+1)$. For Gauss-Seidel, J is the lower triangular ration of A.

3.3 Iterative Solutions to Nonlinear Algebraic Equations: N-R

Solving these many unknown parameters in power flow of power system, we have an another equation: N-R nonlinear algebraic equation, its matrix format is given by $g(x)=\begin{bmatrix} g_1(x) \\ \vdots \\ g_N(x) \end{bmatrix}=y$ (10), There x and y are N vectors and $g(x)$ is an N vector's functions. If we have the value y and $g(x)$, and we will get the value of x. There is an nonlinear equations like: $x(i+1)=x(i)+R^{-1}[y-Ax(i)]=R^{-1}(R-A)x(i)+R^{-1}y$ (11), there R is also $N\times N$ order matrix, which is similar to the Jacobi and G-S methods, for nonlinear equations, the matrix R must be computed.

We have one method to appoint R, called N-R method, it is based on the Taylor series expansion of $g(x)$ an practice point x_0.

$y=g(x_0)+\left.\frac{dg}{dx}\right|_{x=x_0}(x-x_0)...$ (12). If we neglect the higher order terms, we can solve the value of x. $x=x_0+[\left.\frac{dg}{dx}\right|_{x=x_0}]^{-1}[y-g(x_0)]$ (13). In the N-R method, we can use the new value $x(i+1)$ in (13) replaces the old value $x(i)$.

Thus $x(i+1)=x(i)+R^{-1}(i)\{y-g[x(i)]\}$ (14).The $N\times N$ order matrix $R(i)$, whose elements are the partial derivatives shown, is called the Jacobian matrix, The N-R method is similar to the Gauss-Seidel, except that R in(11) is replaced by $R(i)$ in (14).

4 The Control and Solutions of Load-Flow and Load-Flow Calculation Research

Power-flow and power flow calculation research method , further development will make more rapid flow calculation, convergence better, stronger adapt to the system the operation condition of ability, and in some special fields such as optimizing load-flow and online power-flow can get actual application.

4.1 The Control of Load-Flow

There are some means can be used in controlling of power-flow in power system: controlling the prime mover and excitation; capacitor banks and reactors can be switched; controlling the regulated transformers. There is a simple model shown in Figure 1, from the figure, the generator current is

$I = \frac{E_g e^{j\delta} - V_t}{jX_g}$ (15), $S = P + jQ$ (16), the real and reactive powers are delivered: $P = \mathrm{Re}\, S = \frac{V_t E_g}{X_g} \sin\delta$ (17), $Q = \mathrm{Im}\, S = \frac{V_t}{X_g}(E_g \cos\delta - V_t)$ (18).

V_t -final generating voltage, E_g - excitation voltage, δ -power angle, and X_g - positive-sequence reactance. The δ change a little, the load-flow that increase in reactive power Q .The power angle δ increases the real power P also increases. If we increases the reactive power output Q, we can get the conclusion that excitation voltage E_g increases [2]. .

4.2 The Solutions of Load-Flow and Power Flow Calculation Research

(1) Power-Flow and Power Flow Calculation Research by Gauss-Seidel

Because in load buses power-flow buses data contain P_k and Q_k , in voltage-controlled buses power-flow buses data contain P_k and V_k , linear equation format can not be satisfied by the noting equations $I = H_{bus} V$,current vector I is unknown ,so the equations are nonlinear exactly. For we can obtain I_k from $I_k = \frac{P_k - jQ_k}{V_k^*}$ (19) in each load bus, if I_k was given, we can obtain

$V_k(i+1) = \frac{1}{H_{kk}}[\frac{P_k - jQ_k}{V_k^*(i)} - \sum_{n=1}^{k-1} H_{kn} V_n(i+1) - \sum_{n=k+1}^{N} H_{kn} V_n(i)]$ (20), for a voltage-controlled bus, Q_k is unknown, however, we can get the values from the equations like:

$$Q_k = V_k(i)\sum_{n=1}^{N} H_{kn}V_n(i)\sin[\delta_k(i)-\delta_n(i)-\theta_{kn}] \quad (21).$$

(2) Power-Flow and Power Flow Calculation Research by N-R

There is a nonlinear equation $y = g(x)$, can be solved by N-R, For the load-flow calculation research we can definie the x, y, and g vectors as :

$$x = \begin{bmatrix} \delta \\ V \end{bmatrix} = \begin{bmatrix} \delta_2 \\ \vdots \\ \delta_N \\ V_2 \\ \vdots \\ V_N \end{bmatrix}, \quad y = \begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} P_2 \\ \vdots \\ P_N \\ Q_2 \\ \vdots \\ Q_N \end{bmatrix}, \quad g(x) = \begin{bmatrix} P(x) \\ Q(x) \end{bmatrix}. \quad (22)$$

$$y_k = P_k = P_k(x) = V_k\sum_{n=1}^{N} H_{kn}V_n\cos(\delta_k-\delta_n-\theta_{kn}) \quad (23)$$

$$y_{k+N} = Q_k = Q_k(x) = V_k\sum_{n=1}^{N} H_{kn}V_n\sin(\delta_k-\delta_n-\theta_{kn}) \quad (24)$$

For each voltage-controlled buses-link, we know magnitude V_k already, the function $Q_k(x)$ can be neglected. Select case information to see the initial values, we can view the initial matrix in case information of power-flow.

5 Conclusion

This article emphasizes the theoretical research from the formula and principle aspects. Especially, when we control the solutions of power-flow, when we compute the load-flow ,we applied the Gauss elimination, Jacobi and G-S, and N-R to research this calculation method. In wide range to effectively control the power flow are important to power grid and to maintain bus-link voltage. The power system subjected to various disturbances was also effective to the nonlinear simulations [5]. In the power system stability of economy, has an important significance. And then, artificial neural network technology and Genetic series function, and fuzzy series also have an important role in the power-flow calculation gradually.

References

[1] Sun, G.-Q., Wei, Z.-N.: Power System State Estimation with Unified Power Flow Controller. In: DRPT 2008, April 6-9 (2008)
[2] Glover, J.D., Sarma, M.S.: Power system analysis and design. China machine press, China (2004)
[3] Niu, H.-S.: Power system power flow calculation and simulation analysis. Modern Business Trade Industry (16) (2009)
[4] Zhao, B., Guo, C., Cao, Y.: Optimal Power Flow Using Particle Swarm Optimization and Non-Stationary Multi-Stage Assignment Penalty Function. Transactions of China Electrotechnical Society 19(5) (May 2004)
[5] Parimi, A.M., Sahoo, N.C., Elamvazuthi, I., Saad, N.: Transient Stability Enhancement and Power Flow Control in a Multi-Machine Power System Using Interline Power Flow Controller, 978-1-4673-0136-7/11/$26.00 © (2011) IEEE

The Analysis and Design of Human Resources Information Sharing System in University

Jiang Yan

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
jxncjiangyan@126.com

Abstract. Human resources is one of the most important resources in universities. Making full use of human resources information for leadership to provide decision-making services, can effectively improve the level of running a college. In this paper, human resources information sharing system in university is analyzed, and in light of the specific situation of university, the demand analysis is done for the human resources information sharing system in university, the preliminary design of the system is done by using the object-oriented method theory and modeling tools.

Keywords: human resource, information sharing, decision support, object oriented, design analysis.

1 Introduction

In the process of daily management in the university, human resources mainly refers to teaching staff that teachers are the main body, the human resource information does not only refer to the personnel data managed by university, but also is largely consisted by the teachers' teaching, scientific research and other work information. With the rapid development of science and technology, the traditional human resource management system has been unable to meet the demand of social development. Human resources information sharing system in university, improves the efficiency of human resources management and perfects the original daily management. In addition, the human resources information sharing system in the process of establishing, is mainly based on personnel system, connected with teaching system, research system and foreign affairs system, so that it is a unified whole system in the management process, in promoting the smooth flow of information, can achieve the purposes of sharing information. In this paper, the following analysis is done towards problems of human resource information sharing system in university from needs analysis, problems, outline design, the new structure and characters, relationships between modules.

2 The Demand Analysis of Human Resources Information Sharing System in University

With the rapid development of social economy, colleges and universities as the base of talents cultivation in China, is not only about our talents training, but also related to

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 447–452.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

China's social progress and development. In the process of social development in twenty-first Century, universities should actively advocate "people-oriented" development strategy, in daily management, should give full play to people's enthusiasm, initiative, creative, try to tap the human potential. If you only use the old management pattern, it is unable to keep up the pace with the society development, affects severely to the daily management. Along with the rapid popularization of the computer network, it is changing people's lives and promoting social progress. In human resources management in universities and colleges, combined scientific management method with IT technique and applied to human resources management, not only improves the human resources management level, but also promotes the development of colleges and universities, enables it to keep up with the pace of social development.

In the whole process of the establishment of human resources information sharing system in university, the management system is different because of management process, nature, management scope, scale. In view of software design in human resources information management in the current market, the majority towards the top enterprises, minority towards colleges and universities, in the process of selection, which requires the combination of university management, but also needs to select suitable for the software, regardless of the cost of investment or system construction, university is much more difficult than enterprise. In addition, colleges and universities as personnel training and talents gathering place, whether from the humanistic quality requirements or from knowledge level, its requirements must be higher than society enterprises, and based on the universities education characteristics, personnel management is also very different. Thus it should not blindly copy the enterprise management software in using process, rather combine with the university's own development, to establish a suitable digital management platform, to meet the needs of scientific management and promote the development of colleges and universities construction.

With the rapid development of social economy and IT technology, intelligent tools are instead of humans to solve some problems that human can not solve in a certain extent. Intelligent tools in the using process, have certain advantages, but largely depend on human management. In the face of the current human resources management in universities in China, most colleges are still using the decentralized management model, and management system is mainly campus network, LAN and single microcomputer in the management process. The system in use process, are mainly independent system operation, rarely exchanges information between each other. At the same time, these systems are not set corresponding data exchange program, also cannot realize information sharing. Therefore, to establish a college human resource information sharing system, plays a very important role in the daily development.

3 Problems of Human Resources Information Management System

In twenty-first Century, the rapid development of science and technology, the traditional management model has been unable to meet the development of our

society, blocks development of colleges and universities, also indirectly affects China's personnel training. Related questions of human resource management information system mainly include the following aspects:

The Self Closed System. In the current management system in university, the biggest problem is that all systems are working independently, without affecting each other, so the automatic information conversion and sharing, does not achieve the related requirements. Transferring the related information needs frequently to switch back and forth among all systems, it wastes a lot of time, and the result is not ideal. In addition, although administrative staff in universities has collected a large number of business data, the data is not processed and finished, so it brings information, also brings a lot of junk information.

A Large Amount of Information Maintained. Facing the increasingly fierce social competition, all businesses in the system design process, usually use short, flat, fast design scheme, this scheme in the use of the process, can save time, but this approach lacks survey and overall optimization, causes a lot of resources not to be used normally, this wastes manpower, material and financial resources, at the same time, still largely increases redundancy. Once a system problem occurs, it needs related person to maintain, so it increases maintenance costs. In addition, the system has its own C / S structure, it needs to maintain the server and client in the using process.

Insufficient Information. In the human resources information sharing system in university, management and decision makes as a unified whole, cannot be separated in use process. However, people in the development of the system, often neglect the importance of the problem. In human resources management in universities and colleges is no exception. This directly causes information inadequacy in human resource information management in university, which can not timely provide management information to the management staff, which blocks daily management of colleges and universities, directly affects the future development of colleges and universities.

The Lack of Humanity. In daily management, human resource information management not only relates to the college teaching order, but also directly affects the future development of colleges and universities. However, in the current human resources information management system, the design symbolic collects and manages information date, doesn’t do any reminder for the management flow and scale. For example, a teacher's tenure is over and needs for renewal, system does not prompt remind managers, it can only rely on manager's memory, once forgot, it will directly impact the working state and emotion on both sides.

4 The Design of Human Resource Information Sharing System in University

In the design process of human resources information sharing system in university, it make full use of the campus network environment, effectively exchange human

resources information in the original basis, so that it can make information smoothly normal, timely, accurately transfer within the university scope, avoid loss of information, also strengthen the school management system in a certain extent, provide reliable basis for the school leader to do next decision. The greatest advantage of this scheme in the use of the process is the small design workload, to maximize the promotion of information exchange, sharing in the information exchange process, at the same time, related functions be able to give full play to their own use, no matter in the information expansion or statistical functions, higher than the original system. Thus it is favored by university.

Information exchange and share system as the core part of human resources information sharing system, management system can make process in time by their own analysis. Information unit in the use process, can service to the entire information exchange platform system, and information exchange system in the process of renewal, is through specific information unit to complete own maintenance function.

In the university human resources information sharing system, various information units in information sharing system will be based on their function and management, store consistent structural information and content, and exchange information with other units through their own interfaces, in order to realize information exchange and accelerate the information sharing platform construction. In view of human resource information sharing system, information units as important part of whole system, not only provide corresponding management information, but also largely relate with system information collection and overall. System is mainly through the information units to gather information, and the information unit has independent information collection structure in the operation process, so the operation process can be run independently, without mutual interference, to ensure information record accurately and support the transfer information timeliness and completeness.

5 Design and Features of the New System's Structure and Relationships between Modules

In the whole design process of the system's structure, we should fully consider the administration, the scale, direction and management goals and many other aspects of the university and develop a suitable system for university to reflect the characteristic of a university and promote university's development. In view of the current old management system in university, there are two main management scheme, the first scheme retains relevant database content on the basis of the original system; the other links the old system to the new system, the new system reorganizes and divides the original old system, and expands the system according to the system's needs. This model can optimize basic data in a certain extent, reduce redundancy, ensure the concise data structure, expand data timely.

The system maintenance module mainly includes the following aspects: firstly, it realizes the information sharing, completes cross library statistics and query. On one hand it makes the original data to be fully used, on the other hand it can save money, shorten the development cycle. Secondly, it provides seamless succession with the Word2000 and Excel2000 tables and documentation tools. Thirdly, compared with the traditional human resource management information systems, human resources

information sharing system can fundamentally reflect the humanistic management. On Teachers' Professional Title Evaluation and of related management events in colleges and universities, the system can remind management staff advancely to be ready timely. Finally, different with the traditional information system, highly modular in the use process, along with the university human resources management needs, can expand their range of administration at any time, not subject to time and place restrictions in the adding data process, it fundamentally avoids the insufficient information and resources waste.

Throughout the system design process, statistical analysis is one important part of the system, could do prediction and assistant for the next step of management model according to the relevant information, and carry on the real-time monitor in the information management process. Management personnel can use statistical procedures of the system to statistic and analyze all types of data in university management while operates the system, which improves their own quality management and their management level.

6 Conclusion

With the rapid development of social economy, colleges and universities as one of the most important parts of social construction in China, not only has relation with the rapid society development, but also has relation with Chinese personnel training. The design of human resources information sharing system in university improves human resource management quality and largely promotes the development of colleges and universities. This requires colleges and universities to design a suitable management system according its scale and management in the using process, only in this way, can we play the system's practical function, can we fundamentally promote the development of colleges and universities.

References

[1] Liang, Z.: The application situation and prospects of data mining technology on human resource management. Modern Computer (Professional Edition) (2) (2009)
[2] Liang, Z.: The application of UML modeling method based on RUP in personnel recruitment system. Modern Computer (Professional Edition) (6) (2009)
[3] Chen, L.: The design and study of E-HRM system in digital campus platform. Journal of Southwest University of Science and Technology (Philosophy and Social Science Edition) (4) (2008)
[4] Wang, J.: The research of human resources management system in the university. Developing (9) (2008)
[5] Hao, Y., Wang, W., Gao, X.: The analysis and design of human resources management information system in university. Chinese Education Informatization and High Vocational Education (6) (2008)
[6] Huang, W.: The research of framework on human resource management system. Journal of Jilin Teachers Institute of Engineering and Technology (10) (2007)

[7] Liang, L., He, Y.: Discussion construction of China's civil servant flow management information system. Science & Technology Progress and Policy (5) (2007)
[8] Chen, X., Shang, J.: The development of human resources management information system (HRMIS). Journal of Chongqing University (Natural Science Edition) (4) (2006)
[9] Gao, F., Zong, Y.: Construction exploration of agricultural cooperation economic organization ERP. Journal of Shanxi Agricultural University (Social Science Edition) (3) (2006)

The Application Research of Multimedia Database in Innovative Venturing Course

Wu Binying, Chen Shiqiang, and Fang Li

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
wu_binying@126.com, chen_shiqiang88@126.com, fangli818@163.com

Abstract. With the development of internet and communication, people's living becomes rich and colorful. They are not satisfying the past character information and statical information browses. On the contrary, they are more likely to enjoy the video, audio, and other multimedia resources. Multimedia database is the combination of multimedia technology and database technology. This article introduces the basic principle, existing question, and key technology of multimedia database. Moreover, combine with the practice, the writer designs one multimedia database that using in the undergraduates' innovative venturing education. The multimedia database obtains the preferable effect during the practical application.

Keywords: multimedia database, multimedia information search, DBMS, OODM, innovative venturing course.

1 Introduction

At present, the internet has infiltrated into people's living with the computer technology development and network popularity. The primary target of the internet provides convenience for the information communication between scientists and only transports the simple texts. Following with the network development, the internet has been applied in more and more fields. Moreover, people have the increasing requirements on the internet such as online shopping, information consultant, interaction, the entertainment of audio and video and so on.

It is very clear that people's requirement of the internet is not only limited in the static information browse such as character and image. They are more likely to enjoy the huge number of video, audio, and other multimedia information. All these can make people's living and learning become more rich and colorful. The electronic resource can save the character, image, graphic, sound, video and many other information forms into the disc, light disc, tape and other storage medium through the binary data. Moreover, the resources can reappear by the network communication, computer, or the terminal. As the new type of information carrier, the electronic resource has the large amount of information, vivid content and various information representation forms.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 453–459.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

In the network environment, the educational resource includes various teaching materials, teaching coursewares, and education software. Its present form is character, image, graphic, video, audio, and other multimedia forms. Establish the abundant and superior education resource database can provide the favorable platform for teachers' teaching and it is helpful for the independent learning and exploration study of students. Moreover, the establishment of these education resource pools must depend on the database technology.

During the students' innovative venturing education, we need to take full advantage of comprehensive multimedia technology and database technology to create the education resource pool. It is benefit for the students' development and training of the self-dependent innovation. The establishment of education resource pool that based on the multimedia database technology can provide rich and colorful learning content to the students. However, the increase of multimedia resource number will bring the management difficulties at the same time.

The practice proves that designing the multimedia database reasonable can effectively manage the huge number of educational resources.

2 Multimedia Database

2.1 Multimedia Information and Multimedia Data

Multimedia information means the information transmission media is not limited in the character. It comes down to image, graphic, sound, video and other multiple media. The multimedia information can supply sense organ through sound, light, electricity, and other multiple ways. By this way, people will be apt to accept and have a deeper impression about the information.

Organic integrate the multiple media such as figure, text, image, graphic, sound and video. Moreover, save these data into the computer. This is the multimedia data. In these data, there has the formatted data such as figure, character and the non-formatted data such as text, image, graphic, sound and video. The characteristic of multimedia data is the various expression forms among each data. The relations among different data are very completed. Besides, the multimedia dataset has the huge data quantity.

2.2 Multimedia Database and Its Data Model

With the rapid development of multimedia technology and the support technology, the multimedia application is more and more wide. Therefore, the high efficiency of multimedia information management is urgent. For the various expression forms among the different multimedia data, the relationship is very complicated with the large number of data quantity. The traditional database ability is not equal to the multimedia information management and process. Therefore, the multimedia database arises at the historic moment.

Multimedia database is the integration of many multimedia objects. These data objects are organized under the relative way that can share with other applications.

Multimedia database can effectively achieve the multimedia data storage, read and search. It can carry forward some advantages of the traditional database. Moreover, it can synchronize and manage the data, which has the space-time relationship. Data model is the core center of database. In the totality development, the data model of multimedia database can divided into 3 types: ① relational data model; ② OODM (Object Oriented Development Method); ③ expanded relational data model.

The relational data model is under the theory basement of relational algebra. It can perfect manage the traditional structured data after the development. However, the multimedia database has so many images, graphics, video, sound, and other unstructured data. These data structures are complicated that cannot expressed by the relational model. Therefore, the relational data model is not suitable in managing the above data.

Object Oriented Development Method is convenient for the expression and operation of unstructured data. However, the technology is not as mature as the relational data model. There still have some questions in the theory research and application exploration. We need to rewrite the code with huge development work and long term. For these limitations, the multimedia database that explores by the present used OODM is applying in the special subject. This is regard with the development project of multimedia database system.

Apply OODM in the exploration is not suitable for the technology and economic condition. Use the extended relational data model in the exploration of multimedia database system is the most mature method. Bring the thinking of OODM, hypertext and hypermedia into the traditional relational data model, we can solve the questions that image, graphic, sound, video and other unstructured data cannot expressed by the relational model.

2.3 The Existing Questions in the Multimedia Database Design

First, the information media variation. This is not only the numerical data and character data. It will expand into the storage, organization, usage, and management of multiple media data. These multimedia data have huge quantity and great differences among each media. We select the reasonable logic structure and physical structure to do the reasonable organization. There must ensure the ordered and effective information storage and the rapid search.

Second, solve the multimedia data integration or expression integration. Achieve the cross-call and fusion among the multimedia data. Integrate the more fine grading will obtain the stronger multimedia integration with wider applied value. Each kind of multimedia data type needs to support some standard operations except the basic operation and function, reasonable data structure and storage method. The standard operations include the various universal operations and integration of different new types.

Third, the interaction between multimedia data and people. The user interface of multimedia database cannot be described by one simple table. For example, the description of media content, the space description and the time description. The multimedia requires the new method to browse, search and realize the content that

users can convenient describe the search requirement and find the relative data. In many cases, facing at the multimedia data, some users even do not know what they are searching for and how to describe the research. Therefore, the multimedia database requires the user interface not only accept the user description. It is necessary to help the user to describe the thinking, find out the required content, and display it on the interface. The research of multimedia database is not only the traditional table, it also expresses the abundant multimedia information even the result that integrated by the computer.

3 Key Technology of Multimedia Database

3.1 Establish the Multimedia Database on the Relational Database

The simple structure of relational database is the single two-dimensional table. The date type and length is limited in the small subset. It cannot support the new data type and data structure. Moreover, it can hardly achieve the space data and temporal data and lack of deduct and reasoning operation. Therefore, the function of expressing data characters is limited. Use national model in the multimedia database system, there must to expend the present relational model. It can support the formatted data and manage the non-formatted data. The expansion technology for the relational database model has three types.

(1) Combine the file systems in the relational database management system and operation system. Achieve the non-formatted data management.
(2) Integrate the formatted data and non-formatted data in the relation tuple to form the complete tuple and save it at the data page or in the page groups.
(3) Divide the formatted data and non-formatted data into two parts. One is the formatted data and the other is the reference of non-formatted data.

3.2 Establish the Multimedia Database Based on OODM

Different from the expanded relational database system, the OODM database starts from the data model. Reconsider about the entire structure that different with the traditional DBMS system, the storage structure of object class hierarchy, storage method and inheritance achievement, user definition of data type and method management, the necessary version control and friendly user interface. Build the completely new DBMS.

The database structure of OODM:Based on the system model function, design the relative system structure is the important part about the achievement of OODM management system. The present OODM system functions are various and provide the different system structures.

Storage structure and access method about the OODM system: The OODM system is managing the integrated objects of multimedia that saved on the disc. Therefore, design the effective storage structure and multimedia data storage is the most important question. From the present model discussion of oriented object system, we

can see the present storage structure application can divide into two types. One is the storage structure that based on the present relational model. The other one is the new design method that more correspond with the multimedia object character.

4 Multimedia Database Application in the Innovation Venturing Education

4.1 Multimedia Database Designs

The multimedia database is combined with main troll model, login control model, user information management model, and multimedia database.

The main control model as the main surface of system requires better visual effect. The login control model can improve the data security in the multimedia database. Different users have the various permissions about the multimedia database. The system key model is the resource management model. It includes multimedia information search, multimedia information addition, multimedia information delete, multimedia information browse, multimedia information download, and other functions. The user information management groups the users with different permissions. The multimedia database stores the multimedia information.

4.2 Apply Multimedia Database System in the Innovative Venturing Education

The writer utilizes multimedia database system in the college teaching of innovative venturing course that obtained the perfect effect. The main expressions are in the following aspects.

(1) Utilize multimedia to establish the scene that can encourage the learning interest and develop the students' learning effect. The object-teaching mode of multimedia technology can promote students to image in the cubical space and input into the new learning process as quickly as possible. The strongly recommend image and text, both improvement of sound and image, the representation combination of dynamics and statics can mobilize the positivity and initiative of teachers and students. Moreover, it can optimize the teaching process and education effect. The visual, interesting, visible, and dynamic teaching content guide the students to initiative learn and increase the training of innovative ability. The integration of character, sound, image even the animation of teaching content can vividly and directly create the detailed environment that textbook cannot supported. Students will greatly acknowledge the innovative venturing textbook that can effectively achieve the teaching target of innovative thinking training.

(2) Utilize the multimedia technology, abundant lesson planning resources, renew the lesson planning method, and organize the class management. Under the traditional mode, the lesson planning is writing (write down the relative contents about the teaching plan). Teacher is the leading of the whole lesson (teacher speak, write, question and answer). It is greatly restrain the lesson preparing. Under the multimedia environment, the innovative venturing education has expended and the teacher choice becomes wider. Moreover, the self-judge for teachers becomes more and promotes the

positivity. It is true that lesson planning cannot leave the write. However, we have the new method such as download, copy, electronic text input and so on. The class can lead with the teacher, it can also lead by the media (multimedia can speak, broadcast with sound, light, and color that can partly replace the teachers' labor). The participation of multimedia leading can regulate the interest between teach and learn. In addition, it can optimize the class structure.

(3) Utilize the multimedia technology to replenish and expand the listen, speak, read and write is the basic circle of innovative venturing education. The traditional method is using the blackboard, textbook, and the students' note. These kinds of education information express the serious teaching stylization. Bring the multimedia technology into the innovative venturing education we can use the media to play the courseware, short file and so on. The successful cases and passion speech by the venture success person will arise the strong interesting of the students. It can lead the students' positive association and bring the resonation of thinking and emotions. The influence aspect includes slide show, project, and computer. They can play some flash or animation, lead students to do the visual, listen test, and train the imaginative thinking. Then stimulate their association and create the venture thinking and behaviors follow the stimulus information implication. Some music can let them calm down and focus on the positive association. The images can stimulate the imagination, and lead them thinking with innovation. Otherwise, the electronic writing is different from the traditional writing with the character of abundant resource, strong interactive, simple storage and modification that students are interested in it.

5 Summary

In order to apply the multimedia database system into the innovative venturing education, we can establish the relative independent learning environment for the students through the multimedia technology. This method can support every student the fully chance to communicate with teachers and friends. Moreover, the students can obtain the vast quantities of information on the network and they can learn with each other and help each other through the network. For the teachers, multimedia technology is the channel of information transmission. It can increase the education information capacity, expand the teaching methods, improve the education efficiency, and train the cooperation learning ability among students. From the above, the multimedia database technology has the enormous positive function for the innovative venturing education.

References

[1] Li, Y.: Multimedia database technology, pp. 1–115. China Machine Press, Beijing (2004)
[2] Feng, Y.: Research Dynamic of multimedia database. China Computer World (3), 12–13 (1997)

[3] Wang, S., Xiao, J.: The multimedia MIS functional requirement of MDBMS and OODBMS support to the multimedia MIS. Computer Engineering and Applications 33(10), 51–54 (1997)
[4] Zhong, Y., Li, S., Lin, F., Ran, J.: Multimedia computer technology, pp. 56–78. Tsinghua University Press, Beijing (1993)
[5] Zhang, B.: Keep pace with the times—realize the service mode innovation of news information. The World New Service Research (2), 15–16 (2002)
[6] Li, X.: The brief evaluation of key problems in the multimedia database. Network Application (2006)

Design and Realization of WCDMA Project Simulation System

Yehui Liu and Huaiqun Wang

Beijing Polytechnic College, Beijing China 100042

Abstract. WCDMA is the most widely used standard of 3rd-generation (3G) mobile communication in the world. The construction of WCDMA wireless network is complex. In order to meet the needs of the network construction in China, as well as the simulation teaching of Communication Profession in college, we designed and developed the software, "WCDMA Project Simulation System". The software uses geographic information technologies and raster algorithm, and analyses the main processes of wireless network construction, stimulating the key processes such as site location selection, coverage planning, and capacity planning and so on. This software can not only assist the network planning engineers in their works, but also be used in teaching of Communication Profession in a clear way, so it is with a high practical value.

Keywords: WCDMA, Simulation system, 3G.

1 Problem

The third-generation mobile communication technology (3G) currently has three standards, namely the WCDMA (Wideband Code Division Multiple Access), CDMA2000 (Code Division Multiple Access 2000), TD-SCDMA (Time Division-Synchronous Code Division Multiple Access). Among them, WCDMA, formulated by 3GPP, is a 3G mobile communication system standard based on GSM MAP core network and taking UTRAN (UMTS Terrestrial Radio Access Network) as wireless interface. It is the most widely used international 3G standard, accounting for more than 80 percent of global market share.

The construction of WCDMA wireless network is complex. Generally speaking, the construction of a wireless network consists of three main steps: capacity planning, quality planning and coverage planning; link analysis and network planning and design; interference analysis and optimization. In the teaching process of Communications Profession, it is difficult to clearly tell the trainees the complete process of system design by simply using teaching materials, but if you can take advantage of the system project simulation software, you can not only reasonably and orderly display the various complex devices and parameters in the intangible and invisible network, but also lively show a lot of invisible concepts and freely change the virtual devices and the environment on the site, thus greatly enhancing teaching effectiveness. Therefore, developing a set of WCDMA system project simulation

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 461–466.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

software and applying it in the teaching and training of Communication Profession at a low-cost and low-risk way is of great practical significance and application prospects, and also plays a guiding role in live WCDMA system project practice.

2 Structure and Function of the Simulation System

In order to comprehensively simulate the construction process of wireless network, the WCDMA system simulation software is divided into five independent functional modules, namely network planning, computer room construction, device operation, signaling process analysis, and network optimization. In the network construction process, key steps such as base station site location selection, coverage and capacity planning, network basic topology definition, network element design and connection configuration, antenna feeder system selection, signaling analysis simulation, and drive test optimization are allocated in the five modules in accordance with the project progress, as shown in Figure 1:

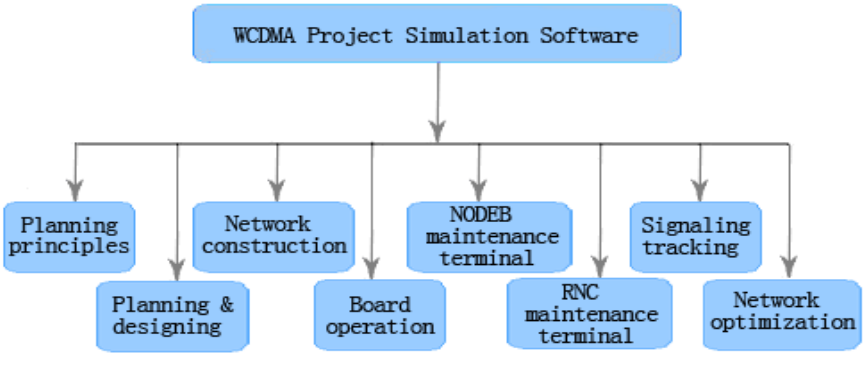


Fig. 1. Software function structure chart

3 Software Key Technologies

Geographic information technology (GIS) is a core technology of this software. The targets processed and managed by GIS are data of a variety of geospatial entities and their relationships, including spatial location data, graphic data and attribute data, etc, which are used to analyze and handle various phenomena and processes distributed in a certain geographic area.

During the base station site location selection, the network base station data (including cell data) has become part of the geographical environment due to its relatively fixed location, so it can be seen as a point feature in the geographical environment. Therefore, geospatial information of the base station can be stored in the form of point geometry, and non-spatial attributes of the base station can be attached with graphics via the base station ID.

During coverage planning, the signal strength distribution data can be regarded as continuous coverage in a certain area, and signal strength distribution data can be accessed by Raster in geographic information, and easily be retrieved and queried, as well as be displayed in the form of image..

During wireless network planning, the planning function is to calculate the coverage of the planning area by using wireless signal communication models. The communication models can be divided into outdoor communication model and the indoor communication model by applicable environment; and empirical model and deterministic model by source of the communication model. Through the analysis of the communication environments, previous researchers and engineers have put forward a number of communication models to predict median field strength of signals accepted. Currently, widely used communication models include Okumura-hata model, COST321 Hata model and general models.

The communication model used by the system is a universal model, whose coefficient is derived by Hata formula. The general model is determined by the following equation:

Path_Loss=K1+K2*log(d)+K3*(Hmeff)+K4*log(Hmeff)+K5*log(Heff)+K6*log(Heff)*log(d)+K7*(Diffraction_Loss)+Clutter_Loss

The meaning of the parameters in the formula is explained as follows:

Path_Loss: communication path loss (dB)
K1: attenuation constant
K2: distance attenuation coefficient
K3, K4: mobile station antenna height correction factor
K5, K6: base station antenna height correction factor
K7: diffraction correction factor
Diffraction_Loss: diffraction loss
Clutter_Loss: surface features attenuation correction value
d: the base station and mobile station distance (km)
Hmeff: mobile station effective antenna height (m)
Heff: base station effective antenna height (m)

For the simulation of phenomena with continuous distribution in a certain space like signal strength, field based view is appropriate. Field model can be expressed as the following mathematical formula:

$$z : s \rightarrow z(s)$$

Where, z is a measurable function, and s represents the position in space. In this system,

Formula: signal strength = sector power + antenna gain - path attenuation, that is, measurable function. The system uses a raster method to calculate the signal strength, in which raster data model is based on a continuous coverage, and it discretizes a continuous space and covers the entire contiguous space with two-dimensional coverage or division.

4 Software Functional Modules

The WCDMA project simulation system manages wireless network construction and simulation teaching through eight main modules. The eight modules are planning principle, planning and design, network construction, RNC board operation, NodeB maintenance, RNC maintenance, signaling tracking, and network optimization.

Network planning module includes the description of planning principles and planning simulation, of which the description of planning principles designs an interactive multimedia teaching framework according to WCDM planning principle. Trainees can independently design the steps of the whole planning principle process based on their understanding and mastery of WCDMA planning principles. The system automatically determines whether every step of the process is correct or not, and guides trainees into the next process based on the results and commence the introduction to the process. All the processes are correct, the system will automatically generate a network planning report. The effect of the description of planning principles is shown in Figure 2.

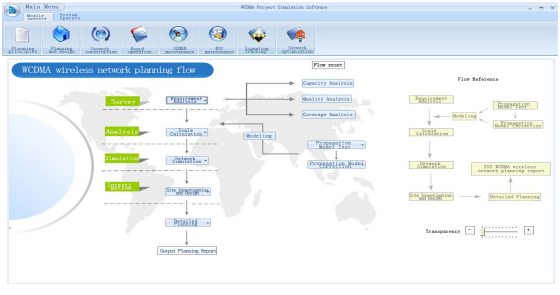


Fig. 2. Operation of Planning Principles module

Planning simulation part provides the virtual geographical environment of the designated areas, and allows trainees to conduct base station site location selection and capacity planning in a virtual geographic environment, set the parameters of the base station sector and simulates coverage planning for the set base station data. The effect of planning simulation part is shown in Figure 3.

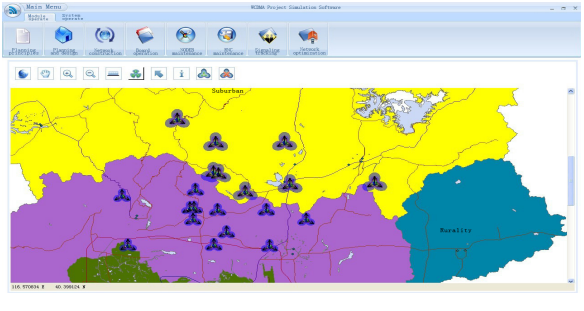


Fig. 3. Coverage planning results

The room construction module is mainly used to simulate the process of defining network topology, designing network elements, configuring links and selecting antenna feeder system. It displays the construction of a virtual room in the form of multimedia animation, and the above-mentioned simulation, trainees are allowed to conduct independent operations, and test whether their independent operations are correct according to the results of the network planning. The effect of the room building part is shown in Figure 4.

Fig. 4. Simulation of room building

5 Summary and Outlook

WCDMA project simulation system, relying on geographic information technology, decomposes engineering construction of the entire wireless network into several relatively independent parts that are correlated and echo each other and form an integral system. The system simulates important links of network optimization, network regulation, and signaling tracking, and can truly reproduce the scenarios of network optimization and network regulation, so that it can both help network optimization engineers work in real practices, and be used in teaching of wireless communication profession in an intuitive way, thus having a high practical value.

The system has been used in the experimental program of the communication vocational education, and achieved good results. How to further strengthen and improve the simulation capabilities of the system to make the simulation more realistic is the direction of future improvement and extension. In addition, considering that the current 3G mobile communications systems also include TD-SCDMA and CDMA2000 in addition to WCDMA, and these two standards are equally important in teaching, we will continue to research and develop system engineering simulation software for the two other standards.

References

[1] Zhang, C.: WCDMA Communication Network Planning and Design. People's Posts & Telecom Press, Beijing (2007)
[2] Pang, W.: On the Planning of WCDMA Wireless Network. Mobile Communications (1), 132–133 (2004)
[3] Zhou, Y., Zhang, J., Tao, J., Zhang, M.: Exploration of WCDMA Network Capacity Optimization Strategy. IT 31(3) (2007)
[4] Bi, M., Wang, S.: How to Balance Coverage and Capacity in WCDMA Network Planning. Design Technology of Posts and Telecommunications (10), 26–30 (2004)

The Application of Computer Technology in the Interior Decoration Project

Qiu Yu[1] and Qiu Jing[2]

[1] School of Arts and Design, Wuhan Institute of Technology, Wuhan, Hubei Province, China
[2] Architecture and Urban Planning School, Huazhong University of Science and Technology, Wuhan, Hubei Province, China

Abstract. With the rapid development of computer technology, nowadays, computer has swept the globe. The Upholstery in numerous metropolises of China has generally realized electronic information technicalization in decoration designing, auxiliary projects, decoration products etc. As computer technology has continuously improved and design software endless develop, these new high-tech assistive technologies make the upholstery more modernized. In the meantime, it also gives interior decoration designer a broader space to thinking and plentiful ways to creation.

Keywords: Computer Technology, Aided Design, Interior Decoration, Application Scope.

Interior decoration aims at providing more comfortable and better indoor living environment. Generally, environment in a broader sense contains natural environment, artificial environment and indoor environment. It is the condition that closely related to the human being and exists in the outside of humanity itself. After human society entered into industry era, and natural environment became more serious, the development of industry brought us science and technology which has promoted social progress. In interior decoration process, we should lift the high-technology contents and take full advantage of advanced technology of the social development to create higher quality and more artistic indoor living environment to meet the rising requirements for indoor environmental quality. The rapid development of modern technology has provided more convenient conditions for people; it makes interior decoration more suitable to human aesthetic and the intention of user. What's more, it also creates healthy atmosphere and more leisure living interest for human beings. (Figure 1; Figure 2)

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 467–473.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Fig. 1. The Hall of YunMengYuFeng Hotel

Fig. 2. The Hall of YunMengYuFeng Hotel

1 Computer Technology Apply to Interior Decoration Designing

With the development of computer technology up to today, its application scope has stretched form initial numerical calculation, word processing to knowledge, information processing and broader areas of development. Picture and graph are more meaningful and intelligible than word and numeric. Even though they are more complex to describe than word and numeric, when compared to traditional manual and mechanical ways, computer is faster and more accurate in graphic processing.

Generally, there are four categories of software which interior design most applied to: modeling software, 2D charting software, rendering software, and post-processing software. The application of computer charting software in interior designing has improved complicated and trivial work of designer hand-painted. Charting software has powerful simulation capability; it could almost describe the true effect close to the real

object. This would be helpful to show the designers' creative capability and be easier to appreciate for people. This skills close to the real could express the expected imagination about interior decoration of designers meticulously and concretely, showing a vivid, exquisite and stereo performance.(Figure 3, Figure 4)

Fig. 3. The Resturant of YunMengYuFeng Hotel

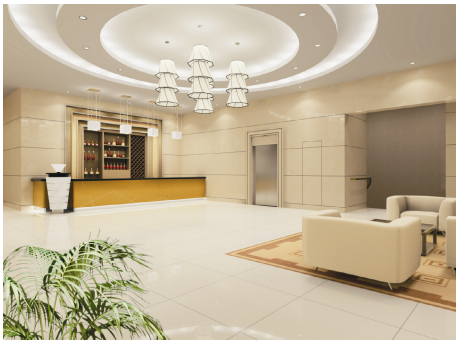


Fig. 4. The Reception Centre of YunMengYuFeng Hotel

In the interior process, renderings, the 3DS, floor plan, the AutoCAD. Besides applying to autodesk directly, we also choose architectural CAD of domestic software developers, includes the PKPM of structure design, the ABD of architectural design, Desai, Yuanfang of the interior design, etc. They are almost the secondary development of synthetic based on AutoCAD, lacking innovative contents and the freedom and the self-design part are quite not as good. Under the traditional CAD-3ds-photoshop mode, the DXF conversion of model design often goes wrong and frequently comes to naught. Renderings are the core of 3D application in the decorative design industry. It ends our

waiting for 3D software based on non-CAD platform until 3DS VIZ turns up. Actually, VIZ is a lite version of MAX. It inherited the merit of MAX, at the same time; it also added what the designers most needed: model library, space capture, sun system, profile production and camera matching. The most of domestic interior designers have completely mastered the skill of applying AutoCAD to draw design drawings and 2D plan diagram. But, the one who could use computer to make conceptual design or draw computer renderings personally among them is few and far between. The main reason is the lack of a set of software easy to operate. At present, most people can only use AtuoCAD and 3DS to render, however, this two application software are independent in many designing details, and cannot be used exclusively for engineering design. Moreover, 3D Studio VIZ's appear which developed in accordance with the need of construction industry completely solved the problem above; designers could use it to design professional 3D modeling, light effect, decoration color, etc. in the same 3D environment. Due to the high visualization level, designers could present architecture plan in their head directly via computer. (Figure 5, Figure 6)

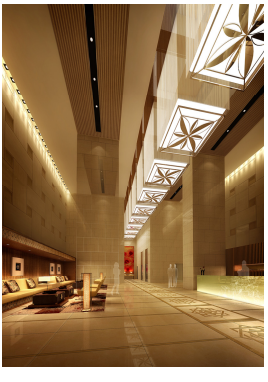


Fig. 5. The Hall of XuHui square in Jiaxing

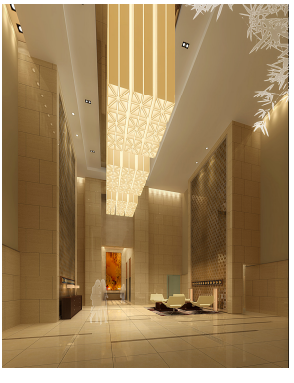


Fig. 6. The Hall of XuHui square in Jiaxing

Because of the commonly used lamplight and model provided by system, VID98, Chinese version design software by Daheng ACE and Kinetix, compatible with tradioray make it become the most efficient and simplest 3D software. When the past software in use, interior designers often cannot be effective to settle a series problems of material quality, lamplight designing, due to the limitations of the software itself. In contrast, 3DS VIZ matching up to VID98 allows designing model both in decoration materials and lighting effects to be real, natural, seamless. Moreover, the excellent operating platform makes VID98's function be unprecedented. Doors, windows, walls, floors and other interior designing are simple to grasp easily, even not need training.

2 Computer Technology Apply to Interior Decoration Construction

（1）Lamp lighting

Computer technology has entered into designing of lighting lamps, and is being integrated into modernized decorative lamps technology, becoming a decorated consumption hotspot. Computer type lamps could regulate light source automatically according to the need, it also could regulate automatically via designing by computer program. Computer color control only need you to touch the lamp shell, you may control the switch and change the intensity. For example, when people go though the dark corridor, staircase, electronic sensors will send a feedback signal to turn on the power automatically as long as human body enters into detection range. That both can improve the lighting quality and prevent thieves to steal effectively. This new infrared sensor lighting without contact could save energy and increase life span significantly.(Figure 7, Figure 8)

Fig. 7. The Elevator hall of XuHui square in Jiaxing

Fig. 8. The Elevator hall of XuHui square in Jiaxing

(2) Shade curtain

Shade curtain sets light operated, remote control in one. According to infrared remote control, it may teleoperate the micro-motor of curtains so as to drive the rack gear complete the order transferring. Press the button gently to turn the remote control on and off to make curtain up and down, keeping the interior light in a light and dark horizontal plane. If matched with original elegant curtain cloth, it may be better to harmonize interior decoration style and makes the room more stylish and elegant.

(3) Sanitary ware

Apply computers, infrared and other advanced technology to sanitary ware, such as automatic flushing toilet has cleaning, drying, heating, function. Then there is a computer steam shower on the basis of sauna. It is also controlled by computer and human body is alternately via hot and cold water cycle to promote blood circulation and relax spirit.

(4) Unit gate control intellectualization

Nowadays, the public safety prevention equipments in the apartments, villas and office premises are continuously improved. The equipment with a multi-functional, intelligent, fire-protection, prevention of burglary, intercom and video phone is popular. This advanced equipment has realized the idea that pressing the button to open the door and keeping people indoors. It could prevent miscellaneous people enter into buildings and be effective to maintain public security and protect residents' property. Building gate intelligent systems is controlled by infrared sensors, when it has received body sensor signals, the electronic chip would control component to drive micro motor immediately, and the mechanical accessory is driven to rotate to make the mobile door makes of stainless steel aluminum alloy back and forth to move. (Figure 9, Figure 10)

Fig. 9. The Room of XuHui square in Jiaxing

Fig. 10. The Room of XuHui square in Jiaxing

(5) Electronic construction detection

Use semiconductor laser launching tube to make the integrated circuits, it would form 2mm vertical line under 500-Lux to replace the traditional measuring tools, line hammer and straightedge. The laser instrument visual inspect objects via vertical method could locate suspended ceiling, partition wall lamps and lanterns, windows, etc quickly. In the process of locating large decorative material cutting line, using the red line in the laser instrument to detect the space vertical region in the construction precisely to achieve the quality construction detecting standard, to some extent, may play a supervisory role in improving quality of interior decoration construction.

3 Conclusion

The advanced technology skill provided by computer technology release the people working on architecture designing from heavy and complex works. No matter in the designing or construction processing, all reflect the superiority of the electronic information age, significantly improved the quality of life and work. It opened up a new era both in interior designing and construction supporting, became the new fashion of interior designing in 21 century.

References

[1] Wang, Y.: The application of First Exploration Psychology in Space designing. Journal of Hubei University of Economics (Humanities and social sciences) (02) (2009)
[2] Wang, Z., Chang, F., Li, W.: The Works of Art and Design Apartment of Wuhan Textile University. Journal of Wuhan Textile University (08) (2008)
[3] Li, W.: Art Skill Technology——Western art historic and related issues. Journal of Wuhan Textile University (01) (2008)
[4] Huang, K., Lan, F.: Analysis of interior lighting design. Journal of Guangxi University (philosophy and social sciences) (S1) (2007)
[5] Li, H., Chen, M.: Analysis of interior space “soft environment”. Journal of Hetao University (natural sciences) (04) (2009)

The Application of Image Recognition Technique in Candidate's Identity Authentication System

Gu Hongjiu

ChongQing Industry Polytechnic College, 401120, China

Abstract. Along with the development of digital image processing technology and computer vision processing technology, using the image recognition technology to replace the human being to carry out artificial intelligent control and information collection provides an necessary technical support for the realization of complete automation technology. To begin with, this paper introduces the current situation of the development of the image recognition technology, then specially describes digital image recognition technique and fingerprint recognition technique, and at the last part, it analyses the application of fingerprint recognition technique, a branch of image recognition technique, in candidate's identity authentication system.

Keywords: Image Recognition Technique, Fingerprint Distinguishing Technology, Identification Authentication, Application.

1 Introduction

At the end of the 1950s, a image recognition technology shaped along with the constant improvement of computer technology and the microelectronics technology. Since the1960's, image recognition technique was gradually accepted by the scientific community and was applied into many fields. The main application object of image recognition technology is "picture", which includes not only a picture, but the ability of identifing and influencing clips. In addition, with the continuous development of science and technology, the researchers put the images and sounds into identification contents of the image recognition technology. Image recognition technology mainly relies on fast data handling capacity of modern computer, that is, making use of the computers' gigantic data processing abilities to replace people to handle the enormous complex physical messages, thus some concrete work can be realized, those of which human was unable to achieve. For instance: in modern factory production workshop, a large number of parts need to be checked every day, which results in the scant accuracy and the importance of image processing technology is highlighted. What's more, Image processing technology is applied to such areas as label identification, two-dimensional scanning, overspeed photographing, seal, fingerprint identification and satellite diagram.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 475–480.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Image Recognition Technique

In present various means of identification technology, comparing with the audition recognition and touch recognition, the visual identification owns more broad application space. From the perspective of informationism, the image recognition technology also has the largest amount of information identifications, among which there are four kinds of "pictures" can be identified, they are Gray scale images, color images, plane pictures and stereo images. Alike the basic principles of human recognition images, the computer image recognition technology also needs to learn the demanded identification images at first, then picks up characteristics owned by these pictures(such as: shape, color, texture, etc), and finally you can conduct the actual operation tasks according to instructions.

One of the key technologies of the image recognition technology is image collection technology. At first, the image collection technology demands using optical lens to collect pictorial information, then changes these information into storaged video data after being managed by the image detector. By the PCI interface port, those video data first deliver the images into the computer in frame and then dispose those collected images with the help of data handles function of computers or directly present on the screen of the Information recognized by image recognition technique can be divided into threes parts according to dimension:

(1) one-dimensional image data: It deals with signal and recognize speech (also named waveform recognitio). This kind of image data mainly solves practical problems of language, radar, electrocardiogram, earthquake waves and so on.

(2) two-dimensional image data: Its tasks contain word processing, images processing, pictures processing, word identification, pictures, images, etc. This kind of image information is mainly used to handle texts, images, remote sensing, X-ray pictures and other actual problems.

(3) three-dimensional image data: Scene analysis and holographic recognition are its missions. This kind of image information material is mainly used to deal with target robot, target recognition, robot vision, and other aspects of the tasks.

3 Digital Imaging Recognition Technology

Digital image recognition technology is composed of digital video surveillance system and image recognition system. First of all, the computer stores the real images intercepted by the digital video surveillance system; then the stored image data will be treated with pretreatment, image characteristics extraction, events judgment and other image recognition processes.

In the process of image information acquisition, "image" information is intercepted by each frame of the streaming video. Those grasp "images" present in the form of frame composed of three primary colors. The three primary color is represented by the 8-brightness value(0～255) of the fundamental color of the red, yellow and blue, which in the programming language is represented by English letters such as R, G and B.

In the process of image and processing, the main functions of the system are removing interference, noise and differences, and turning the retrieved images into the specific formats recognized by the computer, which mainly includes images' transformation, reinforcement, filtration and other steps. Generally speaking, image transformation is achieved by the mathematics function generation(such as Fourier transform, cosine/sine transform, Walsh transformation and Hadamard transform). In virtue of those transformations, the "image" can be converted from time-domain to frequency-domain, which results in the perfection of the picture quality and the speed-raising of the picture processing. Image enhancement is also utilizing mathematical principles and methods to enhance picture sharpness and to filter such interference factors as image grey correction, noise removal, the image smooth, corrosion, sharpening, image edge enhancement, etc. In the process of image characteristics extraction, the chief job is to tidy, to analyse and sum up the "image" data, to extract characteristics can reflect the nature of the image and to get needed parameter in the following processing.

In sentencing phase, that is, a stage that classifies the above received "image" information purposely. During this process, characteristics/parameters and threshold values of each "image" information are being compared and then being categorized to provide information parameter support for the computer operation processing.

4 Fingerprint Distinguishing Technology

In the world, there are no complete same fingerprints, not even the identical twins. Everyone is different from each other in the design, breakpoints and intersection, no exception to the skin textures of fingerprints, which appears the uniqueness and without change forever. Accordingly, we can combine one with his or her fingerprints correspondingly, by comparing his fingerprints with the preserved fingerprint data so that his or her virtual identity can be checked, which is the working principle of the fingerprint distinguishing technology. Fingerprint distinguishing proves operators' identifications mainly based on the human's fingerprint lines and detailed features. Thanks to modern electronic-integrated manufacturing techniques and fast reliable algorithm research, this technology has blended into our daily life, and has become the most mature technology with the deepest research and the most widely use in biological detection.

4.1 The General Characteristics of Fingerprint Identification

General characteristics refer to those features observed by the eyes directly, including grain shape, mode area, core, triangle point and grain number, etc.

On the basis of pundits' long-term practice, fingerprint classification can be divided into three parts in response to the trend of ridge and its distribution situation, which are loop(also named bucket shape), arch and whorl. Mode area refers to an area including the overall features of the fingerprint, so that we can discern which type you belong to according to the region. Some fingerprint identification algorithms only use the data of

mode area while others use the complete fingerprint. Core is the gradual center, acting as a reference point when you reading fingerprints and comparing with them. Many of the algorithm are on the basis of the core, which means that only those own a core can be processed and identified. Triangle point is the first bifurcation point or breakpoint starting from the core, or the convergence of two lines, isolated point and turning point, and the place where point to the singularities. The beginning place of the the count tracking of the fingerprints is offered by the triangulation point and the grain number refers to the number of the fingerprint grains inside the mode area. In general, when you are counting the fingerprint lines, the core and the triangle point will be connected firstly, and the number of the connecting line of which interacting with the the grain fingerprints can be regarded as the fingerprint quantities.

4.2 Fingerprint Collection Technology

How to get a good fingerprint "image" is a very complex problem. The reason why is that the finger mark used for measuring is a fairly small piece of skin, which demands the fingerprint acquisition device be good enough resolving capability to ensure the acquisition of the details of the fingerprint. Currently, the used acquisition equipments are basically in the formation of the following technical foundations: optical technology, semiconductor silicium techniques and ultrasonic technology.

Optical technology: With optical fingerprint technology to collect finger marks is the most remotest historic and the most widespread use technique. Putting your fingers under optical lenses ensuring fingers within the area of built-in lamps illumination, and using the prism to project the above to the charge-coupled device (CCD), thus the black ridge line is shaped and the white valley line is formed. The ridge lines are those of certain width and trends while the valley lines refer to those digital ones and different gray fingerprint "images", also equals to the the lines between depression, which can be managed by the fingerprint instrument algorithm.

Complementary Metal-Oxide Semiconductor (CMOS technology): It developed on the foundation of semi-conductor silicon-capacitance. Silicon sensor becomes one plate of capacitance and the finger is the other one. Using hand fingerprint ridges and valleys, relatively to the capacitance of smooth silicon sensor, 8 bit form of gray "image" will be shaped.

Ultrasonic Technology: It makes use of the the ability of ultrasonic such as penetrating materials and the production of different sizes of the echo as a consequence of different materials. The different echo is the result of different materials, the differences of which result in the various absorption, transmission and reflection. Therefore, with the help of different acoustic wave impedances between the skin and the air, the location of the fingerprint ridges and valleys can be distinguished.

4.3 The Acquisition, Verification and Identification of Characteristic

After being intercepted, a high quality image still needs a series of procedures to transfere his features into a composite template, the process of which is called process characteristics, the core technology of finger scan. After being intercepted, a high

quality image must be converted into a useful format. If the image is gray scale one, the relative shallow part will be deleted and the relative deep part will be turned into black. About 5 to 8 pixels of the ridge will be compressed to one so that the ridge breakpoint and the divergence could be located precisely. Tiny detailed image comes from this processed image. In this process, even the very fine image also has the inevitable deformation details and error details, which must be filtered.

Apart from the application of detail's location and angle methods, the styles and qualities of the details can also be used to differentiate the details. The advantage of this method is the accelerated speed of the retrieval, because the uniqueness of a notable and specific detail make the match successful. In addition, some manufacturers are adopting pattern matching method, a way of deducing data of a specific group to process the fingerprint images.

5 The Application of Image Recognition Technique in Candidate's Identity Authentication System

In the traditional examinations, photos and some effective certificates are the mainly ways of identitifying authentication. With the development of science and technology, the candidates identification precision can not be ensured by the photos and some effective certificates, for many candidates are forging with the help of computer technologies. Therefore, the sponsor have to utilize new identity authentication system, of which the fingerprint identification system is researched based on accurate personal identity.

In large-scale test, it is difficult to use fingerprint identification technology to recognize authentication, because fingerprint identification system is a high-tech equipment, a large amount of fund will be needed if it is installed in all the examination centre. However, in some small-scale test, this can be effectively implemented. What's more, there is another benefit of this system, that is, convenience. The finger mark will not change with the pass of time, and it is easier to collect comparing other ones, and at the same time, a certain scale of sample library of fingerprint identification has been built.

There are two main types of fingerprint identification technology adopted for identity authentication. The one is embedded system, and the other is desktop application system. Embedded system is a relative complete and independent system and is composed of a complete set of system, without the mutual coordination of the outside instruments. It first scans and stores up examinee's fingerprints, turns those data into a certain type recognized by the computer, summarizes the image characteristics and processes digitally, combines the candidate's information with himself or herself and then put those information into the IC card and handle them in a encryption way. When taking the examination, the candidates' identities will be effectivly recognized by the comparison between the carried IC card and his or her fingerprints. By comparison, the desktop application system has a much more flexible system structure, which can connect several external devices and share the fingerprint "image" information, especially of the application of those large-scale database.

6 Conclusion

As a new product of the development of science and technology in new times, image recognition technology will be applied to every aspects. At the same time, with more and more respects have been paid to the exam, more and more large-scale test will take more scientific personal identity authentication technology to ensure the integrity of the exam. Therefore, the image recognition technology will be applied to social life with the continuous development of the computer technology and microelectronics technology.

References

[1] Wang, Y., Li, S., Mao, J.: Computer's Image Processing and Identification Technology. China Higher Education Press, Beijing (2001)
[2] Yin, Y., Ning, X., Zhang, X.: The Development and Application of Automatic Fingerprint Indentification Technology. Journal of Nanjing University 38(1)
[3] Li, S., Yin, C., Shang, Q., et al.: Judging Method for Chaotic Characteristics Based on Image Recognition Technique. In: Proceedings of Chinese Society of Electrical Engineering (2003)
[4] Wu, C.: Design and Implementation of Campus Security System Based on Image Recognition. Digital Technique and Application 37(2), 118–120 (2005)
[5] Yang, X., Ning, X.: Image Segmentation of Automatic Image Recognition System. Journal of Nanjing University 40(4), 424–431 (2004)
[6] Tang, L., Xie, X.: Image Segmentation Methods Based on D-S Evidence Theory. Chinese Journal of Computers 26(7), 887–892 (2003)
[7] Li, J., Ma, X.: Binary Image Method of Fast Dynamic Threshold Based on Digraph. Dalian University of Technology 42(5), 626–628 (2002)
[8] Rao, Z.: Fingerprint Distinguishing Technology and Its Application of in School Exam Management. Coastal Enterprises and Science & Technology (6) (2001)

Research on the Security Risk Assessment and the Solutions of Cloud Computing

Wen Ding

Hunan Mechanical & Electrical Polytechnic,
410151 Changsha, China

Abstract. As a new technology in the field of information technology, cloud computing has brought many benefits as well as big security risk to us. This passage gives a simple introduction to cloud computing. Then the big security risk at present is summarized and the corresponding measures are given.

Keywords: cloud computing, information safety, risk assessment, solutions.

1 Introduction

Cloud computing is the revolutionary technology product in the Information technology age, which can realize fully sharing of network and computing resources. One Founder of IBM Thomas. J. Watson once said:" The global only need 5 computers." And the appearance of the cloud computing make his predict ture and provide the technical support for it. When cloud computing bring a lot of convenience to people, at the same time safety risk accident happened. According To Statistics, Amazon cloud computing platform broke out many safety risk accidents during 2007-2008; The Azure cloud computing platform of Microsoft meet the accident of completely broken in 2009, which lead to the lost of information materials of most users. With more and more accidents of safety risk, cloud computing is experiencing a huge blow right now. "Putting the resource and data on the internet, safety is the problem." This has become the main psychology when the users want to use cloud computing. This article introduces the main facing safety risk cloud of computing platform, and puts forward some concrete measures according to the potential safety hazard.

2 Technical Survey of the Cloud Computing

2.1 Basic Concept of the Cloud Computing

The cloud computing is a kind of internet technology which share the computing resources and deal with relevant data and task using the Internet. The execution of data

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 481–486.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

storage and computation is not in local computer or server equipment but in the Internet in the cloud computing. The basic application mode of the cloud computing is like Figure 1.

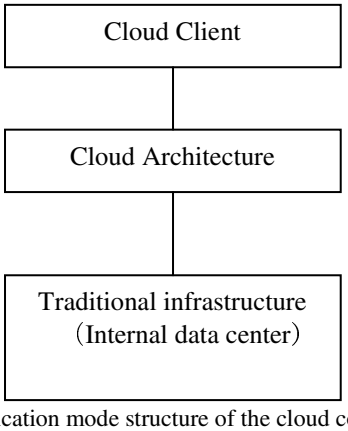


Fig. 1. Application mode structure of the cloud computing

The formation of the cloud computing can be treated as integration development of grid computing, parallel and distributed computing. However the cloud computing is different from any kind of that computing. Taking grid computing for example: grid computing stress to solve data calculation and storage resources distribution, but the cloud computing pay more attention to solve data calculation, storage and sharing of resources. Users don't need provide the resources of computing and storage as using the grid computing, and also don't need to buy developing applications.

2.2 The System Architecture of the Cloud Computing

The system architecture of the cloud computing is divide into: Basic management layer, Application interface layer, access layer(Figure 2).

Configuration as shown, basic management layer is mainly used to process the data sharing in Three layers of system structure of the cloud computing. Application interface layer is mainly used to solve how to provide remote users service; Access layer is mainly used to solve some practical applications by using the cloud computing.

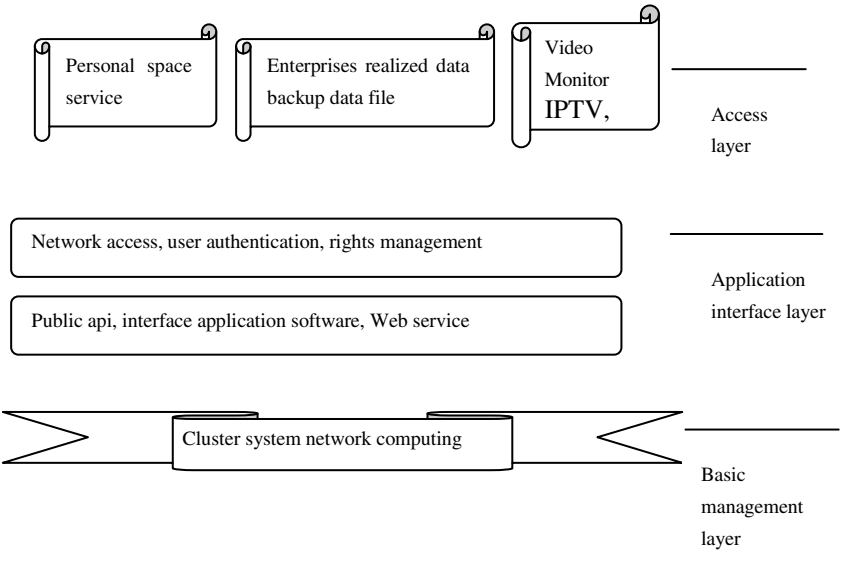


Fig. 2. The system architecture of the cloud computing

2.3 The Main Forms of the Cloud Computing

The main forms of the cloud computing are: software maintenance (SaaS), Platform as a service (PaaS), internal cloud web service, management service provider and utility computing.

Software maintenance is a software layout model, which is design for Function of Network delivery and is convenient for the user to custody, deployment and access through the Internet. The typical representative product of this kind is Google Docs.

Platform as a service is another products form just like software maintenance who regards development environment as a kind of service. The typical representative product of this kind is the platform of Face book and Force.com of Sales force.

Internal cloud web service is a kind of form which develops new applications through API provided by Cloud computing services. This kind of service can don't use all the features in the software development and application. The typical representative product of this kind is the Google Maps & Google Earth.

Management service provider is a traditional form of the cloud computing. It is a kind of service that managed service provides to IT users but not the end users. For example, virus scanning services of email and application of monitoring service. The typical representative product of this kind is Secure Works, Managed Security Services by IBM and Verizon.

Utility computing let the users use data storage and virtual computer provided by the service providers according to their own need. The typical representative product of this kind is S3 and EC2 of Amazon.

3 The Key Factors Leading to Security Risk of the Cloud Computing

3.1 The Disappearance of the Traditional Security Boundary

The safeguarding is controlled by border security and access control in the traditional net safeguarding. However, the traditional border security is disappearing as high degree of integration of computing resources and data storage in the cloud computing. As a result, many network security problems arise.

3.2 The Safety Problems of Virtualization Service

"All problems can be solved by adding a layer of mapping in field of computer science research " According to this perspective, the majority problems can be solved by virtualization in the computer system. Virtualization is the core technology in the cloud computing, which bases on the height integration of storage resources and server resources to develop and use this technology. Virtualization scalability provides the ability of cloud service level by expanding infrastructure and software platform to its max. However, How to solve the problems of the cloud computing center basic network structure, data storage and application services, virtualization delivery put forward higher request for safety design and deployment under this technology environment.

3.3 Safety Problem after the Concentration of the Data

Data storage and computation are concentrated executed on virtual platform in the cloud computing, thus cause both aspects of the security threats. : ① Because of using the expansion of the scale of the server in data storage and computation, basic network architecture strict for unity which led to the higher demand of their own safety protection.② Because of any user operation of cloud computing system is related to each other, how to prevent threat caused by more users to make sure the identity authentication management control of the safety of the task execution is becoming the potential safety problems affecting factors in the cloud computing.

3.4 Stability and Reliability Problems

Firstly, the customers' different needs require to be solved directly through the cloud computing system, which will influence the stability and reliability of the cloud computing services. Secondly, because the customer's operation is a variety of high concentration tasks, the cloud computing system faces attack could be far higher than the traditional network platform. So how to make sure the stability
and reliability problems of cloud computing is the problem that should be paid attention to.

4 The Solutions of the Cloud Computing Security Risk

Facing the safety risk brought by its own technology of cloud computing,
how to solve the science and prevent the security hidden danger effectively has become one of the primary content of cloud computing technology research. According to cloud computing technology development present situation in domestic, the following measures are put forward for the current domestic cloud computing platform.

(1) Perfect laws and regulations of cloud computing technology

To prevent and solve the hidden danger of cloud computing security, first is to start from the law level. Through the clear cloud computing operators and the safety of user's duties, the cloud of security threat is to be limited.

(2) Strongly advocated the technical innovation in science and technology of
the cloud computing safety protection

Cloud computing is a new information technique, many key problems are still waiting to be solved. The development of more security protection core technology is main means to protect the enterprise itself and the national information security under the future cloud computing environment. Only mastering more and more core technology of computing clouds, constantly breakthroughs can be made of cloud computing hardware middleware software and network facilities.

(3) Unified planning In the cloud computing platform to ensure the ability of the later safety

At the beginning of the IP agreement to start work, No scientific unified planning and this lead to a large number of patch is needed later. Learning from the lessons of the Internet IP protocol, it is suggested that the safety and security control technical standards by test and verify the cloud computing system in the large-scale development in domestic. For example, it can be realized through authentication encryption isolation network control.

(4) Actively Participates in the international cloud computing's discussion and exchange and the cloud computing related standards

Nowadays, Cloud computing is a new stage without unified standards and protocols. A lot of standardization organizations promote the formulation of relevant standards and norms actively. The relevant technology of the domestic research institutes and universities should take the chance and take an active part in making related iso standards and specifications. We need to be positive both in some of the core technology and API, control interface/agreement aspacts.

(5) To strengthen higher education, provide more and more excellent talents for cloud computing technology in the future

The higher education in our country has made great progress now, but there is great shortage in some of the core technology areas. Our country should make great efforts to strengthen the quality of higher education and perfect the current higher mechanism of talents training. In order to promote Chinese colleges' and universities' research in area of cloud computing technology, cloud computing and safety system authentication evaluation should be established. At last, more talents are trained for cloud computing technology.

5 Summarize

As new information technology products, cloud computing cause a lot of the attention in many field and it will be a common technology in the future. Although cloud computing facing all kinds of safety risk now, as the sound and perfect of the relevant laws and regulations of the cloud computing, the security risk of the cloud computing will reduce to the minimum in the future. The development of cloud computing in China, we should start from the technical standard and key technology of the field. It is not until we control more and more proprietary intellectual property rights and standard that we can ensure the information technology security of enterprises and the whole state.

References and Notes

[1] Varia, J.: Cloud architectures -Amazon Web Services(EB/OL) (March 01, 2009), http://acmbangalore.org/events/monthly-talk/may-2008-cloud-architectures-amazon-web-services.html
[2] Chang, F., Dean, J., Ghemawats, et al.: A distributed storage system for Structed data. ACM Transactions on Computer System 26(2), 1–26 (2008)
[3] Guo, L., Zhang, N., Shang, J.: Cloud Computing Environment Security Framework. Netinfo Security (7) (2009)
[4] Hu, W., Zhong, W.: Cloud Computing on Network Security Problems. China Science and Technology Information (23) (2008)
[5] Gongdeizhong: Cloud Computing Safety Risk Assessment Model Analysis. Hubei Police Academy Journal 123(6), 85–86 (2011)
[6] Xie, S., Feng, Y.: According to the cloud computing and information security. Journal of Beijing Electronic Science and Technology Institute (4) (2008)
[7] Cai, Y.: Based on cloud computing information system safety risk assessment model. China Management Informationization (12) (2010)

The Study of Security Management Platform Technology and Methods Based on Cloud Computing

Zhang Jie

Abstract. Cloud computing is currently the increase, the use and delivery models in Internet-based services which usually include supplying the dynamic, easily-extending and virtual resources by the Internet. Cloud is a metaphor for network. This paper analyzes the working principle and technical features of cloud computing and studies the security management platform technology and methods as well as the safe application performance.

Keywords: cloud computing, security management platform, method.

1 Introduction

Cloud computing is another technological innovation following the revolution of the mainframe computers, the clients and the servers of 1980s. Cloud computing with extremely powerful processing power is integration of grid computing, distributed computing, parallel computing, utility computing, network storage, virtualization and load balancing etc technology. The basic working principle of cloud computing is distributing the computation to a large numbers of distributed computers rather than the local computer or remote server which makes the operation of the enterprise data center more similar to the Internet. Cloud computing helps the small and medium enterprises to utilize the limited resources effectively and access the computers and memory system according to the requirement. We can regard the tendency of the development of traditional computers to cloud computing as the trend of transmission and distribution structure of traditional independent generators towards the transmission and distribution of today's power network. The emergence of cloud computing impacts on the current concept of calculation and data processing in one independent computer, it means that computation and processing ability can be regarded as a kind of power which is allowed to be divided. Cloud computing reduces the unnecessary waste of software and hardware, supplying much broader space for technology, educational research and technical application in the future.

2 Security Management Platform and Its Features Based on Cloud Computing

In 2009, IBM and Sinochem Group jointly held a press conference of the enterprise cloud computing platform. As the first enterprise cloud computing project in the world,

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 487–491.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

China Sinochem Group applied the solutions from IBM cloud computing center to deploy the ERP system in the cloud across two data centers by the opportunity of upgrading the ERP system comprehensively. It not only comes true to upgrade the ERP system smoothly, but also makes the internal IT infrastructure and a variety of software applications to be able to run more flexible.

Currently, the majority of our cloud computing technology is only used for computer anti-virus, security, etc. The core technology of the security management platform based on cloud computing is "cloud security", it's developed on the basis of cloud computing and it monitors the abnormal aggressive behavior on the Internet by using numerous computers or other devices in order to gain the real-time information of Trojan Horse and other malicious programs pushed to the Server for automatic analysis and processing and then distributes the viruses and Trojan solution to each client. The vision of the basic principles of cloud security is that: more user client, more secure for each client. Because of such a large number of clients, we can cover the entire computer Internet, as long as a website has been linked to Trojan Horses or a new virus appears, we can intercept them immediately.

Cloud security management platform requires the following three basic conditions: ① the IT infrastructure of the entire user client cluster should be with centralized pattern; ② it requires cloud security information management platform operation and management unit to be with skilled operational capacity; ③ the entire security management platform staff must understand cloud security and cloud computing with service-oriented architecture (SOA) concept.

Cloud security is the integration of new technologies and concepts parallel processing, grid computing and unknown virus judgment etc. It detects the abnormal behavior of Internet software by a large number of mesh clients. The basic characteristics of the cloud security management platform are:

(1) It has the characteristics of the full range of security services. It is very convenient to use the security services provided by the security management platform based on cloud computing as long as uses only get a PC and network connection. Only from this sense, any computer connected to the Internet has the potential to be cloud computing platform while enjoying the ability to use the cloud security management platform anytime and anywhere.

(2) Cloud security management platform needs to have a certain technical strength. The users has the ability to use the cloud security services, but don't have the technical conditions to be an independent provider of cloud security services. For example: companies need a certain technical processing capabilities, such as 360 security guards, QQ anti-virus software, Rising Kaka 6.0 and other software, not any business or individual can easily build cloud security management platform.

(3) Whether the cloud security users get access to this platform. Cloud security services provided by the cloud security management platform don't care about the personal information in the computer of Internet users. Users can install the cloud security management platform plug-in according to their need without the limitation of Internet and other servers.

With the gradual development of Internet technology, anti-virus software won't deal effectively with the growing number of malicious programs. The main threat from the Internet is by computer viruses to malicious programs and Trojan, in this case, the use of the feature library discrimination law is clearly outdated. After applying the cloud security technology, the recognizing and killing the virus is no longer only relying on the local virus database in the hard disk, but on a vast network of service real-time collection, analysis and processing. The entire Internet is a great anti-virus software and the more participants there are, the more secure each participant is. The entire Internet will be more secure.

3 Framework and Solution Based on Cloud Computing Security Management Platform

The framework and solution of security management platform based on cloud computing is to build a management center based on cloud computing, and the center consists of user authentication system, data integrity verification system, key management system, encryption/decryption system and log management system.

The basic solutions for the security management platform system based on cloud computing include the following seven aspects.

(1) The framework of security management platform based on cloud computing is mainly to build a cloud security management center based on cloud computing and the center here must connect to the Web server, the front desk terminals, backend service terminals on the cloud computing platform. This center also shoulder the duties of checking the users' login authentication or identity verification and ensuring the integrity of data of front servers and backend servers in cloud security management center.

(2) Cloud security management platform consists of front servers and backend servers of cloud computing center. Among them, the front ones connect to the Web server and connect to the backend ones via the gateway device. Once the users of the cloud security management platform run the cloud security software on the backend servers, the cloud security management platform system will send the cryptographic files to the backend servers via the gateway device and after processing the file, this platform sends them to the front servers in the cloud computing center through the Internet Web server.

(3) Each client on the cloud security management platform can apply the smart chip devices to create the sub-encryption system of cloud security management platform client. On the smart chip, we should build the personal identity authentication protocols, digital signature protocols, signature verification protocols and data encryption/decryption protocols of the cloud security management platform client users.

(4) On the smart encryption chip based on the cloud computing security management, we need to set relatively known cryptographic algorithms, key generation algorithms and a fixed symmetric storage keys. In the information database of security management center based on cloud computing, it stores the identity information of

cloud security management platform client users and cipher text corresponding to the key seed data.

(5) Security management center based on the cloud computing security management platform consists of IPC and the smart encryption chip hardware. Here we mainly apply the hardware as the identity information management terminals of cloud security management platform users.

(6) In the cloud security management platform user terminals, there are smart chip devices which have a relatively complete set of encryption system. It can encrypt the file the cloud security management center transmitting to the front servers through the Internet Web server and after transforming the cipher text the smart chip can recognize, the system upload it to the cloud security management platform.

(7) Create personal log libraries of client users in cloud security management center, record the real-time login, storage, anti-virus etc operation of users and cloud security management platform workers in order to get the first-hand information for safe handling off when other client users experience the same situation.

4 Conclusions

Security management technology based on cloud computing is a new security management tool using the power of the entire Internet computers to detect the computer viruses, unlike the traditional firewall and anti-virus software. It adopts a new security detection, anti-virus, protection mode both reducing the memory usage of anti-virus and enhancing the individual and enterprise computers' real-time protecting capabilities. It is a very useful security management platform. It needs the practical application results and feedback of the computer security corporations and users to promote and apply the cloud security management platform for a large area. It requires the collaboration of the security management platform and the clients to make the Internet into a huge computer virus defense system to better protect the security of computers and the Internet.

References

[1] Seal, R.A.: The information commons: new pathways to digital resources and knowledge management. Preprint for the 3rd China/U.S. Conference on Libraries, Shanghai (March 2005)

[2] Gong, C., Liu, J., Zhang, Q., et al.: The characteristics of cloud computing. In: 39th International Conference on Parallel Processing Workshops, pp. 275–279 (2010)

[3] Xue, J., Zhang, J.-J.: A brief survey on the security model of cloud computing. In: Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science, pp. 475–478. IEEE Computer Society (2009)

[4] Hu, X., Zhao, G.: A CSK-based Solution for Person Authentication. In: The Seventh Wuhan International Conference on E-Business: Unlocking the Full Potential of Global Technology (2008)

[5] Zhao, G., Hu, X., Li, Y., Du, L.: Implementation and Testing of an Identity-based Authentication System. In: ISECS International Colloquium on Computing, Communication, Control and Management (2009)
[6] Du, L., Hu, X., Li, Y., Zhao, G.: A CSK Based SSL Handshake Protocol. In: The IEEE International Conference on Network Infrastructure and Digital Content (2009)
[7] Yang, X.: Analysis of cloud security. Information Security and Communication Confidentiality 9 (2009)

Data Mining in the Application of E-Commerce Website

Gu Hongjiu

ChongQing Industry Polytechnic College, 401120, China

Abstract. With the development of computer technology and Internet technology, the Internet has affect all aspects of people's lives greatly, One of the issues in the applications of Internet technology is how to pick up useful information that is benefit for us from these complex data sets. This article describes the process, methods, and specific applications of data mining in e-commerce site, and it will provide services that grasping dynamic changes, tracking changes and making the development strategy of the future for enterprises.

Keywords: E-commerce sites, Data mining, Web, Application.

1 Introduction

With the development of economic globalization and trade liberalization, the emerging technology of computer networks has gradually penetrated into various aspects of everyone's life , and the e-commerce industry is generated as a platform. As a kind of new business model, e-commerce has changed people's opinion of the traditional commerce and trade, the business philosophy and the method of payments, and it has injected fresh blood into today's business community and brought a revolutionary technical impact to the traditional business model. On one hand, the computer network technology bring detailed business information to people, but on the other hand, it also brings some problems to e-commerce at the same time. One of the practical problems about e-commerce technology is how to collect useful information on their own when we are facing so much business information. Therefore, data mining as a kind of technology about data analyzing and finishing is in urgent need in e-commerce. It will processes and analysis a large amounts of information on the Internet for enterprises in the normal e-commerce trade more effectively, and it will provide business model , marketing strategy and decision-making enterprises in the future as a more accurate technical and information support.

2 Data Mining Techniques in E-Commerce

Data mining is a process that extracts implicit in which people do not know in advance, but is potentially useful information and knowledge from a large number of incompletely, noisy, fuzzy, random data. Data mining is a kind of new business

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 493–497.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

information processing technology, its main feature is that extract the key data which is the auxiliary business decisions through extracting, conversing, analyzing, modeling and processing a large number of business data in commercial databases. Data mining is a kind of deep-level approach to data analysis. The data analysis itself has had many years of history, but in the past the purpose of data collection and analysis is used for scientific research. Due to the limitations of computing power, the large amount of data analysis of complex data analysis method is very limited. Nowdays, thanks to the automation of various business sectors, the commercial sector to produce a lot of business data, these data are no longer collected for the purpose of analysis, but because of pure opportunity (Opportunistic) commercial operations. Analysis of these data is no longer only for the needs of research and mainly for business decision-making that needs truly valuable information, and it will profit us. All enterprises are facing a common problem: the data of enterprise is largely, but valuable information is rarely.From large amounts of data through in-depth analysis, you will find information that is conducive to business operations and competitiveness, and it is like alchemy from the ore. Data mining hence obtaining its name. Therefore, data mining can be described as: enterprises established business objectives to explore and analyze a large number of enterprise data and to reveal hidden, unknown or to verify the known regularity and further modeling advanced and effective method technology used a lot in data mining, which mainly has four kind of technologies: statistics, machine learning method, the neural network method and the database .Statistics can be divided into: regression analysis, discriminant analysis, cluster analysis, exploratory analysis method; Machine learning method also can be divided into: inductive learning method (decision trees, rule induction), learning, genetic algorithm based on the example methods; Neural network to also can be divided into: cashbox neural network (BP algorithm), self-organizing neural network method; Also of the database can be divided into: multi-dimension data analysis and OLAP method.

3 Data Mining in E-Commerce Website

The application of Web data mining technology is data mining in e-commerce site, the English name is"Web Data Mining", a technique that is developed based on a Web environment.It is potentially useful model or information able to collect from complicated Web documents and sites. Web Data Mining technology is an integrated technology, not only related to the computer network technology and artificial intelligence technology, and also involves the discipline of computational linguistics, information science and statistics. Web data mining technology were used for three types of Web data forms : Web content mining, Web structure mining and Web usage mining methods.

E-commerce is a wide range of business types now, and the orderly conduct of electronic commerce on the Internet can't be separated from the support of data mining technology. From the point of view of data mining technology, e-commerce has a sufficient condition for data mining (eg: there are data source richly and reliable data automatically collected and other conditions).And in e-commerce

application of data mining technology, it can provide firsthand information in time for enterprises in business investment, so as to the future development of enterprise decision-making. Here are some key e-commerce website applications, web data mining - data source.

Data source of Web mining in e-commerce is a necessary condition for data mining, and because the Internet is filled with large data sets, so it reflects the necessity and value of data mining. Web data mining techniques are applying the e-commerce website, all related information and data can be structured analysis and processing, and different types of data can be collected and summarized. The six types of data collection can take advantage of Web data mining techniques to produce the different modes of knowledge following.

①Internet server data

Users of Internet visit the site of e-commerce sites leave time data automatically on the Web server and these data is often stored in the form of Web text files on Internet servers (such as: Severlogs the Error logs, cookies, Logs).

②Query information data

Query data is generated on the server layer of the e-commerce Web site for a class of data format, such as: if the online customer to browse through an e-commerce website, you can also search for products and advertising messages that the user viewed before, this style of data on the server through cookies or viewing the registration Information indirect link to the top of the server's access log.

③Online market information and data

This type of data usually refers to traditional relational databases which hold related information of website that related to e-commerce, information of customer purchase and commodity information and so on.

④Web page

Web page of the e-commerce site is the HTLM with XML page contains content information, the main ones being: Web page of the article, picture, voice, image and other content.

⑤Hyperlink between Web pages

Hyperlink relations of this page is the hyperlink relations between the different Web pages in e-commerce site, it is also a very important information and data resources.

⑥the registration information that e-commerce website users want to access

In the daily operation of e-commerce site, thousands of users access to sits every day, and the large number of registration information will be automatically stored. These information and data is entered by the user browsing the Web page or be submitted to the e-commerce site server information, and the information and data is usually related to the quantity and characteristics of users and customers' personal information. In the application of Web data mining technology, registration information and data need to integrate data that users want to access, so that will improve the accuracy of Web data mining, but also facilitate e-commerce website businesses a better understanding of their own customers.

Web data mining technology acquisition mode of knowledge in e-commerce website are: path analysis, discovery of association rules, sequential pattern discovery,

classification rule discovery and clustering analysis of pathways.The applications of Web Data Mining technology in e-commerce site are reflected in the following three aspects:

Firstly, Web data mining need to explore the potential customers initiatively

The browsing behavior of Internet users in e-commerce website reflects the interest and purchase intent of the site's merchandise, for the enterprises engaged in e-commerce website, and the understanding deeply and concerning about such potential customer have good business prospects. At first, you can identify these potential online customer,and then promote them further to become registered customers on their website through commercial means, this behavior is also similar to membership card spending strategy in the real business conduct.If a firm has more registered customers, it indicates that the future of the enterprise increased trade and improving efficiency.

Secondly, we need to improve the design of e-commerce server site

To improve the design of e-commerce site server site, we can start mainly from the following three directions. (1) First of all, through Data Mining of the Web Log, you can selected the correlation between the user page from e-commerce Web site to access , and then add a hyperlink between these there are some associated Web page for later convenience when the user browses ; (2) Web data mining techniques can also be applied in the path analysis, such techniques can analyze the most frequently e-commerce sites Web sites and access path that the user access to, so you can use these information, you can put information of goods and advertising that the bussiness think of the more important in the Web page,at the same time,you can improve Web page and the website structure design, which can increase the e-commerce sites to user's appeal, and it take the convenience of our customers browse, choose and buy, at last,it improves the enterprise' sales, increases the income of the businessman; (3) In data mining to Web Log,we can also analyze the expected position of the registered users of e-commerce site, if access to a desired position in e-commerce website over its access frequency. In order to achieve the optimization of the structure of the Web site e-commerce website, the actual location of the e-commerce site managers can consider to set up navigation links between the desired position and actual position.

Thirdly, clustering e-commerce sites on the customers

Many businesses have customers, market, sales, service and support information in depth excavation and analysis, and they discover the new market opportunities and increase revenues and profits through classify customer value. In e-commerce, customer clustering is an important aspect. By grouping to the browsing behavior of customers that is similar and analyzing the common features of the group of customers, e-commerce can help the organizers understand their customers more better. timely They can meet customer requirements to provide customers with more suitable, more customer-oriented service through adjustment of the pages and the business activities to a certain extent, It will make business activities more meaningful for customers and vendors.

4 Summary

With the rapid development of computer technology and information technology, e-commerce website play an important role in a virtual business activities and occupy an increasing share of e-commerce in commercial trade. Businesses on the Internet can process and analysis a large number of data through the application of data mining technology. And in this way, you can find useful information on the merchants. Businesses can take advantage of these valuable data to grasp the dynamics of customers,to track market changes and to predict the trend of market development in intense competition in the future, to make the most accurate and timely decision-making.It will greatly improve the market competitiveness of enterprises for the enterprises, and lay a solid foundation for the development.

References

[1] Berry, M.J., Linoff, G.S.: Data mining techniques. Machinery Industry Press (2006)
[2] Perner, P.: Advances in Data Mining. Medical Applications, E-Commerce, Marketing and Theoretical Aspects, pp. 20–80. Springer (2008)
[3] Wang, G.F., Xu, D.Q.: Web usage mining in e-commerce personalization service. Hubei University of Economics (Humanities and Social Sciences) (9), 46–47 (2007)
[4] Zhu, Z.-G., Kong, L.: Web mining techniques in e-commerce research and applications. Changsha Telecommunications and Technology Vocational College (1), 32–37 (2007)
[5] Li, S., Shao, L., Li, N.: Artificial neural network BP algorithm. Computer Knowledge and Technical/Academic Exchange (5), 20–22 (2008)
[6] Xiao, H., Hung, F., Zhang, Z.: Such as intrusion detection based on the integration of classification and support vector machine. Computer Simulation 25(4), 130–132 (2008)
[7] Chang, W., Wang, Z., Yan, X.: Intrusion detection system based on integrated neural network research. Computer Simulation 24(3), 134–137 (2007)

Analysis of Security Mechanism in Web Database Application System

Juan Chen

Department of Information Engineering, Hunan Mechanical and Electrical Polytechnic, Changsha City, Hunan Province

Abstract. Although computer technology and database management technology continue to develop..still not to meet the needs of the users of the Web database applications, At the same time, there are serious security risks. this paper introduces the procedure development of the database and architecture of major database system, And to analyze the security risks of Web database application system during operation, the specific measures are proposed to solve these problems.

Keywords: web, database, security mechanism, application system.

1 Introduction

Since the beginning of 60's in the 20 centuries, As the basis of computer data management technology. Database technology provides a complete and comprehensive support for the computer collection, storage, processing data. It played a significant role in promoting the development and application of computers and the Internet. With the continuous improvement of software and hardware, data structures and models evolving. Database technology go through the initial mesh database to level database, developed into a relational database. However, as the popularity and the rapid development of the Internet Relational database in space because of the query database, index. and dealing with problems of higher price limitations. More and more can't adapt to the users to access speed of capacity and needs. A new database (also known as a Web database) based on Web query interface to access database resources become a new favorite.

Web database refers to the backstage that adopt database management system to store the data information. Provide Web page which contains the form as a visiting interface . and the Search results also include datalist in Web page form returned to the user's database system. Web database technology absorbed merits from Web technology and database technology: General browser software to realize the database client function ; Client can use the browser to realize cross-platform and multimedia services but does not take into account the client database design .because database and browser are completely independent, the change of database do not affect the browser software. users also do not affect the operation of the database ; the Web database can be utilized on any platform .so it improved the selection of the software and hardware

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 499–503.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

of freedom . At present, the development of Web database system mainly focused on java platform which is introduced by Sun Microsystems company In May 1995 and NET platform which is introduced by Microsoft company In 2002 . both of them have put into a lot of technical support to realize Web database system development based on B/S/S mode in the two platform. the Java platform of Sun Microsystems company is based on J2EE technology , Due to technical high maturity and can realize the cross-platform development, Web database development have been used extensively in the industry . And net platform Make staff come from different knowledge organization can easily programme as its support for multiple language development.

Together with low cost, good stability, development of high efficiency, make net become the main platform yet in the industry.

2 The Architecture of Web Database System

The database system is composed of database and management software system.. It is a data processing core body which grow at the time when adapt to data processing .it is a practical operation of the storage, maintenance and application of the system that provides the data software system is a aggregate of storage medium, processing object and management system. Database system structure is a relation of computer with database system among all parts.The main task to research database system structure is to study how the hardware to arrange, distribution. How to realize the distribution of each software function. Only from the database of the whole system perspective, as a guide to wholeness, all the parts of the system shall be developed in a well-coordinated manner .The distribution function of hardware, software should be correct .only in this way ,can we establish a database system with clear logic and facilitate with development and maintenance .that is to say, only the correct set of whole system structure can we achieve the whole database system. successfully

The system structure of the database system and the system structure of computer are closely linked . As data calculation mode development , The system structure of database system experienced Centralized mainframe structure, C/S structure,and distribution.

2.1 C/S Architecture

Along with the computer system and the development of the local area network technology . Centralized structure also revealed many shortcomings. and was replaced by Client/server model.at the same time, Centralized mainframe structure was replaced by C/S architecture(Two layers of C/S structure). C/S architecture refers to Communication system which can mate clients and servers (data, documents and network) and operating systems and processes and can support distributed computing, processing, display and print . The resources of the system and application task was assigned to different client and server , Client and server obey the network protocol and connect with each other . the service that client requests is feedback through the server response after its request to the server.

2.2 The B/S Architecture Based on Web

The B/S architecture based on Web is a good way to realize the cross-platform operation, solve the C/S system structure defects

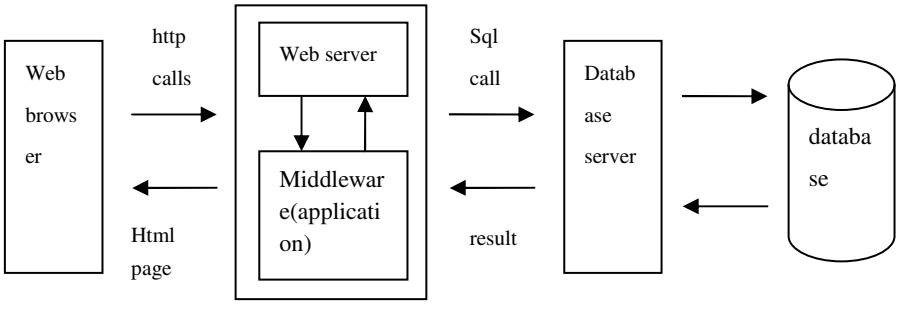
properly. be able to provid clients for text, graphics, image, audio and other services in the best interface. The B/S consider Web browser as a united user interface, provide the data storage and management by Web server, The pattern is also called B/S structure . The B/S structure is a more superior data processing mode which is based on C/S architecture. According to the HTTP protocol , The Web browser could send request that customers submitted to each Web server and receive feedback . Across different platforms to visit the computer program and all kinds of resources. Interface unified, convenient use, information sharing degree higher make Web B/S architecture apply more widely.

2.3 Three Layers C/S/S System Structure

According to the two layers C / S system structure's defects, C/S/S architecture take application server as the middle layer to join between client and server . The parts of the system is divided into three layers according to the function:The client, application server, database server .in C/S/S system structure, To realize the client with thin body and server with more specific function, The system resources and application task was distributed to multiple servers in a more reasonable, scattered way. At the same time also to avoid server getting "fat", the function of each server were scattered to the application server and database server . Such C/S/S structure provides a new mechanism which Could let the application of different components be in different server . Hardware system structure become more concise, more flexibility .

2.4 Three Layers B/S/S System Structure

B/S/S system structure is composed of database server, middleware (system software or service program), Web servers, Web browser and database servers .under the transmission control protocol、HTTP protocol .it Can access the Web server and backstage database server through the browser .



The advantages of B/S/S structure : (1) based on TCP/IP protocol ,it has the strong independence and cross platform. (2) standard unified client program-the browser . (3)it has efficient data storage and management ability , ,to realize the data resources sharing whole hog.

3 The Security System of Web Database System

Web database system consists of the Web browser and Web server, database server and Web server and interface of database management system . Web database system's theory:

The user send request to the Web server through the Web browser. after Server response, connect the database server through the data interface . And perform the corresponding program in the database server. transfer to the customer at last

Web server and the interface of database management system is the key to realize the Web access database. The main function is to establish a connection between Web server and database server and provide the appropriate application program. In view of the different interface technology, there will be different project to support data inquires.

With the development of the Internet. Static HTML technology have already can't meet the requirements. Interactive dynamic network technology be vividly portrayed . after CGI (common gateway interface), Microsoft introduced a kind of typical server web design technology, Namely IIS + ASP solution. Meanwhile, Microsoft push-off the desktop database system-Access system (JET database for the engine),as it has friendly interface, simple operation etc, has a lot of user groups . Microsoft with ASP development based on Web technology of database application use HTTP protocol and HTML format to access data server and the whole system. And can reflect in different Web page. Each page realize mutual visits, calls through the URL. Reflects the strong function of the system. Due to the dynamic server page is based on the server technology. a user want to access the database must go through the Web server . In this way, the application end can control user to act by some means. To ensure that the user can guide and limit access to the database server or system resources . At present, the small and medium website mainly USES the IIS + ASP + Access technology. although the scheme has the advantages of convenient,it also may have the security hidden danger .

4 IIS + ASP+Access Dynamic Website Potential Danger Analysis and Coping Measures

The security hidden danger of IIS + ASP + Access plan mainly comes from IIS server and Access the database's security. Next is the ASP page security holes in the process .for IIS speaking, On the one hand, the earlier version has more security holes . it is necessary for further Web server installation update to fix holes. On the other hand, after installation, IIS attributes will set the default writing, reading, index, log access,

etc . So the wrong Settings may let the client realize the illegal operation website. network database is the soul of Dynamic web site. For the Access database speaking, Once the database was download cracked, the all security mechanism will "non-existing". if there is a professional moral problems, Internet service also can cause the leakage of ASP application source code . In view of the above security measures include:(1)Because IIS is built on Windows operating system , IIS directory permissions rely on Windows NTFS file system access control. Therefore to guarantee the safety of Windows system is the premise and foundation of IIS security.(2) Can use unconventional nomenclature, definite a complex unconventional name for the access database file .meanwhile use data sources (DSN) database connection method prevent effectively Access database be downloaded. (3) Prevent ASP source program and limit ASP page leak access to improve the safety performance of ASP page.

5 Conclusions

Web database application system along with the development of scientific technology develop rapidly . The system resources and application task was more reasonable, scattered distribution to multiple servers, The user's experience is more simple, convenient , But the network environment is full of complexity and variability , Information system also has a greater vulnerability , Network security threats have been and will exist objectively . only we realise dangers, loopholes and the weak link and prevent early, detect early and reorganize early and perfect early. Can we guarantee the safety and reliability of the web site operation.

References

[1] Anido, L., Llamas, M., Fernandez, M.J.: Developing WWW-based highly interactive and collaborative applications using software component. Software Practice and Experience 31(9), 845–867 (2001)

[2] Wang, J.Z., Yin, G.F., Hu, X.B., et al.: Implementing method of remote cooperative design system based on browser/server model. Computer Integrated Manufacturing Systems 6(6), 52–55 (2000)

[3] Saimi, A., Syomura, T., Suganuma, H., et al.: Presentation layer framework of web application systems with server-side Java technology. In: Proceedings of The 24th Annual International Computer Software and Applications Conference, Taipei, pp. 473–478 (2000)

[4] Tang, F., Yao, J.: The analysis and research of Web database access technology. The Computer Engineering 27(7), 88–89 (2001)

[5] Li, C., Xu, S., Yang, Y.: The computer network practical tutorial, vol. (2). Higher Education Press (2004)

The Research and Summary of Evolutionary Multi-objective Optimization Algorithm

Xu Jingqi

Hunan Mechanical and Electrical Vocational Technical College, Hunan Changsha 410151

Abstract. Evolutionary multi- objective optimization mainly studies how to use evolutionary calculation method to solve the multi-objective optimization problem.it has become a hot research topic in the field of evolutionary computation. however, multi-objective evolutionary algorithm based on the concept of Pareto optimal is the research hotspot of current evolutionary calculation. Based on the comparison and analysis of multi-objective optimization evolutionary algorithm. Introduce some major technology and the theoretical results of multi-objective evolutionary algorithm which is based on the concept of Pareto optimal .

Keyword: Multi-objective, Pareto, Ooptimal calculation.

1 Introduction

As a kind of heuristic search algorithm , evolutionary algorithm has been applied successfully to the multi-objective optimization field, developed to a relatively hot research direction-- evolutionary multi objective optimization, Referred to Optimal problem is one of the main problem form in engineering practice and scientific research. Among them, the optimization problem only has one objective function is called single objective optimization problems. the optimization problem has the objective function more than one and need to be handled is called multi-objective optimization problems. For multi-objective optimization problems, A solution, to a specific goal, it is good, but to other target may be poor. Therefore, there is a compromise solution set, call Pareto optimal set or non-dominated solutions. At first, the multi-objective optimization problems are often transformed into a single objective problem by the way of weighted, then use mathematical programming method to solve, every time can only get the optimal solution in a weight. At the same time, as the objective function and constraints function of multi-objective optimization problems may be non-linear, non-differentiable or discontinuous. Traditional mathematical programming method often is low efficiency, and is more sensitive for weights and given target sequence. Evolutionary algorithm achieved global search by the populations composed with potential solution in generation by generation, the method, from population to population, is very useful for Pareto optimal solution set which search multi-objective optimization problems.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 505–512.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 The Mathematics Description of Multi-objective Optimization Problems

Multi-objective optimization problem, also known as multi- standard optimization problems. It do not break the general, a multi-objective optimization problem With an n decision variables, m target variables can be expressed as

$$\begin{cases} \min y = F(x) = (f_1(x), f_2(x), ..., f_m(x))^T \\ s.t. g_i(x) \le 0, i = 1, 2, ..., q \\ h_j(x) = 0, j = 1, 2, ..., p \end{cases}$$

Among them, ,x=(x1...xn)∈ $X \subset R^n$ is an n-dimensional decision vector, X is a n-dimensional decision space. y=(y1...ym)∈ $Y \subset R^m$ is the m dimension target vector, Y is the m dimension target space. The objective function F (x) defines m the mapping function from decision space to the target space. $g_i(x) \le 0, (i = 1, 2, ..., q)$ defines the q inequality constraints. $h_j(x) = 0, (j = 1, 2, ..., p)$ define the p equality constraints. On this basis, gives the following important definitions.

Definition 1 (feasible solution). For some x ∈ x, if meet the (1)constraint conditions , $g_i(x) \le 0, (i = 1, 2, ..., q)$ and $h_j(x) = 0, (j = 1, 2, ..., p)$, Said the x feasible solution.

Definition 2 (feasible solution set). All the feasible solution set of X is called a feasible solution collection, notes for Xfand $X_f \subseteq X$ 。

Definition 3 (Pareto dominant). Hypotheses that XA, XB are two feasible solution which showed in type (1) multi-objective optimization problems, said XA is Pareto dominant compared with XB. When $\forall i = 1, 2, ..., m$, $f_i(x_A) \le f_i(x_B) \wedge \exists j = 1, 2, ...m$, $f_j(x_A) < f_j(x_B)$ notes for $x_A \rangle x_B$, Also known XA dominate XB.

Definition 4 (Pareto optimal solution). A solution x*∈Xf is called Pareto optimal solution (or non dominated solutions).when and only when the following conditions are satisfied: $\neg\exists x \in X_f : x \rangle x^*$

Definition5 (Pareto optimal solution set). Pareto optimal solution set is all Pareto optimal solution set, defined as follows

$$P^* = \left\{ X^* \middle| \neg\exists x \in X_f : x \rangle x^* \right\}$$

Definition 6 (Pareto frontier). The curved surface of target vector corresponding to the all Pareto optimal solutions in Pareto optimal solution set P * is called Pareto frontier PF *:

$$PF^* = \left\{ F(\mathrm{x}^*) = \left(f_1(x^*), f_2(x^*), ..., f_m(x^*) \right)^T \middle| x^* \in P^* \right\}$$

3 The Main Algorithm of the Evolution Multiple Objective Optimization

3.1 The First Generation of Evolution Multi-objective Optimization Algorithm

The first generation of evolution multi-objective optimization algorithm is started by Goldberg . In 1989, Goldberg suggested to use non dominated sorting and niche technique to solve multi-objective optimization problems, the non dominated sorting process is: for the current population of non-dominated individuals, we assigned rank 1; and move it from competition. Then selected non-dominated individuals from the current population, and to assign rank 2. This process is not end until all individuals in a population are assigned. Niche technology is used to keep the diversity of population, prevent premature. Although Goldberg didn't put his ideas to the specific implementation of evolutionary multi-objective optimization, it has instructive meaning for the later scholars. Subsequently, some scholars putted forward the concept of MOGA, NSGA and NPGA based on the thought.

(1)MOGA

Fonseca and Fleming putted forward MOGA in 1993. The method, grade for each individual, all non-dominated individual level is defined as 1, the other individual level is that dominate its individual number plus one. the individual with the same level select by fitness sharing mechanism. Its fitness distribution obey the following way . : first of all, the population rank by grade, then, distribute fitness level to all individuals. Use linear or nonlinear interpolation method presented by Goldberg to

distribution, with the same level of the fitness values are the same. use random sampling to select through fitness sharing mechanism. MOGA is dependent excessively

on sharing function selection, and may has a greater selection pressure, leading to the premature convergence.

(2)NSGA

NSGA is also designed on non dominated sorting thinking, at first non dominated solution is determined, then was assigned a large virtual fitness value. In order to maintain the population diversity, these non- dominated solutions share with their virtual fitness . These non-dominated individuals are disregarded temporarily. identify the second batch of non-dominated individuals from the remaining populations. And then they are assigned a virtual fitness value which is smaller compared with the minimal fitness value that non-dominated individuals has shared . These non-dominated individuals also does not consider temporarily, identify the third batch of non-dominated individuals from the remaining populations. The process will last until the entire population were divided into several grades. NSGA copy a new generation by proportional selection. NSGA's calculation complexity is O(mN3), in which, M is the target number, IV is the population size. Its computational complexity is high, and need to confirm a predetermined sharing parameter .

(3)NPGA

NPGA designed tournament selection mechanism, which is based on Pareto dominance relationships. The concrete thought are as follows: selected randomly two individuals from the evolutionary population and then selected randomly a comparison set from the evolving population. If only one of the individuals is free from the comparison set control, the individual will be selected to go into the next generation. When all of them dominated or be dominated to the comparison set, use a niche technology to choose one of them sharing into the next generation . the algorithms is to select individual with big shared adaptation value into the next generation . In this algorithm, the selection and adjust of J, radius habitat are difficult .and need a suitable choice of the comparison set scale.. The first generation of evolutionary multi-objective optimization algorithm is charactered by non dominated sorting selection and diversity based on sharing function. with the development of the first generation evolution multiple objective optimization , some of the problems need to be solved have raised . First of all, by finding alternative for niche (sharing function) to maintain the population diversity. Fitness sharing is proposed by Goldberg and Richardson for multimodal function optimization. Usually need to have hypothesis of prior knowledge about peak limit number and the space habitat uniform distribution. For a multi-objective optimization problem, and also need to make sure priori information of sharing radius, the computational complexity is the square of population size.

3.2 The Second Generation of Evolution Multi-objective Optimization Algorithm

From the beginning of the 20th century, the field of evolution multi-objective optimization has changed a lot .in 1999, Zitzler putted forward SPEA, The method enables the elite reservation mechanism become popular in the field of evolutionary multi-objective optimization. the born of second generation evolutionary multi-objective optimization algorithm is mark of the elite reservation strategy.

In the field of evolutionary multi-objective optimization, elitism strategy refers to use an external population (relative to the original individual populations) to retain the non-dominated individuals.

(1)SPEA and SPEA2

Zitzler and Thiele put forward the SPEA algorithm in 1999, In this algorithm, the fitness of individuals is also known as the Pareto strength. individual fitness of non dominating set is defined as the innervations of the individual in the proportion of groups, other individual fitness is defined as the dominating individual number plus 1, appoint lower fitness individual corresponds to a higher probability of selection. In addition to the evolutionary population, also set up a external population save current nondominated individuals ,when the external population individual number are beyong the specified value, we use clustering technology to cut individual . from evolutionary groups and external select individual stocks into a mating pool by the tournament selection , cross, mutation operation. The computational complexity of this algorithm is high up to the cubic.

Of population size. SEPA2 is the improved version of SPEA which is presented by Zitzler and Thiele in 2001. They improve it in the fitness assignment strategy, individual distribution evaluation method as well as the non-dominated set . In SPEA2, the individual fitness function is f)=R(f)+D(f), Among them, R (I) take individuals control information of individual I into accoun in external populations and the evolution individuals, 09(0 is the crowded degree measurement which is determined by individual I to its K adjacent individual distance. In the construction of new group, the first choice is the environment, and then the mating. In the environment choice, the first choice is the fitness less than 1 to the external population. When the number of individuals are less than external population size, select individual who is low fitness in evolutionary population; when the individual is greater than the number of external population size, use environmental selection to cut. In mate selection, use the championship

mechanism to choose the individual into a mating Pool . SPEA2 introduced environmental choice based on nearest neighbor rule ,simplifies external population regeneration method of SPEA based on cluster. Although the computational complexity is still the cubic of population size, but, the solution distribution uniformity based on the nearest neighbor rule environment selection cannot be transcended.

(2)PAES,PESA and PESA—II

PAES (1+1) use evolutionary strategy to present a solution for mutation, and then evaluate the variation result, comparing it with the variation of the previous individual dominance relation, retain the better through the elite reservation strategy. the classic of this algorithm is to introduce space super lattice mechanism to maintain the diversity of population, each individual distribution into a lattice, The time complexity is O (Nx), where, IV for evolutionary population size, size for population external. The algorithm of space super grid strategy was later adopted by many evolutionary multi objective algorithm. Subsequently, Come. Based on this space super lattice, proposed PESA. PESA set up an internal population and an external population, in evolution, the internal non-dominated individuals incorporated into external populations. When a new individuals come into the external population, at the same time, cut an individual in external populations. The specific method is to look for congestion coefficient maximum individual in external populations and delete it, if the multiple individuals with the same crowding coefficient, delete a randomly. An individual crowding factor refers to the number of individual corresponding super lattice . in 2001,Come further improved the PESA .Called PESA-II. Proposed concept based on region selection, compared with PESA based on individual choice, PESA. II use grid selection instead of individual choice, to a certain extent, improves the efficiency of algorithm.

(3)NSGA-II

NSGA - II is the improvement of NSGA algorithm by Deb,so far, it is the most outstanding evolutionary multi-objective optimization algorithm, NSGA-II has the following advantages: (1)The new hierarchical fast non-dominated sorting decreased the method computational complexity from O (mN3) to O (mN2), of which, m is the target function number , N is individuals number in population.(2) to mark the various elements fitness value after fast non-dominated sorting, at the same time ,make the individual of current Pareto frontier can be extended to the entire Pareto frontier, And even as much as possible.

The algorithm presented crowding distance concept, use the crowding distance comparison algorithm instead of fitness degree sharing method in NSGA, the time complexity of crowding distance is O (m (2N) log (2N)).

Introduced the elite reservation mechanism, the individuals who are selected and attended reproductive to produce offspring with its parent individuals competing together to produce the next generation of population, so it is conducive to maintaining excellent individual, improve the overall evolution level . NSGA-II, SPEA2and PESA-II is the classical algorithm of second generation evolutionary multi-objective optimization.

the algorithm, with elitist strategy as the main character, and most algorithms no longer use fitness sharing niche technology as the means to maintain the diversity of the population.

Some better strategy is proposed, such as clustering, method based on crowding distance, and method based on space super lattice.

4 Conclution

4.1 Research and Application Trends

In real life, a large number of MOEA research has been applied to many fields. According to the available references, it can be divided into: Engineering, industry and science. The important application of engineering field : Electronic Engineering, hydraulic engineering, structural engineering, aeronautical engineering, robotics and control, Industrial applications include: design and manufacturing, scheduling and management; scientific fields including chemistry, physics, medicine, computer science and so many academic research and practice..

4.2 Summary

Each time, the algorithm is limited by research level. The first generation evolutionary multi-objective optimization problems has problems about how to combine the evolutionary algorithm and multi objective optimization. The second generation began to consider the efficiency of the algorithm, and the current time researchers are further improving the calculation efficiency, at the same time, face how to deal with high dimension multi-objective optimization problem.

Evolutionary multi-objective optimization algorithm also has some problems, mainly reflected in the following two aspects: First of all, if the evolutionary multi-objective optimization algorithm use small populations, for some complex multi-objective optimization problem, it will be very difficult to converge to the desired Pareto frontier, butand it is difficult to obtain the uniform distribution of Pareto optimal solution.

Therefore, according to the complexity of the problem to adjust the size of population is the issue that needs further research. According to the non dominated solutions proportion in the current population to adjust population size adaptively may be a feasible direction to the problem, in addition, how to design an effective algorithm to stop criterion is also a basic theory problem of evolutionary multi-objective optimization.

So far, it is difficult to judge clearly that algorithm has reached the optimal or can stop. In the single objective optimization process, if the population 's fitness in algebra did not improve,we can identify algorithm can stop. This stop strategy is not simply applied to multi objective optimization, because multi objective optimization get is not a solution, is Pareto optimal solution set. Even if it has already converged to a Pareto frontier optimal solution set, then the solution each generation obtained also is different. Therefore, scholars generally set in advance the evolution algebra or function evaluation times ceiling as the conditions of algorithm stopping. But, how to prove effectively or give a definite algorithm stopping criterion remains to be studied.

References

[1] Xie, T., Chen, H., Kang, L.: Multi-objective optimization evolutionary algorithm. Journal of Computers 26(8), 997–1003 (2003)
[2] Zheng, X., Liu, H.: The research progress of multi-objective evolutionary algorithms. Computer Science 34(7), 187–192 (2007)
[3] Zheng, J.: Multi-objective evolutionary algorithm and its application. Science Press, Beijing (2007)
[4] Jiao, L.C., Du, H., Liu, F., Gong, M.: Immune optimization calculation, learning and recognition. Science Press, Beijing (2006)
[5] Cui, X.: Multi-objective evolutionary algorithm and its application. National Defence Industry Press, Beijing (2006)
[6] Wang, L.: Advances in quantum-inspired evolutionary algorithms. Control and Decision 23(12), 1321–1326 (2008)
[7] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multi-objective genetic algorithm, NSGA—II. IEEE Trans. on Evolutionary Computation 6(2), 182–197 (2002)
[8] Zitzler, E., Thiele, L.: Multi-Objective evolutionary algorithm s: A comparative case study and the strength Pareto approach. IEEE Trans. on Evolutionary Computation 3(4), 257–271 (1999)
[9] Deb, K.: Multi—Objective Optimization Using Evolutionary Algorithms. John Wiley & Sons, Chichester (2001)
[10] Li, Z., Liu, S., Xiao, D.: Multi-Objective Particle Swarm Optimization Algorithm Based on Game Strategies. In: GEC 2009, Shanghai, China, June 12-14, pp. 287–293 (2009)
[11] Wang, L.: Advances in quantum-inspired evolutionary algorithms. Control and Decision 12(23), 1321–1326 (2008)
[12] Li, Z., Rudolph, G.: A Framework of Quantum-inspired Multi-Objective Evolutionary Algorithms and its Convergence Condition. In: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2007), London, UK (2007)
[13] Li, Z., Rudolph, G.: On the Convergence Properties of Quantum-Inspired Multi-Objective Evolutionary Algorithms. CCIS, vol. 2, pp. 245–255 (2008)
[14] Li, Z., Rudolph, G.: Convergence Performance Comparison of Quantum-inspired Multi-Objective Evolutionary Algorithms. Computers and Mathematics Application

A Finite-Time Convergent Variable Structure Guidance Law with Impact Angle Constraint

Zuo-e Fan[1], Xian-jun Shi[1], Chang-peng Pan[2], and Wen-jin Gu[2]

[1] Department of Control Engineering, Naval Aeronautical and Astronautical University, Yantai 264001, China
[2] Department of Commanding, Naval Aeronautical and Astronautical University, Yantai 264001, China
fanfan4417@sohu.com.

Abstract. For attack the maneuvering target at a given direction, on the principle of mathematics, the guidance law based on traditional Lyapunov stability can only converge the line of sight angular rate to the zero, and converge the ballistic azimuth to the expected angle, when the time approaches infinity. Base on the stability theory of nonlinear control systems, a finite-time convergent variable structure guidance law with impact angle constraint is proposed. The simulation of some anti-ship missile against high maneuvering target has made, and the results shown the rightness and effectiveness of the proposed method.

Keywords: Guidance Law, Variable Structure, Finite Time Convergence, Impact Angle, Maneuvering Target.

1 Introduction

For most homing weapons, the goal of guidance systems is to offer the appropriate guidance commands to achieve the zero miss distance. However, in some situations, the zero miss distance may not be the only tactical mission. Sometimes, impact angle constraint is required in guidance applications to enhance the effect of the warhead and increase its kill probability. Many researchers proposed the guidance laws considering the final impact angle. A three-dimensional missile guidance law with angle constraint based on sliding mode control was proposed in [1]. Daekyu Sang proposed the impact angle control guidance law based on Lyapunov stability theory and parameter optimization method, which can minimize miss distance and adjust impact angle to predefined value [2]. Ashwini Ratnoo formulated the guidance problem with impact angle constraint as an infinite horizon non-linear regulator problem, and solved the problem using State Dependent Riccati Equation technique with time varying state weight matrix [3]. All these methods have their own advantages, but on the principle of mathematics, the closed guidance loop can only converge at the equilibrium point, when the time approaches infinity. So there is limitation on the point of theory.

All goals of the guidance system, converge performance is an important goal. On the aspect of optimization, the control method with finite-time convergence is an optimal time method. The problem of finite-time convergence control is, whether the control

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 513–519.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

system can be controlled to converge and keep at the equilibrium point in finite time. According to the characters of nonholonomic mobile robot, [4] proposed a method of finite time tracking control algorithm based on the continuous states. [5] Discussed the finite-time feedback stabilization problem of a class of second-order nonlinear systems, and offered three kinds of global finite-time feedback stabilization approaches based on continuous state feedback. [6] used the finite-time convergence stability theory to design the smooth second- order sliding mode, and offered some principles of second-order sliding mode design. Although there are many references about the design of finite-time convergence controller in different control domain, but the reference about the applications of finite-time convergence stability theory on the missile guidance system is very few. [7] used the principle of second-order sliding mode and applied it on the missile guidance design, but the design process of guidance law is complex, and it is difficult to apply on the engineering. Based on the reference [8], using the finite-time convergence stability theory of nonlinear system, a finite-time convergent variable structure guidance law with impact angle constraint is proposed in this paper. The guidance law is simple and convenience to realize in the engineering.

2 Model of Missile-Target Relative Motion

Generally speaking, hitting targets with impact angle constraints is for the vertical plane, for STT anti-ship missile, the guidance law can be designed separately, and the design process and principle on the vertical plane are the same of the yaw plane. For convenience, the scheme to be designed is based on the yaw plane.

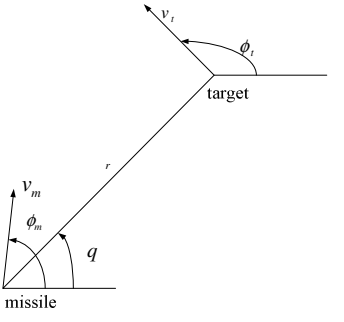


Fig. 1. Missile-target engagement geometry

The kinematical equations between missile and target on the yaw plane can be shown as (1)

$$\dot{r} = v_t \cos(q - \phi_t) - v_m \cos(q - \phi_t) \tag{1-a}$$

$$r\dot{q} = -v_t \sin(q - \phi_t) + v_m \sin(q - \phi_t) \tag{1-b}$$

Where r– the distance between missile and target; q –line of sight (LOS) angle; v_m ,–velocity of missile and target; ϕ_m , ϕ_t –flight-path angle of missile and target. Derivate (1), it can obtain that

$$\ddot{r} = r\dot{q}^2 + w_r - u_r \tag{2}$$

$$\ddot{q} = -2\dot{r}\dot{q}/r + w/r - u/r \tag{3}$$

$$w_r = \dot{v}_t \cos(q-\phi_t) + v_t\dot{\psi}_T \sin(q-\phi_m) \tag{4}$$

$$u_r = \dot{v}_m \cos(q-\phi_m) + v_m\dot{\phi}_m \sin(q-\phi_m) \tag{5}$$

$$w = v_t\dot{\phi}_t \cos(q-\phi_t) - \dot{v}_t \sin(q-\phi_t) \tag{6}$$

$$u = v_m\dot{\phi}_m \cos(q-\phi_m) - \dot{v}_m \sin(q-\phi_m) \tag{7}$$

Where w_r , u_r – the target acceleration component and the missile acceleration component along LOS ; w , u – the target acceleration component and the missile acceleration component prependicular to LOS.

The initial time of terminal guidance is $t = 0$,and the initial states are $r_0 . \dot{r}_0 . q_0$ and $\dot{q}_0$; the terminal states of guidance system are $r(t) . \dot{r}(t) . q(t)$ and $\dot{q}(t)$ .Set the state $x = [x_1 \quad x_2]^T$, where $x_1 = q - q_d$, $x_2 = \dot{q}$, q_d is the expected LOS, $x_0 = [x_{10} \quad x_{20}]^T = [q_0 - q_d \quad \dot{q}_0]^T$,then (3) can be rewritten as the following

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\dfrac{2\dot{r}(t)}{r(t)}x_2 - \dfrac{1}{r(t)}u + \dfrac{1}{r(t)}w \end{cases} \tag{8}$$

3 Finite-Time Convergence Stability Theory

In order to design the finite-time convergent guidance law, the definition of the finite-time convergence stability of the nonlinear system is offered as following [9].

Considering a dynamic system of the form

$$\dot{x} = f(x,t), \quad f(0,t) = 0, \quad x \in R^n \tag{9}$$

Where $f : U_0 \times R \to R^n$ is continuous in $U_0 \times R$, and U_0 is an open vicinity of the origin $x = 0$. We call the equilibrium $x = 0$ is finite-time (local) convergence,

provided that there exists the time $T \geq 0$, which is depend on x_0, such that the solution of (9), $x(t) = \varphi(t; t_0, x_0)$, has definition (maybe not only one) at any given initial state $x(t_0) = x_0 \in U$, and satisfies

$$\begin{cases} \lim\limits_{t \to T(x_0)} \varphi(t; t_0, x_0) = 0 \\ if \quad t > T(x_0), \ \ then \quad \varphi(t; t_0, x_0) = 0 \end{cases} \tag{10}$$

when $t \in [t_0, T(x_0))$, $\varphi(t; t_0, x_0) \in U / \{0\}$. We call the equilibrium $x = 0$ of the system is finite-time (local) stability provided it is Lyapunov stability and finite-time convergence at the vicinity of the origin $U \subset U_0$. If $U = R^n$, then the origin is globally finite-time stability.

Based on the finite-time stability theory of the nonlinear system, there is the lemma as follows.

Lemma. Considering the nonlinear system (9), if there exists a continuous function $V(x,t)$ defines on the vicinity C^1 of the origin $\widehat{U} \subset R^n$, and exists constant s $\alpha > 0$ and $0 < \eta < 1$ such that $V(x,t)$ is positive definite on $\widehat{U}$, and $\dot{V}(x,t) + \alpha V^{\eta}(x,t)$ is negative definite on $\widehat{U}$, then the origin of the system is finite-time stability.

The prove process refers to [9].

4 Finite-Time Convergence Guidance Law

In order to design the finite-time convergent guidance law of the system (8), a sufficient condition that the guidance system is finite-time convergent is offered as follows.

Theorem1. Considering guidance system (8), if the controller u satisfies

$$(\mu x_1 + x_2)[(\mu \dot{x}_1 + \dot{x}_2) + \frac{\beta \operatorname{sgn}(\mu x_1 + x_2)}{r(t)}] < 0 \quad \forall t \geq 0 \quad \mu > 0 \tag{11}$$

Where $\beta, \mu > 0$ is constant, the state of the guidance system is finite-time convergent to zero.

Proof. Choose the smooth positive define function

$$V = (\mu x_1 + x_2)^2, \quad \mu > 0 \tag{12}$$

Based on (11) and (12), we can obtain

$$\begin{aligned}\dot{V} &= 2(\mu x_1 + x_2)(\mu \dot{x}_1 + \dot{x}_2) \\ &\le -\frac{2\beta \operatorname{sgn}(\mu x_1 + x_2)}{r_2(t)}(\mu x_1 + x_2) \\ &\le -\frac{2\beta}{r_2(t)} V^{0.5}, \forall t \ge 0\end{aligned} \tag{13}$$

During the terminal guidance, the following inequality is satisfied

$$V = (\mu x_1 + x_2)^2, \quad \mu > 0 \tag{14}$$

So,
$$\dot{V} \le -\frac{2\beta}{r_2(t)} V^{0.5} < -\frac{2\beta}{r_{20}} V^{0.5}, \forall t \ge 0 \tag{15}$$

According to the lemma, the state x of the guidance system can be convergent to zero in finite-time t_{r1}

$$t_{r1} < \frac{|x_{10} + x_{20}| r_0}{\beta} \tag{16}$$

Based on (8) and (11), the following equation can be obtained

$$(\mu x_1 + x_2)[(\mu \dot{x}_1 + \dot{x}_2) + \frac{\beta \operatorname{sgn}(\mu x_1 + x_2)}{r(t)}] = (\mu x_1 + x_2)[x_2 + \frac{\beta \operatorname{sgn}(\mu x_1 + x_2)}{r(t)} + \frac{2|\dot{r}(t)|}{r(t)} x_2 - \frac{1}{r(t)} u + \frac{1}{r(t)} w] \tag{17}$$

Design the guidance law

$$u = \beta \operatorname{sgn}(\mu x_1 + x_2) + w + r(t)(1 + \frac{2|\dot{r}(t)|}{r(t)}) x_2 + k|\dot{r}(t)|(\mu x_1 + x_2), \quad k > 0 \tag{18}$$

Substitute (18) into (17), we can obtain

$$(\mu x_1 + x_2)[(\mu \dot{x}_1 + \dot{x}_2) + \frac{\beta \operatorname{sgn}(\mu x_1 + x_2)}{r(t)}] = -\frac{|\dot{r}(t)|}{r(t)} k(\mu x_1 + x_2)^2 < 0 \tag{19}$$

Based on the the Theorem1, (18) is the guidance law to converge the state of the system to zero in finite time. β is bigger, the speed of $\dot{q}$ converge to zero and q converge to q_d is faster .

w is the acceleration of target, it cannot be measured accurately, it only be estimated approximately [8], choose a positive constant ε, such that the super of the target acceleration satisfies $|w| \le \varepsilon$, and substitute the w with $\varepsilon \operatorname{sgn}(x_1 + x_2)$ in (18) , the guidance law becomes

$$u = (\beta + \varepsilon) \operatorname{sgn}(\mu x_1 + x_2) + [r(t) + 2|\dot{r}(t)|] x_2 + k|\dot{r}(t)|(\mu x_1 + x_2) \tag{20}$$

Based on the former analyses, we can obtain the following theorem.

Theorem 2. During the guidance process, if $\Delta = \beta + \varepsilon - |w| > 0$ is satisfied at any time, guidance law (20) can converge the state of the system (8) to the zero, that is the guidance law can accurately attack the target at the expected direction.

Proof. Substitute (20) into (8), we can obtain

$$\dot{x}_2 = -x_2 - \frac{k|\dot{r}(t)|(\mu x_1 + x_2)}{r(t)} - \frac{(\beta+\varepsilon)\mathrm{sgn}(\mu x_1 + x_2) - w}{r(t)} \tag{21}$$

According (21) and $\Delta = \beta + \varepsilon - |w| > 0$, the following inequation can be obtained

$$\begin{aligned}(\mu x_1 + x_2)[(\mu \dot{x}_1 + \dot{x}_2) + \frac{\Delta \mathrm{sgn}(\mu x_1 + x_2)}{r(t)}] = -\frac{k|\dot{r}(t)|}{r(t)}(\mu x_1 + x_2)^2 - \\ \frac{(\beta+\varepsilon)\mathrm{sgn}(\mu x_1 + x_2) - w - \Delta \mathrm{sgn}(\mu x_1 + x_2)}{r(t)}(\mu x_1 + x_2) < 0\end{aligned} \tag{22}$$

According to lemma1, the guidance law can converge the state of the system to zero in finite time t_{r2}. And Δ is bigger, the speed of $\dot{q}$ converge to zero and q converge to q_d is faster.

$$t_{r2} < \frac{|x_{10} + x_{20}| r_0}{\Delta} \tag{23}$$

5 Simulation Result

The initialization location of target is (2000km, 0), the target moves at the speed of $v_t = v_{t\ \mathrm{sup}} = 20m/s$, and snaking maneuver is carried out by the target to avoid attack from missile. The azimuth of the target's velocity varies strongly from -600to +600 ruled by sine function.

The initialization location of missile when the guidance system begin to work is (0, -3000km), the velocity of missile is 250m/s, the initial ballistic azimuth of missile is 00. The expected ballistic azimuth of missile is -300. The parameters of guidance law are chosen as $\beta = 20, \varepsilon = 20, \mu = 1, k = 0.03$ according to the principle mentioned above in the system analysis. Substitute the function $\mathrm{sgn}(s)$ with the saturation function $s/(|s| + \delta)$, and the parameter in the saturation function is $\delta = 0.05$.

In simulation, the overload command imposed on missile is restricted 0-8g, considering that the missile's available overload is limited in practice. The simulation result shows in Fig.2-Fig.3.The miss distance is 0.3m, the impact ballistic azimuth is -30.10.The simulation result shows the validity and effectiveness of the proposed guidance law.

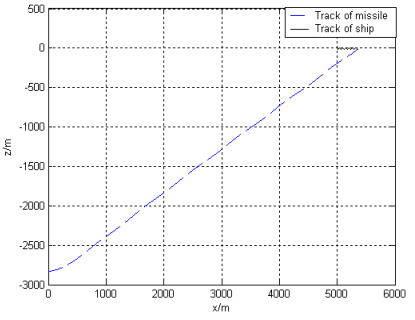


Fig. 2. Track of missile and target

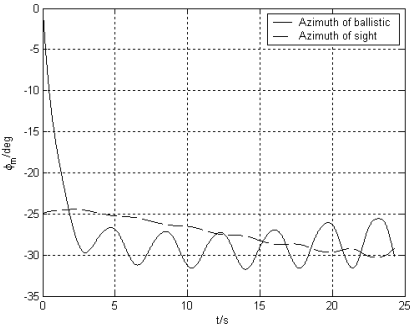


Fig. 3. Time history of sight and ballistic azimuth

6 Conclusion

A new guidance law with angle constraint based on finite time stability theory and variable structure theory is proposed in this article. And the performance is investigated in the presence of target maneuvering. The simulation results have shown the rightness and effectiveness of proposed method.

Acknowledgements. The authors would like to thank the referees, whose suggestions improved the presentation of the paper.

References

[1] Gu, W.-J., Fan, Z.-E., Zhang, X.-J.: Design of variable structure guidance law with constraints on impact angle for anti-ship missile. In: IEEE International Conference on Computer Design and Applications, Qinhuangdao, China, vol. 4, pp. 480–484 (2010)

[2] Sang, D., Min, B.-M., Tahk, M.-J.: Impact angle control guidance law using lyapunov function and PSO method. In: SICE Annual Conference, pp. 2253–2257. Kayawa University, Japan (2007)

[3] Ratnoo, A., Ghose, D.: ADRE based guidance law for impact angle constrained trajectories. In: AIAA Guidance, Navigation and Control Conference and Exhibit, AIAA 2007-6539, Hilton Head, South Carolina (2007)

[4] Li, S.-H., Tian, Y.-P.: Finite time tracking control algorithm for nonholonomic mobile robots. Control and Decision 20(7), 750–754 (2005)

[5] Li, S.-H., Ding, S.-H., Tian, Y.-P.: A Finite-time State Feedback Stabilization Method for a Class of Second Order Nonlinear Systems. Acta Automatica Sinica 33(1), 101–104 (2007)

[6] Levant, A.: Principles of 2-sliding mode design. Automatica 43, 576–586 (2007)

[7] Shtessel, Y.B., Shkolnikov, I.A., Levant, A.: Smooth second-order sliding modes: missile guidance application. Automatica 43, 1470–1476 (2007)

[8] Sun, S., Zhou, D.: A Finite Time Convergent Variable Structure Guidance Law. Journal of Astronautics 29(4), 1258–1262 (2008)

[9] Hong, Y.-G., Cheng, D.-Z.: Analysis and Control of Nonlinear Systems, pp. 225–228. Press of Science (2005)

Highway Traffic Automatic Detection System Based on Video and Image Processing

Jiang Ying

Jiangxi Science and Technology Normal University, Nanchang, Jiangxi, 330013, P.R. China
ncjiangying@126.com

Abstract. Automatic Incident Detection (AID) system, provides a new technical means for the highway intelligent traffic management, improves the efficiency and reduces the work intensity of the highway management department. This paper takes the French Citlog video traffic incident automatic detection system for the technical background, introduces the principle, functions, technical characteristics and the preliminary project implementation plan of automatic detection system.

Keywords: Highway, video and image, traffic incident, automatic detection.

1 Introduction

With the rapid development of our country's highway construction, Highway traffic management scale expands rapidly, totally dependent on human to monitor the health of the vehicle on the road can not meet the needs of current increasing complex traffic management. Intelligent transport system is the development trend of 21 century world road traffic, and advanced emergency management system is an important part. In terms of operation and management of highway, in order to prevent or reduce the losses caused by traffic accidents, timely and effectively rescue and treat for accidents, effectively reduce the traffic delays generated by the traffic accidents and avoid secondary accidents, We must be able to accurately and quickly to detect traffic incidents and traffic anomaly. Fast automatic traffic incident detection system which is established based on video and image processing, makes full use of current highway television monitoring system. It uses image processing methods to process video signals so as to quickly and accurately detect and identify different traffic conditions' traffic incident, and simultaneous detects visibility, snow, rain and other weather anomalies, this will effectively overcome the traditional event detection systems' ills of detecting traffic volume, speed, and share changes to detect event occurring, achieve efficient, accurate and rapid detection of events, provide conditions for events' rapid processing, Significantly reduce traffic delay and the second accident. According to the development trend of current domestic highway, highway intelligent traffic management will become the development direction of modern traffic management. This paper gives a brief description of video traffic incident automatic detection subsystem of the intelligent system which provides reference for the implementation of monitoring and detection system.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 521–526.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 System Overview

Most of the highways in China, uses traditional CCTV highway video monitoring system (For example, Jiangxi Taijing highway whole monitoring system), this monitoring system only sets up series cameras per certain distance along the highway, and transmits the video and images to control center through the communication system, Monitoring staff sit before the monitoring center display wall to monitor the road traffic. This traditional CCTV system is passive, When the video and images displayed on the wall are too much, this Often becomes a decoration, monitoring staff can not monitor a large number of video images in 24 hours. Video traffic incident detection system is different from the traditional CCTV system, Video traffic incident detection system is positive, It uses the latest technology of video and image processing, through the computer to track and identify all the video, so as to automatically detect highway traffic incidents(Traffic jams, suspended vehicles, retrograde vehicles, slow vehicles, the pedestrians, lost goods), It is a breakthrough in video monitoring systems, also is the development direction of highway intelligent management, has great significance in improving highway capacity and ensure traffic safety.

2.1 Video Detection Principle—Trace Detection Technology

Video detection technology automatically detects road accident events in real time through video data acquisition and processing. It takes from the CCTV (closed circuit television monitoring), or video and image which is specifically set and captured by cameras. Camera shoot 5 images per second, from the change of the image sequence select target information for computer processing. From the images，The analyzer bases on algorithm to generate event accident alarm and traffic measurement information，these information is then transferred to the manager of traffic management center .

2.2 The Functions of Video Detection System

The performance of video detection system:

Average detection alarm time: The average time from the occurrence of the event to detection of alarm is 10 seconds; Traffic incident detection rate： detection rate =The number of event detected / The number of event occurring >95%; False alarm rate：False alarm rate = The number of false alarms / the number of days that operation is in use，<one/ per camera in 40 hours

The functions of video detection system: Traffic incident alarm: Depending on the setting, the analyzer can produce a variety of alerts, sends to the data server, forwards to the manager.

Traffic event incident detection and alarm: Analyzer automatically detects vehicles stopping and running，traffic jams, pedestrians, vehicle retrograde, derelicts, smoke and fire. If the fire is detected, promptly notify the administrator, Fire vehicles are generally parked, so Firstly test suspended vehicle. As the visibility of the highway with smoke is very low, Fire will result in reflection in the highway pavement and

wall. The system can grade the priority of the alarm to avoid multiple alarms on the same event accidents.

Parameters alarm: When the measurement exceeds the threshold, automatically generate traffic measurement alerts, such as, the queue length exceeds the threshold, the speed exceeds the threshold. All the thresholds are set when the system is installed, it can automatically adjust according to traffic conditions.

Self-diagnosis alarm of detection system: System runs self-diagnostic program in the analyzer, automatically detects the following errors on the server: Camera position movement; camera signal loss, analyzer error, network communication failure.

Parameter measurement: System provides real-time traffic measurement for each lane: flow, speed, occupancy, headway, queue length, vehicle type. Each measurement result is for a lane or a set of lanes.

3 The Composition of the Video Detection System

3.1 The Composition of Automatic Detection System

Highway video traffic incident detection system has the following components:

Video collection: the cameras set up along the highway have two type: far-range and close-range. Far-range cameras have PTZ, can detect the vehicle stopping and running and traffic jams these two kinds of traffic incident; close-range cameras are the fixed cameras (can be used to replace the current vehicle detectors, the current vehicle detectors system is almost impossible to function properly ,because drive coil damage caused by road damage), can detect vehicle s stopping and running , ttraffic jams, the slow vehicle, pedestrian, vehicle retrograde, loss of goods and other traffic incidents, simultaneously can detect traffic data, such as Flow, speed, occupancy, headway, queue length, etc.

Video Transmission: digital optical and cable

Video processing: Computer systems (including hardware and intelligent image processing, control software), traffic incident analyzer.

Video control and output: Matrix, monitor and displaying walls, etc.

In order to facilitate management of road traffic safety, the two cameras in control point is responsible for the collection of traffic, road share rate and other traffic data, at the same time use its video for traffic event incident detection. Cloud desktop camera can be installed on roads and dangerous sections of accidents. Video monitors camera images, can be used for traffic incident detection.

3.2 The Composition of the Video Processing System

The video processing system is composed of video detection analyzer, data server, manager, digital video recorder, switches and other equipment.

Video detection analyzer is included in the computer unit, is composed of image processing software, video signal digital system and the communication card. A video detection analyzer can simultaneous handle eight fixed cameras' video signals or simultaneous handle four cloud desktop cameras' video signals. According to the number of monitoring points, traffic management control center deploys the corresponding number of image analyzer.

Video detection analyzer uses image processing algorithm to obtain the four or eight digital and synchronous analog video for image processing. On the hard disk of the analyzer, do the video recording, alarms storage, measurements and images, and communicate with the server in cycle.

Data server: This can choose PC server of mainstream configuration, its function is data storage and management (including event accident data, video sequence, traffic parameters, etc.), also equipped with an Ethernet card for communication with central monitoring system, exchanging alarm information, and dumping alert file, Etc. According to the system' storage requirements, extend storage disk can be configured, in order to meet the special needs of different users.

Manager is a standard desktop computer, generally placed on the console command, in most cases configure one, but can support more. It can complete communication with the server, Concentrate alarms, measurements and video sequences, manage, configure and maintain the system, maintain interface remotely. Through the Ethernet switch to form a star LAN, Network connects all analyzer, server and manager. Communication uses TCP/IP protocol (alarm, traffic measurement and video sequences).

Video System provides several video options:Event accident video sequence: The storage and management of event accident video sequence is a powerful tool, it can rapidly analyze problems occurring on the road(Event accident, accident), or record under the external trigger for depth analysi s of events after the accident occurring(Law enforcement needs, The needs of other road users before and after the accident or events).

For each camera, analyzer stores digital video sequences of event and accident occurrence time in the cache (the recording time of events sequence can be adjusted, usually set before the events of 1 minute + 2 minutes after the incident), when an event or accident occurred, save video sequences and extra time span: the given time after events or accident occurred, or to the end of the event or accident.

Permanent digital video: This feature allows the administrator to do digital permanent recording of all the cameras. This video has nothing to do with any event or accident, and the video can be stored in the circular buffers for 30 days (The time depends on circumstances). A specific human-machine interface allows the administrator to manually check the video at any time during 30 days before, without the need for additional digital video recorder.

Administrator video mode: In any time, administrators can manually activate or stop a video camera, view and archive related video recording in the database. This can not rely on accidents or events to video.

3.3 The Camera's Setting Principle and the Number of Installations

The camera's setting principle: Set the interval of monitoring point 800-1200M (to ensure coverage); In the straight sections, winding mountain road, tunnel estuary, bridges, separation embankment' variable road department, black spot and other key locations, set up cameras intensively. A monitoring point set two cameras, each is responsible for one direction (one PTZ and one fixed, and also facilitate the transmission of video, propose PTZ camera to set a preset position).

4 Video and Image Transmission and Power Supply

4.1 The Way of Image Transmission

The image transmission of video traffic incident automatic detection system uses digital optical and cable. The system will transmit video and images to IPC of monitoring points so as to IPC processing the image, and then through data optical transmit to traffic management and control center. The video points which are far from traffic management and control center, send image and data by optical to near toll station, then through multiple-optical send the neighboring images and data together to traffic management and control center. The video points which are near traffic management and control center, send direct to traffic management and control center through optical.

4.2 The Power Supply of Monitoring Points' Device

Monitoring points' devices sett by video traffic incident automatic detection system are generally far from fees management department. Laying of power cables, for the video traffic incident automatic detection systems, the cost of this approach is relatively high. The total electrical load of one single monitoring point of video traffic incident automatic detection system is more than 30 W, so using solar power can be relatively cost-effective.

4.3 The Advantages of Video Detection System

The Traffic event video signal of video detection system, comes from the CCTV surveillance cameras' signal, has no special requirements. When accident occurred in the video detection system, Management and traditional monitoring system can be able to respond immediately. Timely automated alarms, promptly treating, avoid similar accidents to occur one after another. The system's high reliability and low false alarm rate can fully automated carry out traffic control, and process black and white or color camera images. Within the highway video can cover Multi-lane, All the detector and the test parameters are adjustable, distance can up to 500 meters (set position) For weak illumination' highway, the viaduct, ramp and other areas, video detection system can ensure work continuous 24 hours.

Video detection system is a modular, flexible system, gets rid of the limitations of road coil sensors: closure construction, aging of the road easily, Etc. this system is base on "the architecture of "Client - server", uses standard hardware and communication protocols, has a friendly user interface, and has self-diagnostic function. System includes an database of event and accident video sequence, has traffic data of each lane (speed, occupancy, flow, etc.).

5 Conclusions

Traffic event automatic detection system based on video and image, can process video signal, is construction simple, features strong, the system can big or small, has strong expansion. The use of traffic event automatic detection technology based on

video and image, will further improve the intelligent management system of highway traffic monitoring, control and command, and provide a powerful guarantee for the highway operation and management.

References

[1] GA/T514-2004, the engineering acceptance of traffic TV monitoring system (2004)
[2] Liu, W.: The application of video vehicle detector in the highway. Highway (February 2003)
[3] Wang, M., Zhu, M., Wang, H.: Research and implementation of key technologies in Motor vehicle' video speed. Computer Engineering (May 2006)

An Improvement to Leach-Based Routing Algorithm

Liu Re and Wang Xiu-Ping

College of Software and Service Outsourcing, Wuxi Professional College of Science and Technology Wuxi, Jiangsu, China
leach_wq@sina.com

Abstract. Wireless Sensor Networks consists of a huge number of small self-contained devices with computational, sensing and wireless communication capabilities. LEACH is a classical clustering-based protocol in wireless sensor network. In this paper, we propose an improved routing algorithm LEACH-WQ based on the analysis of LEACH protocol, which involves cluster head choosing, multi-hop routing and the building of its path. As a result the number of cluster head is constant and near optimal per round. Simulation results prove that the routing algorithm proposed by us has higher energy utilizing rate, and it helps prolonging the network lifetime.

Keywords: WSN, ROUTING, PROTOCOL, ROUTING ALGORITHM.

1 Introduction

Wireless Sensor Networks (WSN) consists of a huge number of small self-contained devices with computational, sensing and wireless communication capabilities. WSN aims to apperceive in collaborative mode, gather, deal with and send information to observer in network areas. Sensor, sensing object and observer form the three factors in WSN.

WSN protocol stack contains physical layer, data link layer, network layer, transport layer and application layer. According to network architecture, routing protocols are generally classified as plane routing, and level routing. Typical plane routing protocols are DD [1], SAR [2], SPIN [3] and etc; Typical level routing protocols are LEACH [4], PAGASIS [5], and TEEN [6] etc.

LEACH is the earliest cluster-organized routing protocol. Compared to ordinary plane multi-hop algorithm, this kind of low energy consuming, self-adaptive, cluster-organized algorithm, which is specially designed for Wireless Sensor Network, could prolong the lifetime of network by 15 percent.

LEACH is the shortening of Low-Energy Adaptive Clustering Hierarchy. In LEACH the nodes organize themselves into local cluster with one node acting as the local base or cluster-head. LEACH uses randomization to distribute the energy load evenly among sensors in the networks, so the useful system lifetime is extended. However, there are some defects in LEACH worth further improving.

2 Relative Work

Reference [1] and [2] make some improvement for LEACH on routing selection and energy management. Reference [7] calculates the optimum cluster-head number and

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 527–533.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

optimizes the selecting process, but that is only suitable for small-scale networks. The algorithm in [8] takes three factors of a node into consideration which are the remaining energy, the location in the cluster and the distance to sink. Although the compartmentalization of clusters is more uniform, some nodes may die too early. In the course of selecting cluster-head, the method in [9] also considers the remaining energy, however it maximums performance at the cost of increasing the complexity of networks.

Cluster head selection algorithm, adopted by LEACH protocol, avoids fast energy loss of cluster heads and its data aggregation effectively reduces the amount of communication. Therefore, LEACH protocol prolongs the network life time in contrast to plane multi-hop routing and static routing.

However, there are still some improvements to be done in the LEACH protocol. Firstly, the probability in selecting cluster heads is equivalent, without considering the remaining energy of nodes. Therefore, those nodes with less remaining energy may be chosen as the cluster heads which will lead to fast energy loss of these nodes, hence making them invalid. Secondly, the number of cluster heads is fluctuating heavily. The stochastic cluster-head selection of LEACH is prone to result in the unbalanced clusters partition in the network and thus increase the total energy dissipated in system. This paper analyzes the defects of LEACH and proposes an improved protocol: LEACH-WQ.

3 Study on Leach Protocol

LEACH is a kind of self-adaptive cluster-organized topological algorithm. Nodes organize themselves into clusters, one node in every cluster would acts as cluster head [10] [11].

The process is executed in periodical manner; every round is divided in two phases: cluster building phase and stable data communication phase. In the phase of cluster building, close nodes make a cluster dynamically, and one certain be selected as cluster head randomly; In the phase of stable data communication, nodes in one cluster would send their date to the cluster head, then cluster head would use the data and send it to sink node. Because cluster heads need to fuse the data and communicate with sink node, they consume more energy than ordinary nodes.

LEACH algorithm could guarantee that every node in one cluster would be selected as cluster head in equal possibility, which makes every node consume energy relatively equally. The procedure of selecting cluster head in LEACH is like following, Every node produces a random number between 0 and 1, and if this number is less than threshold value *T(n)*, then it pronounces itself as cluster head. In every round, if one node has been cluster head before, then *T(n)* is set to 0, so that this node will not be selected again. For the nodes that have not been selected once, the possibility of being selected is *T(n)*. As the number of nodes which have been cluster head increases, T (n) will increase, so the possibility for the rest nodes to be selected will increase. When there is only one node left, *T(n)* =1, which means this node will be selected surely. *T(n)* could be defined as follows,

$$T(n)=\begin{cases}\dfrac{p}{1-p\cdot(r\cdot \bmod(\frac{1}{p}))}, n\in G\\ 0, n\notin G\end{cases} \tag{1}$$

Where p is the desired percentage of cluster heads, r equals the current round, and G is the set of nodes that have not been cluster-heads in the last $1/p$ rounds. Then, operation moves to the steady phase, the steady phase is divided into frame, where nodes send their data to the cluster head at most once per frame during their allocated transmission slot. After a period, the entire networks are renewed, preparing for the new round. The flow chart of LEACH algorithm is shown in Figure 1.

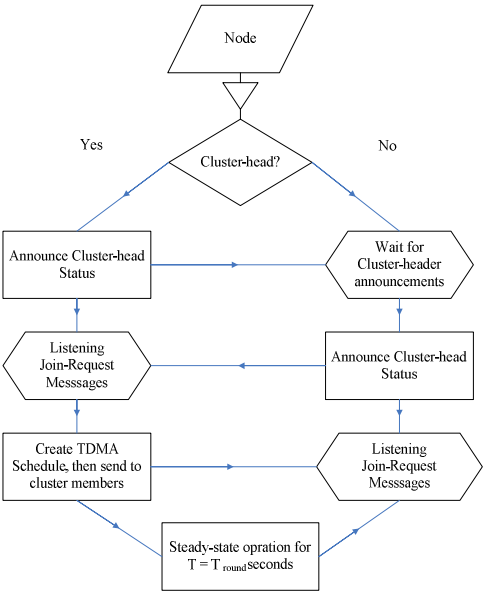


Fig. 1. LEACH Flow Chat

Cluster heads are selected randomly in LEACH, it is possible that nodes with less energy would be chosen, which may lead to these nodes die too fast. In addition, since in LEACH protocol cluster heads communicate with base stations in single-hop range, it is energy consuming and its expandability is limited.

Therefore, it cannot adapt to large network. In the following, we will discuss how to improve LEACH algorithm from two points: the way to choose cluster heads and multi-hop routing.

4 The Proposed Leach-Wq Protocol

In consideration of the defect of LEACH, as description prior, we put forward an improved protocol called LEACH-WQ. This enhanced protocol adds a second selection of cluster heads to modify the number cluster-head in the set-up phase considering the node's residual energy per round.

In order to save the energy consumption and to prolong the life span of the network, the protocol needs to ensure that the partition of cluster is balance and uniform. To achieve this goal, the number of WQs needs to be dominated, and the network needs an

optimal WQs amount. In the process of cluster head choosing, we introduce the concept of Current Energy.

Assume the initiative energy of every node is equal. In LEACH algorithm, the possibility of being cluster head is equal for every node. When a node is selected as a cluster head, it will broadcast this message, which includes the *ID* of cluster head. Other nodes would decide which the cluster to join, according to the intensity of received signal. Then they send Join-Request message, which contain cluster head's *ID* and their selves' *ID* and Current Energy. According to other nodes' current energy, cluster head could work out the average energy of the cluster (Eave). Then, it will broadcast Eave and TDMA time-slice assigning table to every node in the cluster. When selecting a cluster head next time, if one node's randomly produced number between 0 and 1 is less than $T(n)$, it could not be selected as a cluster head immediately.

The Path Weight formula is below,

$$W'_{ch} = \alpha \cdot \frac{E_r}{E_{node}} + \beta \cdot \frac{T_{node}}{\sum T_{node}} + \delta \cdot \frac{d(node, WQ)}{D_{n-ch}} \tag{2}$$

Where d (*node*, *WQ*) is the distance between one node and its cluster-head, which can be calculated by analyzing the signal strength; $D_{n\text{-}ch}$ is a constant that the maximum distance permitted between one node and its cluster-head; α, β and δ are weight factors, $\alpha + \beta + \delta = 1$ and $\alpha, \beta, \delta \in [0,1]$. Where E_{node} is initial node energy; ΣT_{node} is the sum of Trust of all close neighbor; x, y are weight factors, and $x, y \in [0, 1]$; the meaning of other symbols is expressed above. After having chosen the node whose weight is the heaviest, a canceling packet which is used to cancel the node itself as cluster-head can be constructed. Then the node transmits the packet at a time randomly chosen from 0 to T_1.

$$T(n) = \begin{cases} \dfrac{p \cdot w}{1 - p \cdot [r \cdot \bmod(\frac{1}{p})]}, n \in G \\ 0, n \notin G \end{cases} \tag{3}$$

In a stable phase, we introduced a cluster head node to the communication cost $Cost(V_i, CH_j)$, its relationship with the distance from the node to the cluster head $d(V_i, CH_j)$.

The distance to the base station BS $d(CH_j, BS)$, the remainder of the current node and the monitoring of regional energy state of K on (in this case, we denote K as a constant, in fact, with the entire monitoring area is proportional to the square diameter),

$$Cost(V_i, CH_j) = \beta \cdot \frac{d^2(V_i, CH_j)}{Energy} + (1-\beta) \cdot \frac{d^2(CH_j, BS)}{K} \tag{4}$$

In the algorithm of multi-hop routing, not every cluster head is connected to base station. One nearest cluster head *A* will be chose to connect base station, and then the nearest cluster head to *A* (for example, *B*) will be chose to connect to *A*, and so on, until all the clusters are added to this multi-hop route.

We use "head" to represent the set of multi-hop route's cluster heads, and the initial value of "head" is base station, that is head={*BS*}; Assume "path" is the set of route path, and its initial value is NULL, that is path={ }; Assume "*C*" is the number of cluster heads in every round; *d(i)* is the distance between cluster head *i* to base station; *d(i, j)* is the distance between cluster *i* and *j*; the initial value of variable "distance" is infinity. for (m=1; m<=T;++m) {if (d (m) < distance){Distance =d (m);w=m;}} After circulation, "w" indicates the nearest cluster head to base station. So, head= {*BS, k*}, path= {line (*BS,k*)}. Then find the nearest node to w.

for (p=1;p<=C && p!=head;++p){if (d(p, k)<distance){Distance=d (p, k);j=p;}}

After this circulation, the nearest node *j* is added to head set, that is head= {*BS, k, j*}, path= {*line (BS, k), line (k, j)*}; after *k=j* is executed, the circulation continues. In the following rounds, the nodes in set head should not be counted. In the end, head= {*BS, k, j …*}, including *C* cluster heads, and *BS.* And *C* paths are worked out too.

5 Simulation Results and Analysis

Parameters for Simulation We use similar random 100-node networks and radio models with LEACH.

This paper set the following parameters for Simulation:

1) Energy required in sending or receiving 1bit: Eelec =50nJ/bit
2) The amount of data sent by nodes each time: k = 200bit.
3) The initial energy of every node: E = 0.5J
4) Compress ratio during data fusion: r = 0.7
5) Energy consumed in every bit data fusion: EA=50pJ/bit
6) Area: 100*100
7) The location of Sink: (6, 6)
8) The percentage of cluster head: p = 0.06
9) The number nodes: n = 200

Fig.2 illustrates the relation curve which indicates the relationship between the number of nodes in the network with and the remaining time. It can be seen from the Fig.2 that due to energy consumption of the agreement, the time of LEACH-WQ nodes to begin to die later than the LEACH protocol, a longer survival time, mainly due to our first election in the cluster caused by a weight *W* and non cluster head nodes join the corresponding cluster to take into account the communication cost.

We also note that death of some of the steeper more quickly in the relation curve. However, the LEACH-WQ algorithm performs better energy efficiency throughout the network and prolongs the network lifetime. As shown in Fig. 2, we could also see that after the same number of rounds, in the improved LEACH algorithm the total energy consumed is less than in original LEACH algorithm. Moreover, the energy is exhausted after 849 rounds in original algorithm and it is after 1000 rounds in the improved LEACH-WQ algorithm proposed by us. Therefore, we can see that the network's lifetime is prolonged, and the energy is used more effectively.

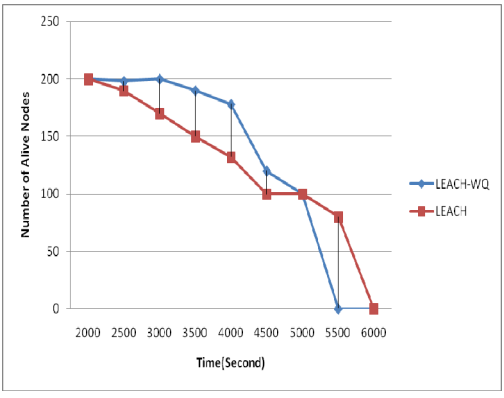


Fig. 2. Time and Number of Alive Nodes

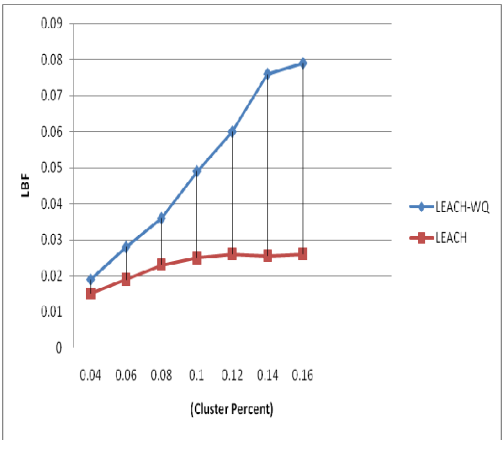


Fig. 3. Cluster Percent and LBF

As shown in Fig.3, the two algorithms under different percentage of cluster heads gain significant different LBF values, the proposed protocol LEACH-WQ performs better load balancing, while the LBF is increasing lineally with the substantial increasing of cluster heads proportion.

6 Conclusion

Wireless Sensor Networks consists of a huge number of small self-contained devices with computational, sensing and wireless communication capabilities. LEACH is the earliest cluster-organized routing protocol. Compared to ordinary plane multi-hop algorithm, it gains characteristics of low energy consuming, self-adaptive and cluster-organized. In this paper, to overcome the weakness of LEACH protocol in cluster head choosing algorithm and multi-hop routing and the building of its path, we

propose an improved algorithm based on LEACH. As proved in simulation in MATLAB, we argue the improved algorithm LEACH-WQ prolongs the lifetime of network and raises the energy effectively.

Acknowledgement. This work is supported by the Qing Lan Project Fund (NO. 200830) of Jiangsu Province, China.

References

[1] Intanagonwiwat, C., Govindan, R., Estrin, D., et al.: Directed diffusion for wireless sensor networking. IEEE/ACM Transaction on Networking 11(1), 2–16 (2003)
[2] Sohrabik, Gaoj, Ailawadhiv, et al.: Protocols for selforganization of a wireless sensor network. IEEE Personal Communications 7(5), 16–27 (2000)
[3] Heinzelmanwr, Kulik, J., Balakrishnan, H.: Adaptive protocols for information dissemination in wireless sensor networks. In: Proceeding of the ACM MobiCom 1999, pp. 174–185. ACM Press, Seattle (1999)
[4] Heinzelman, W., Chandrakasan, A.: Energy-efficient communication protocol for wireless micro sensor networks. In: Proceedings of the 33th Hawaii International Conference on System Sciences (2000)
[5] Lindsey, S., Raghavenda, C.S.: PEGASIS: Power efficient gathering in sensor information systems. In: Williamson, D.A. (ed.) Proc. of the IEEE Aerospace Conf., pp. 1125–1130. IEEE Press, New York (2002)
[6] Manjeshwar, A., Agrawal, D.: TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. In: Proceedings of the 15th Parallel and Distributed Processing Symposium, pp. 2009–2015. IEEE Computer Society, San Francisco (2001)
[7] Xu, D., Zhang, H.: Improved algorithm of LEACH protocol introducing residual energy. Computer Engineering and Applications II 45(28), 115–119 (2009)
[8] Thein, M.C.M., Thein, T.: An Energy Efficient Cluster -Head Selection for Wireless Sensor Networks. In: Proceedings of ISMS 201O (2010)
[9] Heinzelman, W., Chandrakasan, A., Balakrishnan, A.: An Application-specific Protocol Archiecture for Wireless Microsensor Networks. IEEE Transactions on Wireless Communications 1(4), 660–670 (2002)
[10] Anastasi, G., Conti, M., Di Francesco, M., Passarella, A.: Energy Conservation in Wireless Sensor Networks: a Survey. Ad Hoc Networks (2009)
[11] Heinzelman, W.B., Chandraskan, A.P., Blakrisshnan, A.H.: An Application-Specific Protocol Architecture for Wireless Microsensor Networks. IEEE Trans. on Wireless Communications 1(4), 660–670 (2002)

Specifying API Trace Birthmark by Abstract Interpretation

Ying Zeng, Fenlin Liu, Jian Chen, and Bin Yan

Zhengzhou Information Science and Technology Institute, Zhengzhou, 450002, China
zengying510@yahoo.com.cn

Abstract. API trace birthmark is a major class of software birthmarks, where API sequences are defined as software birthmarks to detect software theft. Currently, many birthmarks of this class have been proposed, but the evaluation of these birthmarks is mainly done through experiments and there is no theoretical framework, which makes it difficult to formally analyze and certify the effectiveness of the birthmarks. To solve this problem, an abstract interpretation-based method for specifying API trace birthmark is proposed in this paper. First, API trace birthmark is characterized as a semantic program property by abstract interpretation. Then, the credibility of API trace birthmark with respect to a specific criterion for copy relation is formally analyzed. Finally, the resilience of API trace birthmark is discussed and it is proved that API trace birthmark is resilient to a commonly used program transformation.

Keywords: Software birthmarks, Abstract interpretation, API trace, Program semantics, Credibility, Resilience.

1 Introduction

Software piracy is the act of making unauthorized copies of software or reusing the code of programs without permission. Along with the rapid development of software industry and Internet, software piracy has become a major concern for many software companies and IT sectors. Software birthmark is a property-based technique to detect software piracy [1]. It does not embed extra information into the original program, and instead it extracts inherent characteristics from the program to identify the originality of the program. Currently, most of the works of software birthmark focus on how to define suitable program characteristics as software birthmarks to detect software piracy. According to software birthmarks' representations, we classify software birthmarks into three categories: instruction sequence-based birthmark, API-based birthmark and graph-based birthmark.

API-based birthmark, which is a major class of software birthmarks, is based on the observation that the way a program uses the standard libraries or system calls (we'll collectively call them APIs from now on) is not only unique to that program but also difficult for an attacker to forge [2]. Tamada et al. [2], [3] presented three algorithms for collecting birthmarks, which compute the birthmark from the sequence of method calls within a class, the inheritance path from the root class to a class, and the types a class uses respectively. In addition, Park et al. [4] regarded static API calls as birthmarks of

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 535–545.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Java programs to detect software thefts. And Choi et al. [5] proposed a static API birthmark for Windows executables, which uses sets of API calls statically identified by a disassembler. Besides these statically computed birthmarks, there are some dynamic API-based birthmarks as well. Tamada et al. [6] proposed to compute a birthmark from a trace of well-known library calls that are executed from a run of the program on a particular input. Schuler et al. [7] combined the idea of k-gram with API-based birthmark and presented a dynamic k-gram API-based birthmark, where sequences of well-known API calls executed for a particular input are collected and the birthmark is constructed as a set of k-grams from the API call sequences.

From the above literatures on API-based software birthmark, we observe that the structures of API-based birthmark can be sets, sequences, or graphs. And since sequence structure is more precise than set structure and less difficult to manipulate than graph structure, API birthmarks that use sequence structure attract more concern. In this paper, we collectively call API birthmarks that use sequence structure as API trace birthmark, and focus on API trace birthmark. We specify API trace birthmark based on abstract interpretation, in order to formally analyze and certify the effectiveness of software birthmarks. First, API trace birthmark is modeled as a semantic program property by abstract interpretation. Then, a specific criterion for the copy relation is given and the credibility of API trace birthmark with respect to the given criterion is proved. Moreover, the resilience of API trace birthmark is formally analyzed as well. The commonly used transformation, substitution of equivalent commands, is considered as an attack against API trace birthmark. And it is proved that API trace birthmark is resilient to substitution of equivalent commands.

2 Brief Introduction on Abstract Interpretation and the Programming Language

Since we specify API trace birthmark by abstract interpretation, we first recall some basic results of the abstract interpretation framework [8]. Then we introduce the language to be used in this paper, an extension of the simple imperative language.

2.1 Abstract Interpretation

Abstract interpretation [8] was originally developed by Cousot and Cousot as a general theory for designing and approximating the fixpoint semantics of programs. It provides a method to approximate the concrete semantics, which is often infinite and not decidable, by a sound abstract semantics, so that the results of abstract execution can give some information on the actual computations. The concrete and the abstract semantic domain are related through a Galois connection. In addition, abstract interpretation can also be equivalently formalized by upper closure operators instead of Galois connections. uco(C) denotes the set of all upper closure operator of C.

For a program $P \in \mathbb{P}$ chosen in the set $\mathbb{P}$ of all possible programs, the concrete semantics of the program P is written as $\mathcal{S}[\![P]\!]$. In order to ensure the soundness of the approximation of $\mathcal{S}[\![P]\!]$, the abstract semantics $\mathcal{S}[\![P]\!] \sqsubseteq \gamma\left(\overline{\mathcal{S}}[\![P]\!]\right)$ should satisfy the condition that $\mathcal{S}[\![P]\!] \sqsubseteq \gamma\left(\overline{\mathcal{S}}[\![P]\!]\right)$. For a semantic domain $po\langle \mathfrak{D}, \sqsubseteq \rangle$, the semantic

ordering $\sqsubseteq$ induces an ordering $\trianglelefteq$ on the domain $\mathbb{P}$ of programs, where $P \trianglelefteq Q \triangleq (\mathcal{S}[\![P]\!] \sqsubseteq \mathcal{S}[\![Q]\!])$. Thus, $po\langle \mathbb{P}/\eqsim, \trianglelefteq \rangle$ is a poset, and $\mathbb{P}/\eqsim$ denotes the classes of syntactically equivalent programs, where $P \eqsim Q \triangleq (\mathcal{S}[\![P]\!] = \mathcal{S}[\![Q]\!])$.

Cousot et al. [9] introduced a general uniform language independent framework for designing program transformations by abstract interpretation. According to the framework, programs can be considered as an abstraction of their semantics, so that it leads to the following Galois connection:

$$\mathrm{po}\langle \mathfrak{D}, \sqsubseteq \rangle \underset{\mathcal{S}}{\overset{\mathcal{P}}{\rightleftarrows}} po\langle \mathbb{P}/\eqsim, \trianglelefteq \rangle . \tag{1}$$

The correspondence between syntactic and semantic program transformations is illustrated in Fig.1.

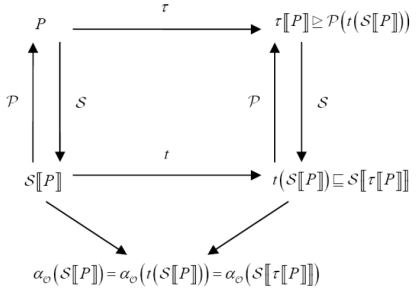


Fig. 1. Syntactic-semantic program transformation [9]

The Extension of the Simple Imperative Language

The language we consider in this paper is an extension of the simple imperative language introduced by [9]. We extend the simple imperative language into a language with procedures. The syntax of the language is given in Table 1.

Table 1. The syntax of the extended simple imperative language

Syntactic Categories		Syntax	
$n \in \mathbb{Z}$	(integers)	$E ::=$	$n \mid X \mid E_1 \; op \; E_2$
$X \in \mathbb{X}$	(variable names)		$(op \in \{+, -, *, \ldots\})$
$L \in \mathbb{L}$	(labels)		
$E \in \mathbb{E}$	(integer expressions)		
$B \in \mathbb{B}$	(boolean expressions)	$B ::=$	$true \mid false \mid \neg B_1 \mid$
			$B_1 \vee B_2 \mid E_1 < E_2$
$M \in \mathbb{M}$	(function names)		
$A \in \mathbb{A}$	(actions)	$A ::=$	$X := E \mid X := ? \mid skip \mid B \mid$
			$call\ M \mid return$
$C \in \mathbb{C}$	(commands)	$C ::=$	$L : A \to L'$
$F \in \mathbb{F} \triangleq \mathbb{M} \times \mathbb{C}$	(functions)		
$P \in \mathbb{P}$	(programs)	$\mathbb{P} ::=$	$\wp(\mathbb{C})$

For a given set $\mathcal{S}$, $\wp(\mathcal{S})$ denotes its powerset, i.e. $\wp(\mathcal{S}) \triangleq \{T | T \subseteq S\}$. Let $var[\![D]\!]$ be the set of variables of an expression or an action $D \in \mathbb{E} \cup \mathbb{B} \cup \mathbb{A}$. $\not{X}$ denotes the undefined variable. $\not{M}$ denotes the undefined function name. A function $F \in \mathbb{M} \times \mathbb{C}$ is a pair $\langle M, C \rangle$, where M is the name of the function F and C is the entry command of the function F. A stop command $L: stop$ is syntactically equivalent to the command $L: skip \to \not{L}$ and $L: true \to \not{L}$, where $\not{L}$ denotes the undefined label. The actions of commands can be classified as assignment, call, return, and test action. We denote the set of all call actions as $\mathbb{A}_c$. The auxiliary functions in Table 2 are useful in defining the semantics of the considered programming language, which is described in Table 3.

Table 2. Auxiliary functions

Label	Variables
$lab[\![L: A \to L']\!] \triangleq L$	$var[\![L: A \to L']\!] \triangleq var[\![A]\!]$
$lab[\![P]\!] \triangleq \cup_{C \in P} lab[\![C]\!]$	$var[\![P]\!] \triangleq \cup_{C \in P} var[\![C]\!]$
Action of Command	Entry Command of Function
$act[\![L: A \to L']\!] \triangleq A$	$entry[\![M]\!] \triangleq C$, where $\langle M, C \rangle$ is a function
Function called in Action and Command	
$cfunc[\![call\ M]\!] \triangleq M$	
$cfunc[\![C]\!] \triangleq \begin{cases} cfunc[\![act[\![C]\!]]\!] & \text{if } act[\![C]\!] \in \mathbb{A}_c \\ \not{M} & \text{otherwise} \end{cases}$	

An environment $\rho \in \mathfrak{E}$ maps variable $X \in dom(\rho)$ to its value $\rho(X)$. The value can be the uninitialized or undefined value $\perp$, so $\mathfrak{B}_\perp \triangleq \mathfrak{B} \cup \{\perp\}$. op is extended to the undefined value $\perp$ as $\perp op \perp \triangleq \perp op\ n \triangleq\ n\ op \perp \triangleq \perp$. $\mathfrak{E}[\![P]\!]$ denotes the set of environments of a program P whose domain is the set of program variables.

Since a function might be called during the execution of another function, a stack is required in order to recover the right return point. Hence, we extend the program states with a stack, which is a finite sequence of pairs that consist of function names and the corresponding return labels. A program state s is a pair $\langle \rho, k, C \rangle$, where C is the next command that has to be executed in environment ρ, k is the current stack. Let $\Sigma \triangleq \mathfrak{E} \times \mathfrak{K} \times \mathbb{C}$ denote the set of all possible states, in particular $\Sigma[\![P]\!] \triangleq \mathfrak{E}[\![P]\!] \times \mathfrak{K} \times P$ denote the set of states of program P.

The transition relation $\mathbf{C}: \Sigma \to \wp(\Sigma)$ between states specifies the set of states that are reachable from a given state. Moreover it can be specified with respect to program P, $\mathbf{C}[\![P]\!]: \Sigma[\![P]\!] \to \wp(\Sigma[\![P]\!])$:

$$\mathbf{C}[\![P]\!](< \rho, k, C >) \triangleq \{< \rho', k', C' > \in \mathbf{C}(< \rho, k, C >) | \rho' \in \mathfrak{E}[\![P]\!] \wedge k' \in \mathfrak{K} \wedge C' \in P\}. \quad (2)$$

Table 3. The semantics of the extended simple imperative language

Value Domains	
$\mathfrak{B}_\perp \triangleq \{true, false, \perp\}$	(truth values)
$n \in \mathbb{Z}$	(integers)
$\mathfrak{V}_\perp$	(variable values)
$\rho \in \mathfrak{E} \triangleq \mathbb{X} \to \mathfrak{V}_\perp$	(environments)
$k \in \mathfrak{K} \triangleq (\mathbb{M} \times \mathbb{L})^*$	(stacks)
$\sum \triangleq \mathfrak{E} \times \mathfrak{K} \times \mathbb{C}$	(program states)
Boolean Expressions $\mathbf{B} : \mathbb{B} \times \mathfrak{E} \to \mathfrak{B}_\perp$	
$\mathbf{B}[\![true]\!]_\rho \triangleq true$	
$\mathbf{B}[\![false]\!]_\rho \triangleq false$	
$\mathbf{B}[\![E_1 < E_2]\!]_\rho \triangleq \mathbf{E}[\![E_1]\!]_\rho < \mathbf{E}[\![E_2]\!]_\rho$	
$\mathbf{B}[\![\neg B]\!]_\rho \triangleq \neg \mathbf{B}[\![B]\!]_\rho$	
$\mathbf{B}[\![B_1 \vee B_2]\!]_\rho \triangleq \mathbf{B}[\![B_1]\!]_\rho \vee \mathbf{B}[\![B_2]\!]_\rho$	
Arithmetic Expressions $\mathbf{E} : \mathbb{E} \times \mathfrak{E} \to \mathfrak{V}_\perp$	
$\mathbf{E}[\![n]\!]_\rho \triangleq n$	
$\mathbf{E}[\![X]\!]_\rho \triangleq \rho(X)$	
$\mathbf{E}[\![E_1 \; op \; E_2]\!]_\rho \triangleq \mathbf{E}[\![E_1]\!]_\rho \; op \; \mathbf{E}[\![E_2]\!]_\rho$	
Actions $\mathbf{A} : \mathbb{A} \times \mathfrak{E} \to \wp(\mathfrak{E})$	
$\mathbf{A}[\![skip]\!]_\rho \triangleq \{\rho\}$ $\quad \mathbf{A}[\![call\ M]\!]_\rho \triangleq \{\rho\}$	
$\mathbf{A}[\![return]\!]_\rho \triangleq \{\rho\}$ $\quad \mathbf{A}[\![X := E]\!]_\rho \triangleq \{\rho[X := \mathbf{E}[\![E]\!]_\rho]\}$	
$\mathbf{A}[\![X := ?]\!]_\rho \triangleq \{\rho' \mid \exists z \in \mathbb{Z} : \rho' = \rho[X := z]\}$	
$\mathbf{A}[\![B]\!]_\rho \triangleq \{\rho' \mid \mathbf{B}[\![B]\!]_{\rho'} = true \wedge \rho' = \rho\}$	
Commands $\mathbf{C} : \Sigma \to \wp(\Sigma)$	
$\mathbf{C}(< \rho, k, L : A \to L' >) \triangleq \{< \rho', k', C' > \mid \rho' \in \mathbf{A}[\![A]\!]_\rho \wedge k' = \begin{cases} < M, L' > k \wedge lab[\![C']\!] = lab[\![entry[\![M]\!]]\!] & \text{if } A \in \mathbb{A}_c, \text{ where } A = call\ M \\ \tilde{k} \wedge lab[\![C']\!] = L_r & \text{if } A \in \mathbb{A}_r, \text{ where } k = < M, L_r > \tilde{k} \\ k \wedge lab[\![C']\!] = L' & \text{otherwise} \end{cases} \}$	

Let $\mathfrak{D}^+$ be the set of all finite state sequences (i.e. traces). The maximal finite trace semantics $\mathcal{S}^+[\![P]\!] \in \wp(\mathfrak{D}^+)$ of program P can be expressed as the least fixpoint $lfp^{\subseteq} \mathrm{F}^+[\![P]\!]$ of the monotone function $\mathrm{F}^+[\![P]\!] \in \wp(\mathfrak{D}^+) \to \wp(\mathfrak{D}^+)$ which is defined as follows

$$\mathrm{F}^+[\![P]\!](\mathcal{X}) \triangleq \mathfrak{I}[\![P]\!] \cup \{ss'\sigma \mid s' \in \mathbf{C}[\![P]\!](s) \wedge s'\sigma \in \mathcal{X}\}. \tag{3}$$

3 Modeling API Trace Birthmark as Semantic Property

In this section, let us see how API trace birthmark can be regarded as an abstraction of program semantics. As we know, software birthmarks are the inherent characteristics of a program and API trace birthmark can also be considered as a program property. Furthermore, from a semantic point of view, each program property can be specified by abstract interpretation of the program semantics. Therefore, we can model API trace birthmark as a semantic program property by abstract interpretation.

Let $\mathcal{X} \in \wp(\mathfrak{D}^+)$ be a set of non-empty traces, $\mathbb{M}_A \subseteq \mathbb{M}$ be the set of all API function names, then the abstraction α_{API} that extracts the sequences of API calls from the maximal finite trace semantics of a program is defined as follow:

$$\alpha_{API}(\mathcal{X}) \triangleq \{\alpha_{api}(\sigma) | \sigma \in \mathcal{X}\} \tag{4}$$

and for traces

$$\alpha_{api}(\varepsilon) \triangleq \varnothing, \tag{5}$$

$$\alpha_{api}(\langle \rho, k, C \rangle \sigma') \triangleq \begin{cases} M_c \cdot \alpha_{api}(\sigma') & \text{if } M_c \in \mathbb{M}_A \\ \alpha_{api}(\sigma') & \text{otherwise} \end{cases}, \tag{6}$$

where $M_c = cfunc[\![C]\!]$ is the name of the function called in command C. Given a non-empty trace σ, α_{api} discards from σ all information about the environments and stacks, only records the sequences of API calls in the trace. By observation, we have the following Galois connection:

$$po\langle \mathfrak{D}^+, \subseteq \rangle \underset{\gamma_{API}}{\overset{\alpha_{API}}{\leftrightarrows}} po\langle \wp(\mathfrak{A}^+), \subseteq \rangle \tag{7}$$

where γ_{API} is the concretization function induced by α_{API} and $\mathfrak{A}^+$ is the set of finite sequences of API calls.

Moreover, let us consider the closure operator $\varphi_{API} = \gamma_{API} \circ \alpha_{API}$ corresponding to the abstraction function α_{API}. Since the closure operator φ_{API} can be regarded as a semantic property that reflects the program's API call sequences, we can then use this closure operator φ_{API} to model API trace birthmark.

4 The Credibility of API Trace Birthmark

Credibility is an important criterion for software birthmarks. In order to use software birthmarks to identify programs in the event of suspected theft, it is necessary for software birthmarks to be credible, which requests that the software birthmarks of two independently implemented programs should be different. In this section, we discuss and analyze the credibility of API trace birthmark. First, we give a specific criterion for the copy relation. Then we prove that API trace birthmark is credible with respect to the given criterion for the copy relation.

The criteria for the copy relation can vary depending on cases. Thus, it is reasonable for us to assume that if the set of all API calls of a program is the same as that of another program, then there is a copy relation between these two programs. We denote this specific criterion for the copy relation as $\mathcal{C}_{API}$. From a semantic point of view, we can characterize this specific copy criterion $\mathcal{C}_{API}$ by the equivalence of a related program property.

Let us consider the set of all API calls in a program as a semantic property, which can be specified by the following abstraction. The abstraction $\alpha_{\mathcal{C}_{API}} : \wp(\mathfrak{D}^+) \to \wp(\mathbb{M}_A)$ abstracts the set of API calls from the maximal finite trace semantics of a program and defines a Galois connection:

$$po\langle \wp(\mathfrak{D}^+), \subseteq \rangle \underset{\gamma_{\mathcal{C}_{API}}}{\overset{\alpha_{\mathcal{C}_{API}}}{\leftrightarrows}} po\langle \wp(\mathbb{M}_A), \subseteq \rangle \tag{8}$$

where $\mathbb{M}_A \subseteq \mathbb{M}$ denotes the set of all API function names.

$$po\langle \wp(\mathfrak{D}^+), \subseteq \rangle \underset{\gamma_{\mathcal{C}_{API}}}{\overset{\alpha_{\mathcal{C}_{API}}}{\leftrightarrows}} po\langle \wp(\mathbb{M}_A), \subseteq \rangle, \tag{9}$$

and for traces

$$\alpha_{\mathcal{C}_{api}}(\varepsilon) \triangleq \varnothing \tag{10}$$

$$\alpha_{\mathcal{C}_{api}}(\langle \rho, k, C \rangle \sigma') \triangleq \begin{cases} \{M_c\} \cup \alpha_{\mathcal{C}_{api}}(\sigma') & \text{if } M_c \in \mathbb{M}_A \\ \alpha_{\mathcal{C}_{api}}(\sigma') & \text{otherwise} \end{cases}, \tag{11}$$

where $M_c = cfunc[\![C]\!]$ is the name of the function called in command C. Given a non-empty trace σ, this abstraction discards from σ all information about the environments and stacks, only observes the information about the API calls in the trace. In addition, let us consider the closure operator $\varphi_{\mathcal{C}_{API}} = \gamma_{\mathcal{C}_{API}} \circ \alpha_{\mathcal{C}_{API}}$ corresponding to the abstraction function $\alpha_{\mathcal{C}_{API}}$. Since the closure operator $\varphi_{\mathcal{C}_{API}}$ can be regarded as a semantic property that reflects the set of API calls in a program, the above criterion for the copy relation $\mathcal{C}_{API}$ can then be specified by the $\varphi_{\mathcal{C}_{API}}$ - equivalence between the semantics of two programs, which is formalized as follows

$$\forall P, Q \in \mathbb{P}: \varphi_{\mathcal{C}_{API}}(\mathcal{S}^+[\![P]\!]) = \varphi_{\mathcal{C}_{API}}(\mathcal{S}^+[\![Q]\!]) \Rightarrow P \equiv_{cp} Q \tag{12}$$

where $P \equiv_{cp} Q$ denotes that there is a copy relation between program P and Q.

Given this specific criterion for the copy relation $\mathcal{C}_{API}$, we can then have the following proposition that shows the credibility of API trace birthmark with respect to the given copy criterion $\mathcal{C}_{API}$.

***Proposition 1*:** API trace birthmark φ_{API} is credible with respect to the copy criterion $\mathcal{C}_{API}$.

***Proof*:** In order to show that φ_{API} is credible with respect to the copy criterion $\mathcal{C}_{API}$, we have to prove that if $P,Q \in \mathbb{P}$ are two independently written programs which accomplish the same task, then $\varphi_{API}\left(\mathcal{S}^+[\![P]\!]\right) = \varphi_{API}\left(\mathcal{S}^+[\![Q]\!]\right)$. As the condition implies that $P \not\equiv_{cp} Q$, we only have to prove $\varphi_{API}\left(\mathcal{S}^+[\![P]\!]\right) = \varphi_{API}\left(\mathcal{S}^+[\![Q]\!]\right) \Rightarrow P \equiv_{cp} Q$.

First, let us show that $\varphi_{API} \sqsubseteq \varphi_{\mathcal{C}_{API}}$. Since we have

$$po\left\langle \wp\left(\mathfrak{A}^+\right), \subseteq \right\rangle \underset{\gamma_r}{\overset{\alpha_r}{\rightleftarrows}} po\left\langle \wp\left(\mathbb{M}_A\right), \subseteq \right\rangle, \tag{13}$$

where $\alpha_r(\mathcal{T}) \triangleq \left\{ M_c \,\middle|\, \exists \sigma \in \mathcal{T}: \exists i \in \left[0, |\sigma|\right) : \sigma_i = M_c \right\}$, then for (7) and (13), we have

$$po\left\langle \wp\left(\mathfrak{D}^+\right), \subseteq \right\rangle \underset{\gamma_{\mathcal{C}_{API}} = \gamma_{API} \circ \gamma_r}{\overset{\alpha_{\mathcal{C}_{API}} = \alpha_r \circ \alpha_{API}}{\rightleftarrows}} po\left\langle \wp\left(\mathbb{M}_A\right), \subseteq \right\rangle, \tag{14}$$

Therefore $\varphi_{API} \sqsubseteq \varphi_{\mathcal{C}_{API}}$.

Then since $\varphi_{API}, \varphi_{\mathcal{C}_{API}} \in uco\left(\wp\left(\mathfrak{D}^+\right)\right)$ and $\varphi_{API} \sqsubseteq \varphi_{\mathcal{C}_{API}}$, we have $\varphi_{\mathcal{C}_{API}} \circ \varphi_{API} = \varphi_{\mathcal{C}_{API}}$. Moreover $\forall P,Q \in \mathbb{P}$, the following holds

$$\varphi_{API}\left(\mathcal{S}^+[\![P]\!]\right) = \varphi_{API}\left(\mathcal{S}^+[\![Q]\!]\right) \Rightarrow \varphi_{\mathcal{C}_{API}}\left(\varphi_{API}\left(\mathcal{S}^+[\![P]\!]\right)\right) = \varphi_{\mathcal{C}_{API}}\left(\varphi_{API}\left(\mathcal{S}^+[\![Q]\!]\right)\right) \Leftrightarrow \varphi_{\mathcal{C}_{API}}\left(\mathcal{S}^+[\![P]\!]\right) = \varphi_{\mathcal{C}_{API}}\left(\mathcal{S}^+[\![Q]\!]\right) \tag{15}$$

According to (12), we have

$$\varphi_{\mathcal{C}_{API}}\left(\mathcal{S}^+[\![P]\!]\right) = \varphi_{\mathcal{C}_{API}}\left(\mathcal{S}^+[\![Q]\!]\right) \Rightarrow P \equiv_{cp} Q. \tag{16}$$

For (15) and (16), we can then conclude

$$\varphi_{API}\left(\mathcal{S}^+[\![P]\!]\right) = \varphi_{API}\left(\mathcal{S}^+[\![Q]\!]\right) \Rightarrow \varphi_{\mathcal{C}_{API}}\left(\mathcal{S}^+[\![P]\!]\right) = \varphi_{\mathcal{C}_{API}}\left(\mathcal{S}^+[\![Q]\!]\right) \Rightarrow P \equiv_{cp} Q. \tag{17}$$

Thus, API trace birthmark φ_{API} is credible with respect to the copy criterion $\mathcal{C}_{API}$.

6 The Resilience of API Trace Birthmark

Since attackers may transform a program to destroy the birthmark of the program, software birthmarks have to be strong enough to endure semantics-preserving transformations. In this section, we analyze and discuss the resilience of API trace birthmark from a semantic point of view, and prove that API trace birthmark is resilient to a commonly used transformation, substitution of equivalent commands.

As semantics-preserving transformations, such as code obfuscation and code optimization, can be considered as semantic program transformations, and API trace birthmark can be modeled as a semantic property, we unify API trace birthmark and semantics-preserving transformation attacks within a semantic foundation. Moreover, given a semantics-preserving transformation attack modeled by semantic transformation $\mathcal{A} \in \wp\left(\mathfrak{D}^+\right) \to \wp\left(\mathfrak{D}^+\right)$, we say that API trace birthmark φ_{API} is resilient to $\mathcal{A}$, if the following holds:

$$\varphi_{API}\left(\mathcal{S}^+[\![P]\!]\right) = \varphi_{API}\left(\mathcal{A}\left(\mathcal{S}^+[\![P]\!]\right)\right) \tag{18}$$

In other words, if the semantic property φ_{API} is preserved by transformation $\mathcal{A}$, then we regard API trace birthmark φ_{API} as a resilient birthmark to transformation $\mathcal{A}$.

To further demonstrate our point of view, we give a concrete example, where the resilience of API trace birthmark to substitution of equivalent commands is shown. Substitution of equivalent commands is a commonly used obfuscating transformation, which replaces a single command with an equivalent one. It can be used by attackers to thwart signature-based malware detection, or to hide the event of software theft. Now let us consider this obfuscating transformation as a program transformation attack against API trace birthmark.

Let $C, C' \in \mathbb{C}$ be two commands, C and C' are equivalent if they always cause the same effect on program states, namely if the following condition holds:

$$\forall \rho \in \mathfrak{E}, k \in \mathfrak{K} : \mathbf{C}(\langle \rho, k, C \rangle) = \mathbf{C}(\langle \rho, k, C' \rangle). \tag{19}$$

Moreover, given a command $C \in \mathbb{C}$, the set of commands that are equivalent to C can be denoted as $eqCom[\![C]\!] \triangleq \{C' \in \mathbb{C} \mid \forall \rho \in \mathfrak{E}, k \in \mathfrak{K} : \mathbf{C}(\langle \rho, k, C \rangle) = \mathbf{C}(\langle \rho, k, C' \rangle)\}$. And we specify a substitution pattern by function $SubP \in \mathfrak{C} \triangleq \mathbb{C} \to \mathbb{C}$, i.e.

$$SubP \triangleq \lambda C \in \mathbb{C}.C', \text{where } C' \in eqCom[\![C]\!]. \tag{20}$$

Let $\mathcal{X} \in \wp(\mathfrak{D}^+)$ be a set of non-empty traces, $\mathbb{C}_1 \in \wp(\mathbb{C})$ be the set of commands that need to be substituted, $SubP$ be the function that specifies substitution patterns for the commands in $\mathbb{C}_1$. Then the abstraction that specifies substitution of equivalent commands is defined as

$$\alpha_S(\mathcal{X})[\mathbb{C}_1, SubP] \triangleq \{\alpha_s(\sigma)[\mathbb{C}_1, SubP] \mid \sigma \in \mathcal{X}\}, \tag{21}$$

and for traces

$$\alpha_s(\varepsilon)[\mathbb{C}_1, SubP] \triangleq \varepsilon, \tag{22}$$

This semantic transformation obscures each trace in the set $\mathcal{X}$. For a non-empty trace, it replaces the command C in state $\langle \rho, k, C \rangle$ with an equivalent one $SubP[\![C]\!]$, if command C is in the set $\mathbb{C}_1$. Observe that this semantic transformation α_S defines a Galois connection as follows:

$$po\langle \wp(\mathfrak{D}^+), \subseteq \rangle \underset{\gamma_S}{\overset{\alpha_S}{\leftrightarrows}} po\langle (\mathbb{C} \times \mathfrak{C}) \to \wp(\mathfrak{D}^+), \dot{\subseteq} \rangle. \tag{24}$$

Based on the above abstraction α_S that specifies substitution of equivalent commands, we then have the following proposition which shows the resilience of API trace birthmark to substitution of equivalent commands.

Proposition 2: API trace birthmark φ_{API} is resilient to substitution of equivalent commands α_S.

Proof: We have to show that if program Q is obtained from program P by transformation α_S, then $\varphi_{API}(S^+[\![P]\!]) = \varphi_{API}(S^+[\![Q]\!])$. According to the condition, we have $Q = \mathcal{P}^+(\alpha_S(S^+[\![P]\!]))$, namely

$$S^+[\![Q]\!] = \alpha_S(S^+[\![P]\!]) \quad (25)$$

Hence, we have to prove

$$\varphi_{API}(S^+[\![P]\!]) = \varphi_{API}(\alpha_S(S^+[\![P]\!])) \quad (26)$$

Following the definition of the semantic transformation α_S, we have that for every $\mathbb{C}' \in \wp(\mathbb{C})$, $SubP \in \mathfrak{C}$, $P \in \mathbb{P}$, and $\sigma \in S^+[\![P]\!]$, such that $\sigma = \langle \rho_1, k_1, C_1 \rangle \ldots \langle \rho_n, k_n, C_n \rangle$, there exists $\delta \in \alpha_S(S^+[\![P]\!])[\mathbb{C}', SubP]$, such that $\delta = \langle \rho_1, k_1, C_1' \rangle \ldots \langle \rho_n, k_n, C_n' \rangle$, where $\mathbf{C}(\langle \rho_i, k_i, C_i \rangle) = \mathbf{C}(\langle \rho_i, k_i, C_i' \rangle)$. Since a API call command can not be substituted by a single command and the stacks of the trace σ are unchanged, we have $\alpha_{api}(\sigma) = \alpha_{api}(\delta)$. Hence, we can conclude that semantic property φ_{API} is preserved by the semantic transformation α_S, i.e. $\varphi_{API}(S^+[\![P]\!]) = \varphi_{API}(\alpha_S(S^+[\![P]\!]))$.

7 Conclusion

We have specified API trace birthmark based on abstract interpretation to formally analyze and certify the credibility and the resilience of API trace birthmark. The proposed method has the following characteristics:

- It unifies the criteria for the copy relation, program transformation attacks and API trace birthmark within solid semantic foundations.
- API trace birthmark is considered as a semantic program property, which can be specified by abstract interpretation of program semantics. Moreover, the credibility and resilience of API trace birthmark are formally analyzed and proved.
- Although illustrated on a low level imperative language, the method is language independent.

Nevertheless, some future work is still necessary in order to validate our method in practice. To make birthmarks less sensitive to program transformation attacks, similarity is usually introduced in the comparison of birthmarks. Thus, one interesting field for future research is how to extend our semantic method for API trace birthmark with similarity, so that makes it more practical.

Acknowledgements. This work was supported in part by grants from the National Natural Science Foundation of China under Grant Nos. 60970141 and 60902102, and the Fund of Innovation Scientists and Technicians Outstanding Talents of Henan Province under Grant No. 094200510008.

References

[1] Collberg, C., Nagra, J.: Surreptitious software: obfuscation, watermarking and tamperproofing for software protection, 1st edn. Addision Wesley, Boston (2009)
[2] Tamada, H., Nakamura, M., Monden, A., Matsumoto, K.: Detecting the theft of programs using birthmarks. Graduate School of Information Science, Nara Institute of Science and Technology. Technical Report NAIST-IS-TR2003014 (2003)
[3] Tamada, H., Nakamura, M., Monden, A., Matsumoto, K.: Java birthmark-detecting the software theft. IEICE Trans. Inf. & Syst. E88-D(9), 2148–2158 (2005)
[4] Park, H., Choi, S., Lim, H., Han, T.: Detecting Java theft based on static API trace birthmark. In: International Workshop on Security, pp. 121–135 (2008)
[5] Choi, S., Park, H., Lim, H., Han, T.: A static API birthmark for Windows binary executables. J. Syst. & Softw. 82, 862–873 (2009)
[6] Tamada, H., Nakamura, M., Monden, A., Matsumoto, K.: Design and evaluation of birthmarks for detecting theft of Java program. In: IASTED International Conference Software Engineering, pp. 569–575 (2004)
[7] Schuler, D., Dallmeier, V., Lindig, C.: A dynamic birthmark for Java. In: 22nd IEEE/ACM International Conference on Automated Software Engineering, pp. 274–283 (2007)
[8] Cousot, P., Cousot, R.: Abstract interpretation: A unified Lattice model for static analysis of programs by construction or approximation of fixpoints. In: 4th ACM SIGPLAN-SIGACT Symposium Principles of Programming Languages, pp. 238–252 (1977)
[9] Cousot, P., Cousot, R.: Systematic design of program transformation frameworks by abstract interpretation. In: 29th ACM SIGPLANSIGACT Symposium Principles of Programming Languages, pp. 178–190 (2002)

Extracting Semantic Information from Chinese Language Patent Claims

Yong Tang[1], Shihan Yang[2], Jiwen Chai[1], and Shanmei Liu[1]

[1] Sichuan Electric Power Research Institute, Chengdu, China
[2] Chengdu Documentation and Information Center, CAS, Chengdu, China
dr.yangsh@gmail.com

Abstract. A big challenge for automatically analyzing patent claims written in Chinese language is how to obtain semantically information from claims which are usually a non-structured, but free text. In this paper, a set of techniques is been provided to extract some valuable semantically information from claims in Chinese language. The method could automatically discovery some usable semantically information from the patent claim texts by means of regular expression pattern and predefined ontology model. It can extract not only surface semantic information but also deeper semantic information. Furthermore, the extracted semantically information is automatically translated into web ontology language (OWL), a machine readable semantic structure specification. The work proposes a potential semantic solution in Chinese language patent analyzing based on patent claims, such as semantic searching, legal status and invalidation checking, and discovering new technique trends. A case study on electronic engineering domain patent claims is also provided.

Keywords: Semantic Information Extraction, Patent Claim; Chinese Language; Patent Automatically Analyzing; Ontology.

1 Introduction

A patent is a set of exclusive rights granted by a national government to an inventor or their assignee for a limited period of time in exchange for a public disclosure of an invention. With the size of patents rapidly growing, and with the written style of them being free, it is a big challenge to conduct analysis, comparison, classification and retrieval of patents at semantically level. Patent legal status checking, patent examination and patent invalidation checking need to analyze claims semantically. These usually can be done by domain experts retrieving and reading all potential relevant patents. It is very tedious and costly. Our purpose of this work is to develop an automatic technique to extract some valuable semantic information from Chinese language patent claim and facilitate the semantically analyzing of that.

FASTUS [1] is a system using cascaded finite state automata to extract some semantic relations from free texts. Three steps are included: (a) phrase recognition; (b) pattern recognition in a phrase; and (c) merge the information found. But it can just find "desired" information, some patterns defining the requirement. On the other words, it can not "discovery" implicit semantic information.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 547–556.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

[2] Proposes a set of techniques to extract the semantic relations from patent claims in English. A domain specific regular expression is employed. The method can find some explicit relations among patent claims, but can not find any implicit relations. Furthermore, English language is very different with Chinese language, so the method can not be facilitated to analyzing of Chinese patent claims.

[3] Introduces regular expressions to extract information from Chinese patent, but it obtains information based on the structure of the whole patent document. It does not consider the meaning of the free text in the patent document, i.e., it is not involved in Chinese text information processing. [4] Employs text-mining techniques to analyze Chinese patents, in order to qualify the novelty degree of a patent. It is also involved in chew on patent claims and abstract, but it just evaluates words (concepts or terms) mined from there. It does not extract and evaluate relations hidden in claims. Our method focus on analyzing patent claims in Chinese, furthermore on extracting semantic relations hidden in that.

2 Chinese Word Segmentation

In this paper, we substitute ICTCLAS [5], the first and the best international Chinese word segmentation tool, for challenging the complex problem of Chinese word segmentation. The lastly version is ICTCLAS2011 [6]. ICTCLAS is a competitive Chinese lexical analyzer, and its frame is the unified HHMM-based (Hierarchical Hidden Markov Model [7] based).

ICTCLAS provides an interface to configure user-defined dictionary. The tool could do better for the special domain word segmentation by means of defining some proper terms in user dictionary. The following is an example, the free Chinese text coming from the claim of Chinese patent No. CN1046065.

Example 1: yi zhong yan shi zi dong guan bi la xian kai guan, you xian la qi gang ti, gui wei tan huang, gu ding huo sai, dan xiang mi feng quan, dong chu dian, jing chu dian, qi guang qian la heng gan, chang kai ka feng deng gou cheng. (The delayed auto power-off electronic pull switch consists of pull cylinder block, return spring, stationary piston, one-way seal ring, traveling contact, stationary contact, cylinder draw-off crossbar and normal open slit.)

The following is result of Chinese word segmentation without user-defined dictionary. The result is trivial, and does not express the precise meaning of fundamental words. The letter(s) behind slash is the mark of part of speech, or user-defined mark in user dictionary. The result shows the segmentation as general morphemes, which can not correctly express the specific concept in the text. For example, the concept cylinder draw-off crossbar is split into cylinder / noun, draw / verb, pull / verb, crossbar / noun, but not treated as a single domain specific noun, so do normal open slit, traveling contact, stationary contact, etc.

yi/m zhong/q yan/vg shi/ng zidong/d guanbi/v laxian/n kaiguan/n , /wd you/p xian/n la/v qiguang/n ti/ng ,/wn guiwei/vi tanhuang/n ,/wn guding/a huosai/n ,/wn danxiang/b mifengquan/n ,/wn dong/v chu/v dian/qt ,/wn jing/ad chu/v dian/qt ,/wn qigang/n qian/v la/v hengguan/n ,/wn chang/d kai/v ka/n feng/n deng/udeng goucheng/n ./wj (one/m kind/q /over/vg time/ng automatic/d close/v pull/n switch/n comma/wd by/p line/n

pull/n cylinder/n body/ng comma/wn return/vi spring/n comma/wn stationary/a pistion/n comma/wn one-way/b seal ring/n comma/wn moving/v touch/v point/qt comma/wn stationary/ad touch/v point/qt comma/wn cylinder/n draw/v pull/v crossbar/n comma/wn normal/d open/v card/n slit/n etc./udeng consist/n period/wj)

A user dictionary like following is defined. One item locates one line, and each item is a pair, which consists of a word and a mark, separated by a blank. Explanations are included in the following parentheses.

yizhong (a kind of) mq (numeral and quantifier)
yanshi (delayed) a (adjective)
zidongguanbi (auto power-off) d (adverb)
xianlaqigangti (pull cylinder block) n (noun)
dongchudian (traveling contact) n
jingchudian (stationary contact) n
qianlahenggan (draw-off crossbar) n
changkai (normal open) d
kafeng (slit) n

With auxiliary of user defined dictionary, the tool gives the result as following, which is closer to the real meaning of the text content than before. And the segmentation provides better lexical expression for the sequential semantically analysis.

Yizhong/mq yanshi/a zidongguanbi/d laxian/n kaiguan/n ,/wd you/p xianlaqigangti/n ,/wn guiwei/vi tanhuang/n ,/wn guding/a huosai/n ,/wn danxiang/b mifengquan/n ,/wn dongchudian/n ,/wn jingchudian/n ,/wn qigang/n qianlahenggan/n ,/wn changkai/d kafeng/n deng/udeng goucheng/n ./wj (a kind of/mq delayed/a auto power-off/d pull/n switch/n comma/wd by/p pull cylinder block/n comma/wn return/vi spring/n comma/wn stationary/a piston/n comma/wn one-way/b seal ring/n comma/wn traveling contact/n comma/wn stationary contact/n comma/wn cylinder/n draw-off crossbar/n comma/wn normal open/d slit/n)

In the work, Chinese patent claims semantically analyzing, we always provide user defined dictionary for the special domain. The preparation work provided by domain experts or came from sharing domain ontology is really needed, because the text of the patent claims might contain a lot of domain specific terminology.

3 Semantically Pattern

We employ two semantically extracting patterns to discovery semantically information after Chinese patent claims lexical analyzed with the ICTCLAS tool. Firstly, regular expression pattern can find the explicit semantic information, which has precise semantically structure, such as subject-predicateobject structure, adjective-noun structure. Moreover, desired semantically pattern can also be made with regular expression. Secondly, ontology schema pattern can the find implicit semantic information, which can not be obtained directly from visual text, but which can be reasoned out. For example, from the free text “wo jiao li qiang, wo ba ba shi li gang

(My name is LiQiang, and my father's name is LiGang.)", we can discovery a "fuzi" (father-son) relationship between "ligang" and "liqiang", which can not be found by regular expression. The implicit semantic information like this can be discovered by ontology and reasoning.

3.1 Regular Expression Pattern

For getting useful explicit semantic information, we utilize domain specific regular expressions. Java regular expression specification (a part of JSR 51 [8]) is adopted in this paper, and the part of speech tags in ICTCLAS [6] is also included in the work. Five kinds of regular expressions have been identified to extract semantic information in Chinese patent claims.

- Type 0 : Claim type

Patent claims are the part of a patent or patent application that defines the scope of protection granted by the patent. The claims define, in technical terms, the extent of the protection conferred by a patent, or the protection sought in a patent application. The claims are of the utmost importance both during prosecution and litigation. There are two basic types of claims: independent and dependent claims. Type 0 regular expression pattern can distinguish them.

The independent claims stand on themselves and do not depend on other claims. Example 2 shows an independent claim from the Chinese language patent CN2852371.

*Example 2 (*independent *claim)*: 1.yizhong zhineng fanghuo kaiguan, ta baokuo duanluqi, tuokou zhuangzhi, dianliu huganqi, qi tezheng zaiyu:…(Claim 1. A kind of intelligent-controlled fireproofing switch consists of disconnector, release device, and current transformer. It features: …)

The dependent claims refer back to another claim and generally express particular embodiments as fall-back positions. They incorporate by reference to prior claims. Example 3 and 4 show two dependent claims in Chinese patent CN2852371.

*Example 3 (*dependent *claim)*: 2. **genju quanli yaoqiu 1 suosu** de zhineng fanghuo kaiguan, **qi tezheng zaiyu**: hai baokuo guangdian chuanganqi, suosu de shuju chuliji (7) de shuru he guangdian chuanganqi de shuchu lianjie.(Claim 2. According to claim 1 with intelligent-controlled fireproofing switch, characterized in that: …)

Example *4 (*dependent *claim)*: 3. **genju quanli yaoqiu 1 suosu** de zhineng fanghuo kaiguan, **qi tezheng zaiyu**: hai baokuo wendu chuanganqi, suosu de shuju chuliji (7) de shuru he wendu chuanganqi de shuchu xiang lianjie. (Claim 3. According to claim 1 with intelligent-controlled fireproofing switch, characterized in that:…)

Some domain specific regular expression for claim types are as follows, where /n is a noun mark, /m a numeral, and /w a Chinese punctuation. Other symbols are standard marks from Java regular expression specification.

regClaimType_Independent = *"yizhong"/n\S* "qi tezheng zaiyu"\n/S*| ("n (d+))"\n/S*;*

*regClaimType _Dependent =/m(/w?)"*genju *quanli yaoqiu" \(d+) "suosu"\S*|\s*;*

- Type 1 : Components

The type of regular expressions is to extract components in the claims. The invention components (concepts) are the most important element in an invention. They are similar to general concept domain term, because applicants wish to have more right protection. Therefore they should be extracted by more flexible expression.

A domain specific regular expression for components of Chinese patent claims is as follows.

```
term = "DT";
regComponents = \S* (/n+ |term+)\nS*;
```

In the regular expressions above, DT is some one domain special term, included in user defined dictionary. Example 5 shows components extraction from Chinese patent CN2852371 in Example 2, where bold-face words represent the extracted information, (here are concepts of electronic domain) similarly hereinafter.

Example 5 (components): 1. yizhong zhineng fanghuo kaiguan, ta baokuo duanluqi, ta baokuo duanluqi, tuokouzhuangzhi, dianliu huganqi, qi tezheng zaiyu: hai baokuo kaiguan dianyuan(10), dianya bianhuan danyuan(1), baojing danyuan(9), tiaojie danyuan(8), qianzhi danyuan(3), shujuchuliji(7), zhixing danyuan(6), dianya bianhuan danyuan(1) de shuchu tongguo qianzhi danyuan(3) yu shujuchuliji(7) shuru xiang lianjie, shujuchuliji(7) de shuchu fenbie yu zhixing danyuan(6), baojing danyuan(9) de shuru xiang lianjie, shujuchuliji(7) yu tiaojie danyuan(8) xiang lianjie, zhixing danyuan(6) de shuchu yu tuokou zhuangzhi(5) xiang lianjie, dianliu huganqi(2) tongguo qianzhi chuli danyuan(3) yu shujuchuliji(7) de shuru xiang lianjie. (switch, disconnector, release device, current transformer, power supply, preposition unit, voltage change unit, warning unit, accommodation unit, data processor, executing unit etc.)

- Type 2 : Attribute

The type of regular expressions can obtain the attribute (has-a and is) relationship. Following shows three regular expression for finding attributes in Chinese patent claims, where /a is adjective mark.

```
regAttribute_1 = /a+ \S*/n;
regAttribute_2 = /n "shi"\S* (/n+|/a+);
regAttribute_3 = /n "you"\S*/n+;
```

Example 6 shows result of the patent CN2852371 that "kaiguan de shuxing shi fanghuo de, shi zhineng de" (the switch is fireproof and intelligent controlled).

Example 6: 1. yizhong **zhineng fanghuo** kaiguan, … (a kind of fireproof and intelligent switch, …)

- Type 3 : Containment relationship

The type of regular expressions can extract the containment (part-of) relationship between components. Regular expressions belong to this type like the following, and its result is shown in Example 7.

regContainment = (/n\S "baokuo"(/n/w+)*)|(/n\S* "you"nS* "zucheng"| "goucheng"/w+);*

Example 7: 1. yizhong zhineng fanghuo kaiguan, ta baokuo duanluqi, ta baokuo duanluqi, tuokouzhuangzhi, dianliu huganqi, qi tezheng zaiyu: hai baokuo kaiguan dianyuan(10), dianya bianhuan danyuan(1), baojing danyuan(9), tiaojie danyuan(8), qianzhi danyuan(3), shujuchuliji(7), zhixing danyuan(6),… (Includes / consists of)

- Type 4 : Spatial relationship

The type of regular expressions is to be used to extract spatial relationships among components from a claim. The following is a sample regular expression belongs to the type, and Example 8 tells the result.

regSpatial = (/n + ["tongguo"]\S "yu"/n + "xiang lianjie"| "jiehe") | (/n + "zai"| "chuyu"\S*/n + "shang"| "zhong"| "xia"| "nei"| "wai");*

Example 8: … dianya bianhuan danyuan(1) de shuchu tongguo qianzhi danyuan(3) yu shujuchuliji(7) shuru xiang lianjie, shujuchuliji(7) de shuchu fenbie yu zhixing danyuan(6), baojing danyuan(9) de shuru xiang lianjie, shujuchuliji(7) yu tiaojie danyuan(8) xiang lianjie, zhixing danyuan(6) de shuchu yu tuokou zhuangzhi(5) xiang lianjie, dianliu huganqi(2) tongguo qianzhi chuli danyuan(3) yu shujuchuliji(7) de shuru xiang lianjie. (voltage change unit is connected with data *processor* by preposition unit, data processor is connected with inputting interfaces of executing unit and warning data, executing unit's outputting is connected with release device, and current transformer is connected with input of data processor also by preposition unit.)

4 Ontology Schema

For discovering valuable implicit semantic information, domain specific ontology has been employed. With reference to the domain specific ontology, we can get other semantic information, which can't be extracted only by regular expression pattern. The main idea is that we apply concepts and facts found by regular expression pattern into the reasoning on domain specific ontology to discovery implicit semantic information of patent claims. For the purpose, three types of semantic expansion operator are defined as follows:

*Definition 1 (*Inheritance *expansion)*: If concept *A* is a child of concept *B (A* $\subseteq$ *B)*, and *x* is an individual of *A (x* $\in$ *A)*, then *x* is also an individual of *B (x* $\in$ *B)*.

In the definition 1, *x* is the fact that been found by regular expression pattern. By the inheritance expansion operator, new concepts from the ontology and new assertions (relations) could be introduced into extracted semantic result. This could increase accuracy and completeness about semantics of patent claims.

*Definition 2 (*Attribute *expansion)*: If concept *C* has *n* attributes, where an attribute a_1 with weight $w_1 = 0.6$, an attribute a_2 with weight $w_2 = 0.2$, an attribute a_3 with weight $w_3 = 0.1$, … , etc., and where the sum of all weights is 1,and *x* has attributes a_i, … , a_j ,*(i, j*$\in$ *[1, n])*, then *x* is an individual of *C* with weight $w_i + w_{i+1} + \ldots + w_j$.

In fact, definition 2 provides a method to decide how much an extracted fact or concept is associated with another concept among domain specific ontology, and the weight of attributes to its concept is used. The weight of an attribute expresses how much it can distinguish its owner with other concepts, a value between *0* and *1* given by experts or computed by machine learning.

*Definition 3 (*Combination *expansion)*: If concept *C* consists of components $C_1, \ldots, C_n$, and *X* is a child of concept C_i, $i \in [1, n]$ (i.e. $X \subseteq C_i$), then *X* is part of *C*.

Combination expansion can introduce new *part-of* relations into the extracted semantic result. All three expansion operators can be compounded.

5 Semantic Information Normalization

The extracted semantic information should be shared as knowledge base, so normalization of semantic information is needed. In the paper, OWL [9] is adopted, i.e., the extracted information is expressed as OWL, which is the most popular standard in the semantic web world.

Extracted and expanded semantic information should be translated into OWL specification for knowledge sharing. Fortunately, the task is not too tough. The result of analyzing semantically Chinese patent claims is represented as a Wgraph [10], which is a simple formal semantic graph, a directed graph. [10] Developed a WGraph editing tool that can automatically translate WGraph files into OWL files effectively. In this research, we use the tool to translate extracted and expanded semantic information into OWL.

6 A Case Study

The process of extracting semantic information from Chinese language patent claims is shown in figure 1.

There are five main steps in the process:

1) Chinese word segmentation. It is difficult to split Chinese sentence up to semantic words, because there is no blank among words like English. The ICTCLAS2011[6] does this job better in general Chinese free text than that in domain specific free text, due to lots of technical terms. So user defined dictionary for analyzing Chinese patent claims is needed. Fortunately, ICTCLAS provides the interface to define user special dictionary for segmentation. In this step, ICTCLAS gets Chinese patent claims free text and user defined dictionary as its inputs, and its outputs is Chinese text annotated by predefined or user-defined lexical marks, just like the example in Section 2.

2) Parsing regular expression. With developed domain specific regular expression pattern, the parser can extract some required semantic information from the annotated patent claim, the output of Chinese word segmentation. The extracted information is found on command, and is usually some facts and terms occurred in the claim text in the view of ontology.

3) Reasoning on the domain specific ontology. To discovery deep semantic information in the claim text, ontology reasoner is a good choice. With the domain specific ontology and facts and terms found from the claim, the reasoner can get more semantic information, such as synonymy relations, inheritance relations, and containment relations.

4) Semantic checking and merging. A weak checking for semantic correctness and repetitiveness is done in this step by analyst manually or machine automatically. The result of semantic information is visualized as directed graph by means of WGraph editing tool [10], and it is easy to be adjusted.

5) Translating into OWL. In the step, the semantic information represented by WGraph is automatically translated into OWL file, which then be syntactically and partsemanticallyly validated by OWL validator. Finally the semantics information of the Chinese patent claim is put into Chinese patent knowledge base for public sharing.

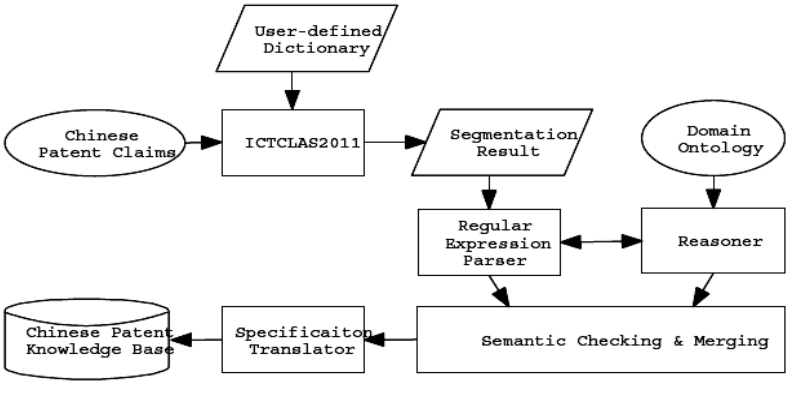


Fig. 1. The process of extracting semantic information

Figure 2 shows the result of extracted semantic information from Chinese patent CN2852371 as WGraph.

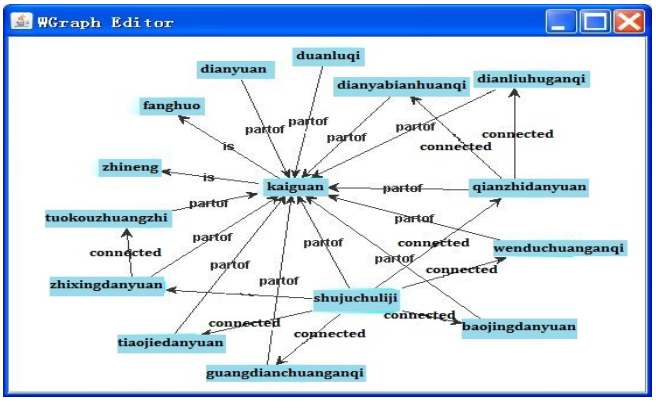


Fig. 2. Visulization of extracted semantic information without expanded

And Figure 3 shows the result after merged and expanded of extracted semantic information, which is the last result of analyzing the patent CN2852371, represented as WGraph file. Some concepts and relations are increased, and shown in the figure with red circles.

The extracted semantic information are confirmed in a sense by electronic domain experts. Then the extracted and expanded semantic information are translated into OWL file by means of the WGraph editing tool. The created OWL file is validated by the online OWL Validator[1].

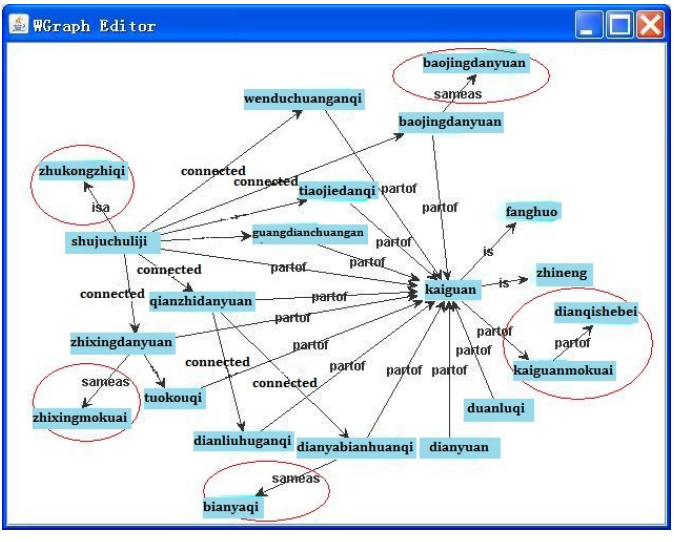


Fig. 3. Visulization of extracted semantic information with domain specific ontology expanded

7 Conclusion and Future works

A set of techniques is proposed to facilitate Chinese patent claims analyzing semantically. The method can not only find surface semantic information by means of domain specific regular expression pattern, in which we introduce 5 types regular expressions, but also discovery deeper semantic information by means of domain specific ontology pattern, in which we define 3 types semantic expansion operators. The result of extracted and expanded semantic information from claims is visually represented as a simple semantic graph (WGraph), which can be effectively translated into ontology language specification (OWL) for public knowledge sharing. An example has also been studied.

Concurrently, the regular expression patterns were devised manually based only on Chinese patent claims in electronic engineering domain, so were the ontology patterns. For other domains, the new regular expressions and ontology should be developed, and this is very tough work for claim analysts. In the near future, some machine learning

[1] http://owl.cs.manchester.ac.uk/validator/

approaches should be employed to building regular expression patterns and domain ontology patterns. Moreover, a better integration tool for analyzing Chinese patent claims will be developed. Based on extracted semantic information from claims, the researches on semantically searching and legal status invalidation checking is also our interest.

Acknowledgements. This research is partially supported by the West Light Foundation of the Chinese Academy of Sciences on the project: Automatically extracting semantically meta-data of digital document resources on science and technology, and by the Talent Training Plan of Chinese Academy of Sciences on the project: Research on patent claims searching and analyzing tools based on formal semantically reasoning.

References

[1] Hobbs, J.R., Appelt, D.E., Bear, J., Israel, D.J., Kameyama, M., Stickel, M.E., Tyson, M.: Fastus: A cascaded finite-state transducer for extracting information from natural-language text. CoRR, vol. cmplg/ 9705013 (1997)

[2] Yang, S.-Y., Lin, S.-Y., Lin, S.-N., Lee, C.-F., Cheng, S.-L., Soo, V.-W.: Automatic extraction of semantic relations from patent claims (2008)

[3] Qingying, Q., Guomin, Z., Pei'en, F., Jianwei, W.: Extraction approach of patent information based on regular expression. Journal of Chinese Machine Engineering 18(19), 2326–2329 (2007) (in Chinese)

[4] Yu-qin, L., Xue-feng, W., Xiao-ping, L.: Quality estimation of patent based on text mining and its empirical research. Computer Engineering and Applications 43(33), 12–14 (2007) (in Chinese)

[5] Zhang, H.-P., Yu, H.-K., Xiong, D.-Y., Liu, Q.: Hhmm-based chinese lexical analyzer ictclas. In: Proceedings of the Second SIGHAN Workshop on Chinese Language Processing, SIGHAN 2003, vol. 17, pp. 184–187. Association for Computational Linguistics, Stroudsburg (2003)

[6] Zhang, H.P.: Ictclas (version 2011), `http://cid-51de2738d3ea0fdd.office.live.com/self.aspx/.Public/ictclas2011/ICTCLAS2011-SDK-release.rar` (2011)

[7] Fine, S., Singer, Y., Tishby, N.: The hierarchical hidden markov model: Analysis and applications. Mach. Learn. 32, 41–62 (1998), `http://portal.acm.org/citation.cfm?id=325865.325879`

[8] Reinhold, M., et al.: Jsr 51: New i/o apis for the java platform (2002), `http://www.jcp.org/en/jsr/detail?id=51`

[9] Motik, B., Grau, B.C., Horrocks, I., Wu, Z., Fokoue, A., Lutz, C.: Owl2 web ontology language profiles, `http://www.w3.org/TR/2009/CRowl2-profiles-20090611/` (2009)

[10] Yang, S., Wu, J.: Mapping relational databases into ontologies through a graph-based formal model. In: International Conference on Semantics, Knowledge and Grid, pp. 219–226 (2010)

Sensitivity Analysis on the Influence Factors of Software Reliability Based on Diagnosis Reasoning

Li Qiuying, Li Haifeng, and Wang Guodong

School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China
li_qiuying@buaa.edu.cn

Abstract. There are some uncertain factors (i.e. environment factors) in the software development process which have direct or indirect influence on software reliability. This paper uses the Bayesian network to construct the software reliability qualitative evaluation topology structure based on the environment factors and analyzes the sensitivity of these influence factors by diagnosis reasoning of the Bayesian network for determining the environment factors which have important influence on the improvement of software reliability. The results of the sensitivity analysis may propose the decision reference for improving the resource allocation and the level of software reliability under the condition of resource restriction.

Keywords: diagnosis reasoning, software reliability, influence factor, sensitivity analysis.

1 Introduction

There are some environment factors which have certain influences on software reliability in the software reliability engineering [1]. Thus the issue how to analyze the influence of these factors and balance the relationship among these factors under the condition of resource restriction should be considered by the project manager. The sensitivity analysis is an approach for uncertain analysis which is usually used in the project evaluation and management decision. This approach analyzes the influence of the change of one or more factors on the object under the condition that the other factors don't change for deciding the importance of the change influences of these factors on the object.

Because software is very complex, the factors which have influences on software reliability are diverse.

The approach that how to analyze the importance of these influence factors for improving the reliability level by improving these important influence factors are significant for the software reliability research and engineering under the condition of resource restriction.

However, in practice, it is very difficult to describe and measure these various influence factors of software reliability [2]. Moreover, the measurement results may be biased and incorrect due to the limitations of the measurement approach and technology. Thus it is more suitable that software reliability prediction should be taken as an uncertain reasoning issue. Bayesian network which is also called as the belief

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 557–566.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

network [3-6] is suitable for the uncertain and probability things and applied into the conditional decision which is dependent of many controlling factors. By using these above research results, this paper aims to construct the Bayesian network based on the environment factors and use the character of diagnosis reasoning of the Bayesian network to analyze the sensitivity of these software reliability influence factors.

2 Analyzing Software Reliability Influence Factors

2.1 The Definition of Environment Factor

Software development process is a very complex and dynamic process including many environment factors, such as human, development tools and application background, which may have some influences on software reliability.

Environment factors can be defined as [7]: the set of the various characters and behaviors existing in the software development, testing and maintenance process which have certain influences on the software's defects and failures.

These environment factors are the basis to study software reliability and have the close relationship with the software development process. Furthermore, the information of these factors is easy to collect in the development process. Thus it solves the disadvantage that traditional quantitative approaches only focus on the failure information and neglect the qualitative information of these influence factors in the software development process.

2.2 Common Environment Factors

Reference [8] made an investigation on the factors which have influence on software reliability during the software development process of 13 organizations, such as AT&T, BellCore and Chrysler. As a result of this investigation, [8] proposed 32 main influence factors and graded these factors according to the influence level of these factors on software reliability. Based on the results of [8-10], this paper presents 30 commonly used environment factors, redefines and classifies these factors as the research objects of this paper which are shown in Table 1.

2.3 Classification of Environment Factors

Software reliability is mainly decided by the software development, testing and software product. Thus the 30 environment factors are classified into three type as shown in Table 1, namely product factors, resource factors and process factors.

- Product factors are the characters or properties relevant to the software size, document and structure. Product factors are the static index and can be collected from the materials such as design documents.
- Resource factors can be divided into three types: 'human', the reusable components and the soft-hard environment.
- Process factors are those characters and behaviors in the development, testing and maintenance processes.

Table 1. Commonly used environment factors

No.	Environment Factors	No.	Environment Factors	No.	Environment Factors
1	Software complexity	11	Testing environment	21	Reliability design method and technology
2	Reusable codes (modules)	12	Technology level and general quality of the testers	22	Development design method and technology
3	User quality object	13	Hardware resource allocation	23	The operation condition
4	Software programming language	14	Coding and testing cost	24	Software project management
5	The characters of failure and fault	15	Testing fault-removal and fault-detection	25	The construction of the documents and the evaluation criteria
6	Technology level and general quality of the developers	16	Testing coverage and adequacy	26	Software requirement analysis and detailed design quality
7	The level and size of the coding and testing teams	17	Testing case	27	Software volatility
8	Software development environment	18	Testing method	28	Development programming difficulty
9	The general quality of user	19	Testing workload	29	The schedule
10	Testing tool	20	Collecting data information	30	Programming workload

Notes:1~5 belong to the product factors, 6~14 belong to the resource factors,15~30 belong to the process factors.

3 Constructing the Topology Structure of Bayesian Network Based on Environment Factors

3.1 Bayesian Network

Bayesian network is an organic model for the expression and reasoning of uncertain information which combines the Bayesian probability method and the network topology structure of the directed acyclic graph. It is a powerful tool for the data analysis and prediction and is widely used for describing the uncertain information which can describe the data items and their relationships, and construct the graph according to the probability relationships among the various variables. Thus the Bayesian network can combine the expert experience, history data and various incomplete or uncertain information for improving the modeling efficiency and credibility.

3.2 Principle of Using Bayesian Network for Software Reliability Evaluation

The basic idea using Bayesian network for software reliability evaluation is that firstly describing the environment factors which are the most relevant with reliability to construct the Bayesian network topology structure according to the casual dependent relationships of these factors, then determining the conditional probability distribution of the node and its father nodes, based on the above analysis results, constructing the Bayesian reasoning model to yield the final software reliability evaluation results by the reasoning results of this reasoning model.

3.3 Constructing Bayesian Network Topology Structure of Software Reliability Evaluation

Defining nodes of Bayesian network

- The source of the network nodes

We select the environment factors as the input nodes of the Bayesian network model.

- The selection method of the environment factors

This paper uses the following selection criteria: A) Relevance (to reliability): this criterion reflects the relationship between metrics and software reliability; B) Experience: this criterion reflects the degree to which this metric has been used and recognized; C) Correctness: this criterion includes: i) Objectivity. The input and results of this metric can't be easily influenced; ii) Justness. The metric is not partial to any specific result; iii) Precision; D) Practicality: this metric should be concerned and required in development; E) Feasibility: this criterion means that i) the formula of this metric should be understood easily supported by tools; ii) data collection should be easy; iii) the results of this metric can be evaluated and confirmed conveniently.\

- Defining states of the nodes of Bayesian network

Here we divide the space state of the selection metrics (i.e. the input nodes) into three degrees, namely low, middle and high, which are corresponding to the values (0-1/3, 1/3-2/3, 2/3-1) of these metrics. Similarly, the space states of the other nodes are also defined as these three degrees.

- Constructing the Bayesian topology structure (i.e. defining the relationship of the nodes)

The construction approach for the Bayesian topology structure can be divided into the following two steps:

- Constructing the topology structure of the Bayesian network according to the causal relationships of the selection nodes which are analyzed by the experiential knowledge;
- Studying the topology structure of the Bayesian network based on the sample data. It can be seen that the construction of the topology structure of the Bayesian network should consider both the sample data and the prior knowledge. When the sample size is small, the prior knowledge is leading. With the growth of the sample size, the constructing process is more dependent with the real data. It is not only consistent with the general cognitive law, but also can reflect the influence of these real data on the decision of the mutual relationship more adequately.

3.4 Software Reliability Evaluation Based on Bayesian Network Topology Structure

Set the selection node set is: X={x1, x2, x3, x4, x5}, and the space state of each node is {low, med, high}. Then the Bayesian network topology structure can be obtained by this node set as shown as the following figure.

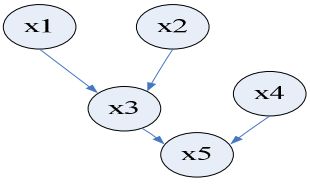


Fig. 1. A simple Bayesian network topology structure

- Determining the probability distribution of the Bayesian network nodes
- The probability distribution of the nodes is used to describe the influence degree of the reasons on the results, and can be divided into the marginal probability distribution and the conditional probability distribution. With the sufficient data, the probability distribution can be obtained directly by the maximum likelihood evaluation. With the insufficient data, the probability distribution can be determined by the expert experience as well as the study process. In this paper, we only introduce how to use the expert experience to evaluate the required probability parameters.
- Determining the marginal probability distribution. As shown in Figure 1, we call the node without the father nodes as the leaf node. Thus, x1, x2, x4 are the leaf nodes. The marginal probability distribution of the leaf nodes should be determined first. For example, the three states of the x1 are (low,med,high), then the corresponding probabilities of these three states determined by the expert experience are respectively 0.2, 0.5 and 0.3. Then we can conclude that the probability that the level of x1 is middle is higher. According to the various development teams, projects and the development environments, the probability distribution of the leaf nodes is also diverse.
- Determining the conditional probability distribution. As shown in Figure 1, considering the dependent relationships among x1, x2 and x3, if the probability that x1 and x2 are low is higher, then the probability that x3 is low is also higher. For expressing the conditional probability distribution of the nodes more conveniently, we present the conditional probability table as shown in Table 2. W= {w1, w2…wa} means a space status of the node x1, and A={a1,a2,…ak} means the conditional probabilities of the node D within the uniform distribution of its father nodes. Each group of the father nodes is corresponding to a group of probabilities and the sum is 1.
- The causal reasoning based on the Bayesian network

Table 2. The general condition of the conditional probability table

Father nodes	Leaf nodes of X1-Xn				
X1-Xn	S1	S2	……	Sk	∑
(w1,p1,…t1)	a1	a2	……	ak	1
……	……	……	……	……	1
(wa,pb,…tc)	……	……	……	……	1

The commonly used reasoning methods conclude the causal reasoning, the fault reasoning and the support reasoning. Due to the limited space, here we only introduce the main idea of the causal reasoning: Utilizing the Bayesian network structure to allocate a processor to each node. When the Bayesian network receives the evidences, the probability value of the evidence node will be changed. The processor of this node will transmit this change to its neighbor nodes. The neighbor nodes will recalculate their posterior probability values after their processors receive the transition information and then transmit the results to their neighbor nodes until the influences of the evidences have been transmitted through all nodes in the Bayesian network. Consequently, the probability distribution under various states of the required node can be obtained. If the probability of the required node is the largest for the certain state, this node will the most probably occur under this certain state. As a result, the causal reasoning results can be obtained to get the final software reliability evaluation results.

4 Sensitivity Analysis of the Reliability Influence Factors Based on the Bayesian Network

4.1 The Diagnosis Reasoning of the Bayesian Network

After the software reliability evaluation results are obtained, if we want to improve the reliability level, we can use the diagnosis reasoning of the Bayesian network to find the critical factors which may affect the reliability results.

Suppose A_1 , A_2 …are the 'reasons' which cause the results and the event B is obtained after the experiment, the posterior probability $P(A_i|B)$ reflects the probability of each 'reason' after the experiment. If we want to use the Bayesian formula to calculate the corresponding probability, the prior probability $P(A_i)$ should be determined first, then $P(B|A)$ should be determined, finally the $P(A_i|B)$ can be calculated by the Bayesian formula which can be shown as follows:

Taking Fig. 1 for example, calculation steps for determining the influence factors x_i (i=1, 2…4) which causes x5=y1 are shown as follows:

- Calculating the values of P(x5=y1|x4=y1) and P(x5=y1|x3=y1);
- Calculating the value of P(x4=y1|x5=y1) and judging the value is whether more than 0.5. Because the sum of P(x4=y1|x5=y1) and P(x3=y1|x5=y1) is 1, if the value of P(x4=y1|x5=y1) is more than 0.5, it shows that the influence degree of x4 that causes x5=y1 is larger then go to the end; if the value of P(x4=y1|x5=y1) is less than 0.5, it shows that the influence degree of x3 that causes x5=y1 is larger and then go to step 3);
- Repeating the above two steps, then determining the influence degree of x1, x2 and x3;

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{i=1}^{\infty} P(A_i)P(B \mid A_i)} \tag{1}$$

Finally the node which has the largest influence degree will be obtained and should be considered as much as possible in the practice. The quality of the most important factor should be ensured as much as possible. If a node has not less than two father nodes, the probability of each state of each father node should be calculated under the condition that the results are satisfied, and the node which has the largest probability is the required node.

In the second step, the case that x4=y2, x4=y3, x3=y2 and x3=y3 are not considered, because x5=y1 is shown here, we only consider the factors which have larger influence on the results (i.e. x5=y1). Similarly, if x5=y2 is shown here, we also only consider the factor which has the largest prior probability when its state is y2, and this factor is the main reason that causes x5=y1 or x5=y2 and should be considered more carefully.

4.2 The Diagnosis Reasoning and Result Analysis

If the software reliability's level is not satisfied with the requirement, or the reliability level should be improved, the diagnosis reasoning can be used to find the critical factors which have influence on the reliability evaluation results. The better improvement results can be obtained if we take some behaviors to these critical factors. The application process of the whole method can be shown by the following case study.

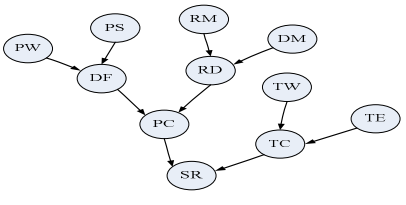


Fig. 2. The Bayesian network topology structure of the case study

- Determining the network nodes.

We select the following environment factors to get the node set of the Bayesian network: programming workload (PW), technology level and general quality of the developers (PS), testing workload(TW), testing environment(TE), reliability design method and technology(RM), development design method and technology(DM), software requirement analysis and detailed design quality(RD), development programming difficulty(DF), software complexity(PC), testing coverage and adequacy(TC), software reliability(SR), that is X={PW,PS,TW,TE,RM,DM,RD,DF,PC,TC,SR}.

- Determining the states of the nodes: for the convenience, we define the state of each node as (high,med,low).
- Constructing the Bayesian network topology structure: the detailed construction result is shown as Fig. 2.

- Determining the probability distribution of the nodes of the Bayesian network

According to the history data as well as the expert experience, the probability distribution of each metric value can be evaluated. The marginal probability distribution of each leaf node in Figure 2 is different for the various software projects, development teams, testing teams and the development environments. Supposing the marginal probability distribution which is obtained by the expert experience is shown in Table 3. The required conditional probability table of the non-leaf nodes is obtained by the real data of the project report and through the analysis of the project managers and experts. Taking DF for example, its conditional probability table is shown as Table 4.

Table 3. The marginal probability distribution of the leaf nodes

The metric	High	Med	Low	∑
programming workload(PW)	0.2	0.6	0.2	1
Technology level and general quality of the developers(PS)	0.25	0.55	0.2	1
Testing workload(TW)	0.68	0.23	0.09	1
Testing environment(TE)	0.33	0.51	0.16	1
Reliability design method and technology (RM)	0.37	0.42	0.21	1
Development design method and technology(DM)	0.41	0.32	0.27	1

Table 4. The conditional probability table of the development programming difficulty (DF)

DF	PW=L			PW=M			PW=H		
	PS=L	PS=M	PS=H	PS=L	PS=M	PS=H	PS=L	PS=M	PS=H
H	0.000	0.155	0.349	0.171	0.000	0.665	0.427	0.478	1.000
M	0.000	0.634	0.453	0.287	1.000	0.231	0.342	0.408	0.000
L	1.000	0.211	0.198	0.542	0.000	0.104	0.231	0.114	0.000

- The causal reasoning calculation

The causal reasoning calculation is obtained according to the marginal probability and the conditional probability shown as Table 4, and the following results can be obtained (only taking DF for example):

$$P(DF = H) = \sum_{PW,PS} P(DF = H \mid PW, PS)P(PW)P(PS) = 0.25$$

$$P(DF = M) = \sum_{PW,PS} P(DF = M \mid PW, PS)P(PW)P(PS) = 0.551$$

$$P(DF = L) = \sum_{PW,PS} P(DF = L \mid PW, PS)P(PW)P(PS) = 0.199$$

The calculation results of the other nodes are shown in Table 5. The final calculation results are: The probability that the reliability level is high is 0.332, middle is 0.451 and low is 0.217. Thus we suggest that the probability that the reliability level is middle is the largest.

Table 5. The probability table of the other nodes

The Node	High	Med	Low	∑
The quality of the requirement analysis and the detailed design(RD)	0.404	0.444	0.152	1
Software complexity (PC)	0.335	0.425	0.240	1
Testing coverage and adequacy(TC)	0.343	0.411	0.246	1
Software reliability (SR)	0.332	0.451	0.217	1

- Diagnosis reasoning calculation

Then we can get SR=M and determine the influence level of each influence factor on this result. The determining process is: first comparing the influence level of PC and TC, the calculation process is shown as follows.

$$P(SR = M \mid PC = M) = \sum P(SR = M \mid PC, TC = M)P(TC \mid PC = M) = 0.54$$

$$P(SR = M \mid TC = M) = \sum P(SR = M \mid TC, PC = M)P(PC \mid TC = M) = 0.6117$$

$$P(TC = M \mid SR = M) = \frac{P(TC = M)P(SR = M \mid TC = M)}{P(TC = M)P(SR = M \mid TC = M) + P(PC = M)P(SR = M \mid PC = M)} = 0.67$$

Because 0.67>0.5, the influence level of TC on SR is larger than PC, then determining the influence level of TW and TE on TC. Similarly we get the following result $P(TW = M \mid TC = M) = 0.54$. Because 0.54>0.5, the influence level of TW on TC is larger than TE.

According to the above results, we consider that TC and TW have the more important influences on the final result, i.e. SR=M. For improving the software reliability level better, we can get the significant results by taking some measures on these two critical factors.

5 Conclusion

This paper constructs the reliability evaluation model based on software reliability influence factors by the Bayesian network. The Bayesian network topology structure can be constructed by the causal relationships between the environment factors and software reliability. Then the Bayesian network is used to get the software reliability qualitative evaluation results by reasoning. Finally, the sensitivity analysis for various factors can be obtained by the diagnosis reasoning of the Bayesian network. The sensitivity analysis results can help the manager find the corresponding factors which should be controlled to guide the resource allocation and achieve the optimal management of software reliability.

References

[1] Michael, R.L.: Handbook of software reliability engineering. McGraw-Hill and IEEE Computer Society Press, New York (1996)

[2] Fenton, N.E., Neil, M.: A critique of software defect prediction models. IEEE Transactions on Software Engineering 25(5), 675–689 (1999)

[3] Neil, M., Fenton, N.E.: Predicting Software Quality using Bayesian belief networks. In: Proceedings of the 21st Annual Software Engineering Workshop, 1996, pp. 217–230. NASA Goddard Space Flight Centre (1996)
[4] Wooff, D.A., Goldstein, M., Coolen, F.P.A.: Bayesian Graphical Models for Software Testing. IEEE Transactions on Software Engineering 28(5), 510–525 (2002)
[5] Pai, G.J., Dugan, J.B.: Empirical Analysis of Software Fault Content and Fault Proneness Using Bayesian Methods. IEEE Transactions on Software Engineering 33(10), 675–686 (2007)
[6] Wu, G.Q., Wei, J., Huang, T.: A Dynamic QoS Assessment Approach for Internetware with Uncertainty Reasoning. Journal of Software 19(5), 1173–1185 (2008) (in Chinese)
[7] Li, H.F., Lu, M.Y., Li, Q.Y.: Summary of the Research on Software Reliability Modeling Considering Environment Factors. National Doctoral Academic Forum of Aeronautics & Astronautics (2008)
[8] Zhang, X.M., Hoang, P.: An analysis of factors affecting software reliability. The Journal of Systems and Software 50, 43–56 (2000)
[9] IEEE Std 982.1-1988, IEEE Standard Dictionary of Measures to Produce Reliable Software
[10] Hoang, P.: Software reliability. Springer, Singapore (2000)

Study on Method of Layer Modelling in Geological Body Visualization

Niu Wenjie

College of Electromechanical Engineering, Petroleum University of China (East China), Dongying City, Shandong Province, 257061, China
niu_wj67@126.com

Abstract. Dealing with the present situation of geological body visualization, Firstly Delaunay triangulation principle and its region boundary segmentation algorithm in the presence of a large number of fault lines, complex boundary conditions is discussed. Then, Geostatistics, Kriging interpolation method is used to create a mathematical model of layer. In the end, with the modeling process, a specific example is used to better verify the geological layer model of the proposed modeling method.

Keywords: geological body visualization, layer model, Delaunay triangulation under restricted condition, Kriging.

1 Introduction

Data visualization has been widely used in medicine, meteorology, environmental protection and agriculture [1,2]. Rolling in oil exploration and development process, because the underground stratum structure, invisible reservoir distribution and awareness about the uncertainty of the data visualization decisions in the process of petroleum exploration and development plays a very important decision support effect. Data visualization can be artificial seismic wave, logging and drilling and other means to obtain a reflection of stratum structure, reservoir depth, sand thickness, porosity, permeability and other parameters of the samples in order to provide a continuous distribution of the geological graphics workers, as their judgments, decision-making.

Petroleum geological data visualization often needs to create geological models. In the process of Geological body modeling, there are three steps to establish layer model. Firstly the region studied has to be subdivision [3], and then the variable is interpolated to establish mathematical models, and finally the traditional methods of computer graphics is used to display the model. At present, for containing the complex boundary, fault lines and other aspects of regional geological constraints triangulation method is not perfect, we use the boundary subdivision algorithm to deal with the problem. In addition, Interpolation methods to establish mathematical model of geological layer, the domestic general use the Inverse distance weighted method[4,5] or the Least square method. These methods can not accurately reflect the features of geological layers, we propose the geostatistical, the Kriging method to solve this problem [6]. The following, based on the oil geological visualization background, we specific study the method of geological layer modeling and use examples for specific applications.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 567–574.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Triangulation of the Study Region

We adopt Delaunay triangulation technology to Subdivide the exploration target region in Geological map, since Delaunay triangulation [7,8] is the Dirichlet / Voronoi diagram of the dual geometry, mathematics well, Delaunay triangulation is unique grid cell automatically equilateral triangle approximation algorithm has a good local changes, these excellent properties to a large extent to meet the requirements for geological regional subdivision in the process oflayer modeling.

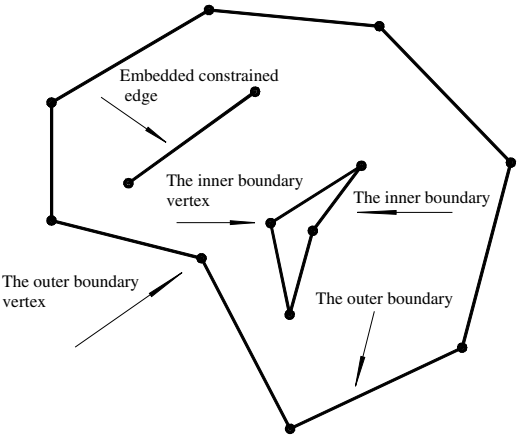


Fig. 1. The domain constraints

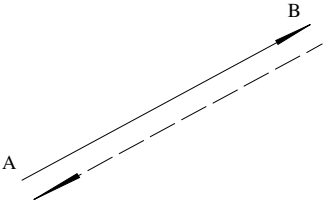


Fig. 2. Embedded constrained edge of regional triangulation

2.1 Delaunay Triangulation Algorithm Based on the "Cavity"

Delaunay triangulation is divided into point set triangulation and regional triangulation. Point set triangulation is that a given set of P to construct triangular mesh, and grid node mast be sample points. Regional triangulation is to construct triangular mesh only by given boundary region, including the generation of the mesh node and element connection. For the constraint regional Delaunay triangulation, broadly speaking, the restriction element for the region, the internal and external boundary, namely:

(1) Constrained point: vertices on internal and external boundary, edge vertices embedded constraints.
(2) Constrained edges: outer boundary, inner boundary, embedded constraint edge (inner boundary).

Figure 1 and figure 2 show the situation of above (1),(2), where embedded constraint edge AB can be seen as a boundary which area is zero.

2.2 Delaunay Triangulation of Point Set Based on the "Cavity" Algorithm

For Delaunay triangulation of point set there are many kinds of algorithms, the "cavity" algorithm is simple and efficient. Based on the "Cavity" algorithm, regional Delaunay triangulation algorithm with constraint can be conducted. The algorithm first generates a triangle that contains all points, and as the initial grid, and then gradually inserts each point in the grid as the new vertex.

In order to set a point in the Delaunay triangulation by adding a new vertex, it must delete the triangle where its circumcircle contains the new vertex to form a hollow. This hollow is called " Delaunay cavity", and the newly added point is called "Delaunay cavity nucleus" (see Figure 3).

Delaunay cavity boundary is formed by line segments, and connection the Delaunay cavity nucleus with each segment to generate new triangles is in line with Delaunay criterion [5]. Therefore, using this method can obtain a new Delaunay triangulation including the new vertex (see Figure 4).

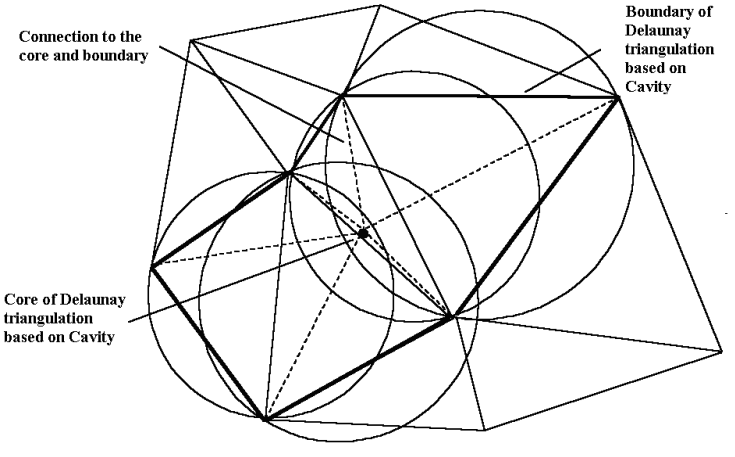


Fig. 3. Delaunay cavity (by the circumcircle contains the core of the triangles)

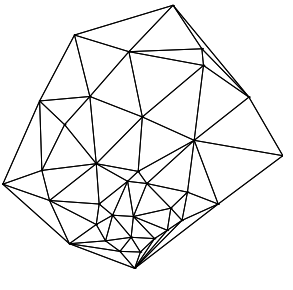


Fig. 4. Result of Delaunay triangulation based on Cavity

2.3 The Principle of Regional Delaunay Triangulation with the Constraints

Regional Delaunay triangulation with constraints is a more difficult problem to solve. There are lots of research and ways on this at home and abroad. Among them, the "boundary segmentation" algorithm [5] is an effective method.

Delaunay constraint domain with boundary subdivision ideas are as follows:

(1). First, the boundary line, limited edge is subdivided into small segments.
(2). Then, the vertex set of the small segment is Delaunay triangulated by " Delaunay cavity" method.
(3). For each small segment, check the existence of triangular grid.

If a small segment in the grid does not exist, the midpoint is inserted to subdivide the segment, and added to the point set of triangulation, the center of the circle circumcircle containing vertex of the segment is used to subdivide the triangle and added to the point set of triangulation also.

(4).returns (2), re-use of "cavity" algorithm, until all the small segment in the presence of the grid algorithm is ended.

Fig. 5. Survey line data distribution in the study region

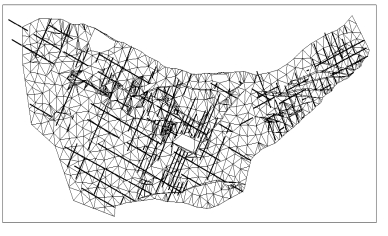


Fig. 6. Result of constrained regional Delaunay Triangulation

Constrained regional Delaunay Triangulation is generally used for the study area complex, uneven distribution of sample data points or number of very small, triangulation with more complex internal and external boundaries, the case of the embedded side constraints. Figure 5 is the seismic line data in a local area in China's eastern oil fields during the oil exploration and development, the data capacity of more than 170,000, concentrated in less than one hundred of the survey line. There are many complex subdivision restrictions - the fault line in the study region. In this case, we use the regional Delaunay triangulation with the constraints and figure 6 shows the result

3 Interpolation to Establish Mathematical Model of Layer

After geological layer in the study region was triangular meshed, we need to use mathematical tools to interpolate triangular mesh vertex to establish mathematical model of geological layer based on the known spatial sampling data (in this case the seismic line data). Interpolation of data visualization methods commonly used include:

the Inverse distance weighting method, the Polynomial least squares method. Geostatistics, Kriging is an unbiased, optimal interpolation method. Unbiased means expectation of interpolation results equal to the expectation of theoretical value, and optimization is the meaning of Interpolation results of the minimum deviation. It is proved [4] that, Kriging interpolation result is better than that of the Inverse distance weighted method, and the Polynomial least squares method. To adapt to different conditions, Kriging interpolation method is divided into a number of ways, such as: Ordinary Kriging method, Universal Kriging method. When the sample data (seismic survey line data or drilling data) does not meet the Second-order stationary assumption, and present drift (defined below) Universal Kriging method can be used. Here Universal Kriging method is discussed necessarily.

4 Universal Kriging

We set the study region is A, the regional variables (ie, the depth strata variable) is $\{Z(x) \in A\}$, and x is spatial location of sampling points in the attribute value (depth value), then under the principle of Kriging interpolation, the estimate value of unsample point x_0 is sum of value of known sampling points with weight, namely:

$$z(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) \tag{1}$$

$\lambda_i, i = 1, 2, \cdots n$ is the unknown weights.

If the variable in study region does not meet the following second-order stationary hypothesis:

(1) Mathematical expectation of $Z(x)$ exists and is equal to a constant:

$$E[(Z(x)] = m \quad (\text{constant}) \tag{2}$$

(2) Covariance of $Z(x)$:

$$Cov(x_i, x_j) = E\{Z(x_i) - m][Z(x_j) - m]\} \tag{3}$$

exists and only the relative position between two points.

Or the variable does not meet the following intrinsic hypothesis:

(1) Expected increment of $Z(x)$ to 0:

$$E[Z(x_i) - Z(x_j)] = 0 \tag{4}$$

(2) Incremental variance of $Z(x)$ exists and is stationary:

$$Var[Z(x_i) - Z(x_j)] = E[Z(x_i) - Z(x_j)]^2 \tag{5}$$

That regionalization variable $Z(x)$ in the entire study region exist drift components: $E[Z(x)] = m(x)$

At this point, $Z(x)$ can be assumed the generalized linear model:

$$Z(x) = \sum_{i=0}^{K} f_i(x)\beta_i + R(x) \tag{6}$$

Of which:

$$E[Z(x)] = m(x) = \sum_{i=0}^{K} f_i(x)\beta_i \tag{7}$$

usually to be polynomial $(0 \le p \le 2)$, $R(x)$ is assumed to meet the second-order stationary regionalized variables, known as the residual of $Z(x)$, and:

$$\begin{aligned} E[R(x)] &= 0 \\ E[R(x_i), R(x_j)] &= Cov(x_i, x_j) \end{aligned} \tag{8}$$

Thus, $Z(x)$ the generalized linear model consists of two parts: one is the deterministic component, that is, the mathematical expectation of $Z(x)$, the other is the residual $R(x)$.

Since:

$$E[z(x_0)] = \sum_{i=0}^{K} f_i(x_0)\beta_i \quad , \quad E[z^*(x_0)] = E[\sum_{i=1}^{n} \lambda_i z(x_i)] = \sum_{i=1}^{n} \lambda_i [\sum_{j=0}^{K} f_j(x_i)\beta_j] \tag{9}$$

Based on unbiased terms, $\lambda_i, i = 1, 2, \cdots, n$ can be obtained by weighting coefficient to be determined to satisfy the relationship:

$$\sum_{i=1}^{N} \lambda_i f_k(x_i) = f_k(x_0) \qquad k = 0, 1, \cdots, K \tag{10}$$

Under constraint of unbiased condition, it can be obtained solving equations of weights $\lambda_i, i = 1, 2, \cdots, n$ by making the Universal Kriging interpolation variance minimum:

$$\begin{cases} \sum_{i=1}^{n} \lambda_i C(x_i, x_j) - \sum_{k=0}^{K} \mu_k f_k(x_j) = C(x_0, x_j) \\ \sum_{i=1}^{n} \lambda_i f_k(x_i) = f_k(x_0) \qquad k = 0, 1, \cdots, K \end{cases} \quad j = 1, 2, \cdots, n \tag{11}$$

Corresponding Kriging interpolation variance is :

$$\sigma_{UK}^2 = C(x_0, x_0) + \sum_{i=1}^{n} \lambda_i C(x_i, x_0) - \sum_{k=0}^{K} \mu_k f_k(x_0) \tag{12}$$

Based on the above principle we can use Kriging method to interpolate mesh nodes shown in Figure 6 to establish Mathematical model of geological layer by seismic line data. Practice found that the data does not meet the second-order stationary hypothesis, and exist drifts, so we use Universal Kriging interpolation method.

5 Geological Layer Model Visualization

After Kriging interpolation method is used to interpolate every triangle mesh vertex, and mathematical model of geological layer is established, the model needs to be displayed on the computer. The way of display can be the contour map, the three-dimensional model map or the gradient map. These ways above are all based on the principle of the traditional computer graphics. Figure 7 shows the contour map of model of geological layer after the regional triangulation and interpolation shown in Figure 6. Figure 8 is the three-dimensional grid map of the model, figure 9 is color map of the model, and figure 10 is the gradient map of the model.

Fig. 7. Model for the layer contour map

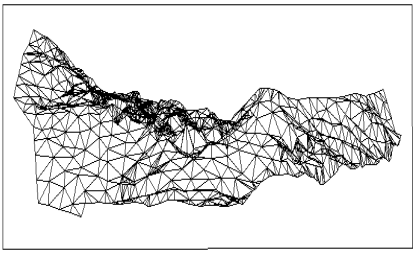


Fig. 8. Three-dimensional grid of layer model

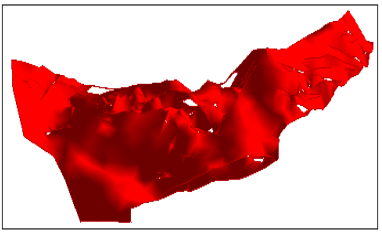


Fig. 9. Color display map of layer Model

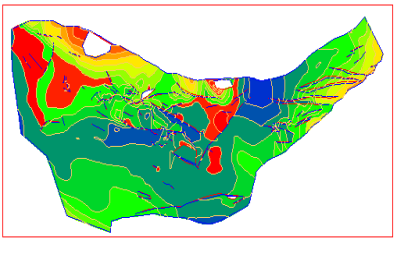


Fig. 10. Gradient map of layer model

6 Conclusion

In petroleum exploration and development process, the use of visualization tools to model the structure of underground formations, reservoir physical parameters (porosity, permeability, etc.) distribution models in an intuitive graphical way to show up for the petroleum science and technology workers to make the right decision-making program. In this paper, visualization of geological bodies, we studied the modeling method of geological layer, presented the modeling process and the steps, and used regional Delaunay triangulation method - "boundary segmentation" algorithm to deal with region subdivision containing a large number of fault lines in the complex boundary. On the mathematical model of the geological layer, we used the Kriging interpolation method. Through specific example better verified the layer modeling results.

References

[1] Tang, Z.S.: Three-dimensional data visualization. Tsinghua University Press, Beijing (1999)
[2] Tang, R.X.: Computer Graphics (revised). Science Press, Beijing (2001)
[3] Rosenblum, L.J.: Research Issues in Scientific Visualization. IEEE Computer Graphics & Application, 61–85 (March 1994)
[4] Franke, R.: Scattered Data Interpolation:Test of Some Methods. Mathematical Computation 38, 181–200 (1982)
[5] Foley, T.: Interpolation and Approximation of 3D and 4D Scattered Data. Computers and Mathematical Application 13, 711–740 (1987)
[6] Niu, W.J.: The study on theory and application for Kriging modeling and its algorithms for 3-D data field visualization, pp. 86–91. Beijing University of Aeronautics Doctoral Thesis, Beijing (2002)
[7] Yang, Q.: Restricted Delaunay triangulation, pp. 145–182. Beijing University of Aeronautics Doctoral Thesis, Beijing (2001)
[8] Li, Z.G.: Broad domain in the field of general three-dimensional scattered data visualization technology, pp. 105–132. Beijing University of Aeronautics Doctoral Thesis, Beijing (1997)

Design and Realization of a Novel Low-Voltage High-Efficiency On-Chip Current Sensing Technique

Wei HaiLong, Geng ZengJian, Liu Zhi, and Liao Xue

Institute of Xi'an Microelectronics, Xi'an, 710054, China
weihl007@126.com

Abstract. In this paper, a novel Low-Voltage, High-Efficiency on-Chip Current Sensing Circuit is designed and realized for the Current Mode Control BUCK DC/DC Converter.Differ from the conventional current-sensing techniques, the proposed circuit utilize the sensing of power MOS turn-on resistance Ron technique which lower the sensing dissipation and improve the power efficiency, meanwhile, the current-sensing circuit without OP amplifier so that it can work at the 2.2V supply voltage. The Chip is simulated and fabricated based on the 0.5μm 2P3M Mixed Signal CMOS process. Test result shows, under heavy load conditions, that efficiency can reach 93%.

Keywords: Dc/Dc Converter, Low-Voltage, High-Efficiency, Current-Sensing.

1 Introduction

Motivated by the rapidly development of today's consumer market, the portable electronics devices become increasingly popular. Particularly, Tiny size, Low-cost, Low-voltage, High-efficiency DC/DC converters are more important for the application systems. Monolithic CMOC/BICMOS BUCK converters which can qualify these requirements become one of the most important issues in the Smart Power Management Integrated Circuits(IC). [1]

Current-Mode Control (CMC)DC/DC converters success in Power Management since it takes the advantages of automatic over current protection, and faster dynamic response than its counterpart, Voltage Mode Control, which is easy to design but has slower transient response. [2]

In this paper, The BUCK DC/DC converter apply Constant off-time PWM technique combines with PWM/PFM dual operation mode and a Low-Voltage, high-efficiency current-sensing circuit without OP amplifier is presented.

2 Conventional on-Chip Current Sensing Techniques

Figure.1 illustrates the Classic on-Chip current sensing technique, uses M7 to monitor power MOS (M8) current (Ipass), which is K times of M7. [3] When M8 and M10 are turned on (Ctr=0), Vb is equal to Vx and Vb is copied to Va by OP amplifier. At on state (Ctr=0), IM7 would vary with IM8 and IM7 is the summation of Iref and Isen, which Iref is a constant current. Therefore, Isen is 1/K of Ipass and passes Rsen to

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 575–582.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

provide sensing voltage Vsen for the current loop use. At off state(Ctr=1), M12 turns on and provide a high voltage for Vb to avoid OP amplifier cut-off. This architecture needs an OP amplifier to act as a voltage mirror. Its finite gain and bandwidth would limit the accuracy of the sensing circuit. In this current sensing circuit, a positive feedback loop exists in the loop (Vb-Vc-M4-Vb) and, therefore, the current sensing circuit has to design carefully to avoid oscillation. The sensing circuit is not suitable for low voltage application because M3 and M6 apply cascade architecture, which needs two Vgs voltage in Vc.

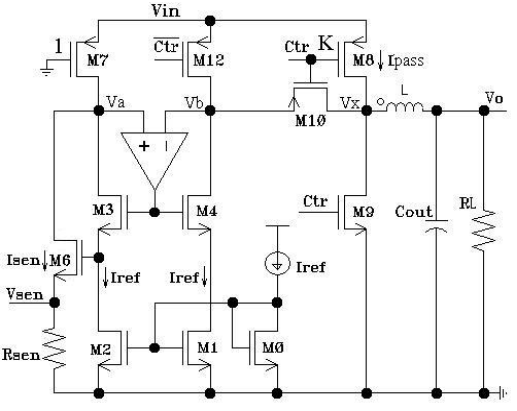


Fig. 1. Classic on-Chip current sensing technique

Another low-voltage on-chip current sensing circuit, shown in Fig. 2, is proposed to deal with the disadvantages of supply voltage limitation in above architecture and its operation is similar to the previous circuit[4].

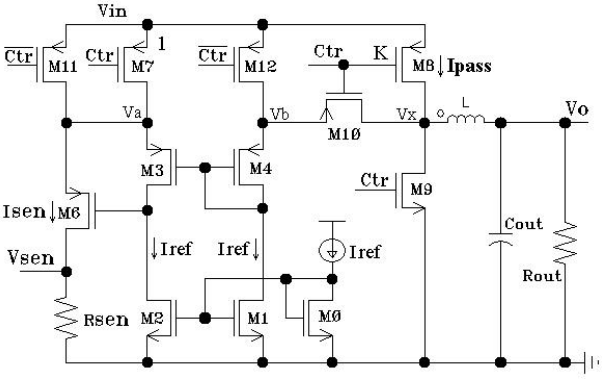


Fig. 2. Low voltage CSC without OP amplifier

During ON-state(Ctr=0), the power transistor M8, the sensing transistor M7, and the switch M10 all turn on, and Va ≈Vb≈Vx.

If the relationship of the aspect ratio between the sensing transistor M7 and the power transistor M8 is 1: K, then the current of IM7 ≈Ipass/K, and IM7 is the summation of Iref and Isen. So the Isen varies with the IM7 since which varies with the Ipass. After sensing the power MOS current Ipass, Isen will pass Rsen to provide sensing voltage Vsen for the current loop use.

During OFF-state, M7, M8, and M10 all turn off, and M11 and M12 turn on to provide the current path for source M1~M4 with normal operation. When M11 and M12 turn on, Va and Vb will pull up to Vin. At the OFF-state, the sensing current will drop to zero.

This sensing circuit doesn't use any OP amplifier to mirror Vx and run the risk of oscillation caused by the positive feedback. Substitute PMOS M6 in Fig.2 for NMOS M6 in Fig.1 can further reduce the power supply by (2Vgs-Vsat) which more suitable for Low-Voltage application.

The existence of Isen will increase power dissipation and lower the efficiency of converter, the reduced efficiency could be expressed as follow:

$$\Delta\eta \approx \frac{V_{in}}{V_{out}K} \tag{1}$$

From the above equation designers can promote efficiency by increasing the aspect ratio K. However, due to the limitation of fabrication process, this factor is chosen from 100 to 1000 for reliability which, at least, lower the efficiency by 0.1 to 1 percent.

3 Low-Voltage High-Efficiency On-Chip Current Sensing Technique

The proposed Low-Voltage, High-Efficiency on-Chip Current Sensing circuit is shown in Fig.3, When M7 and M8 are turned on (Ctr=0), assuming M6 and M7 have the same turn-on resistance Ron, Vy can be obtained as:

$$V_{in} - V_y = \frac{R_{on} I_{pass}}{2} \tag{2}$$

Where Vin is the supply voltage, RM8 is the turn-on resistance of POWER MOSFET M8, and Ipass is the output current of the BUCK converter.

Resistors R1~R3 combine with the transistors M0~M5 form the differential amplifier which senses the signal between Vin and Vy. The current IR1, across R1, is the current summation of Iref and Isen. if the relationship of the aspect ration between R1 and R2 is 1:1, then the voltage generated across R1 by Isen equals the voltage drop across M6, namely, Vin minus Vy. Choose the R3 is N times of R1, then the relationship between sensed voltage Vsen and sensed current Isen can be represented as follow:

$$V_{sen} = \frac{N}{2} R_{M8} I_{pass} \tag{3}$$

When M7 and M8 are turned off (Ctr=1), M6 still keep conduction which provides the current path for the source M1~M4 with normal operation. At the same time, Vy will be pulled up to Vin and the sensing current drops to zero.

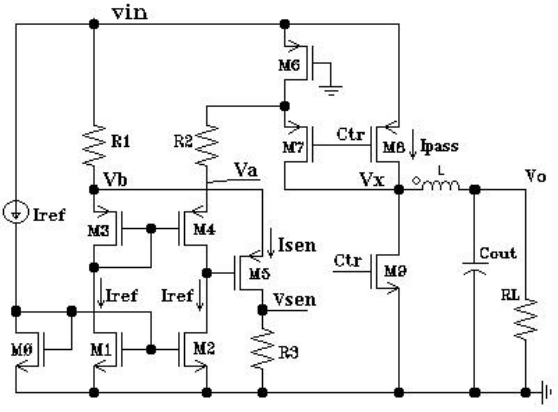


Fig. 3. Proposed on-Chip Current Sensing Technique

3.1 Low-Voltage High-Efficiency Features

The proposed current sensing technique without OP amplifier and arrange PMOS M5 and R3 as common source structure forces the supply voltage can reach it's minimum value in theory, expressed as follow:

$$V_{in-\min} = V_{sen-\max} + V_{DS-M6} \tag{4}$$

Where Vsen-max is the maximum swing of Vsen, determined by the circuit designer, and VDS-M6 is the voltage drop on PMOS M6 which works in triode region. In the proposed current circuit, the VDS-M6 is about 0.2V,which makes the minimum supply voltage can reach 2.2V.

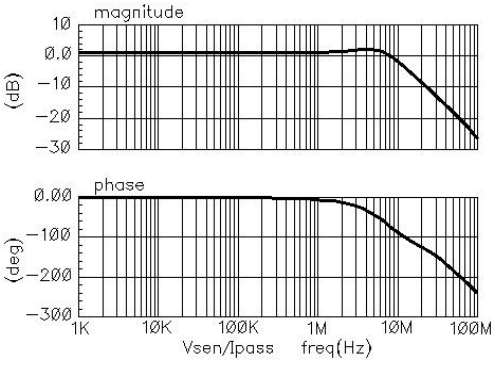


Fig. 4. Frequency response of the proposed CSC

Existence of R1 and RM8 will decrease the efficiency of proposed topology, for modern monolithic DC/DC converters, RM8 within the range of 0.1Ω ~ 0.5Ω. The worst case occurs when RM8 is at its maximum value, namely, 0.5Ω. Suppose a 1A

output current and R1 is 5K, the sensed current Isen will only be 50uA which is 5 percent of the sensed current in Fig.2 (1mA). The current across M6 and M7 will be added into output current which improves efficiency of the converter further.

Fig.4 shows that the frequency response of the proposed current-sensing circuit. The -3dB bandwidth is 12MHz that can meet the requirements of the most monolithic DC/DC converters which switching from 0.5 MHz to 5 MHz.

3.2 Matching of Transistors and Resistors

Proper Layout selection can eliminate the random and system errors stem from mismatching. Common-centroid Layout designs for M1 and M2, M3 and M4 can suppress the fluctuations from fabrication process and temperature gradients. Resistors matching for R1~R3 including construct matched resistors from a single material, make matched resistors the same width and identical geometries, orient resistors in the same direction , place them in close proximity and interdigitate arrangement, etc.

4 Design and Measurement Results

4.1 System Architecture

Fig.5 illustrates the system architecture of the proposed monolithic DC/DC BUCK converter with on-chip current-sensing technique. This architecture uses a constant off-time, current mode PWM/PFM control which avoids the efficiency rolls with the decreasing input voltage in the conventional fixed frequency current mode control.

Current control via an inner control loop including Current Sensing Circuit, Current Comparator, OFF-TIMER, Gate driver, POWER MOS M8 and M9, Inductor, Output Capacitor and Rload. Voltage control via an outer control loop composited by the feedback resistors RA/RB, Error Amplifier and the current loop .[5] The system can switch automatically between PWM and PFM depending on the loading current, .for an heavy load the chip operations with PWM, when the load is relatively light, the chip automatically switches into PFM operation to optimize the efficiency.

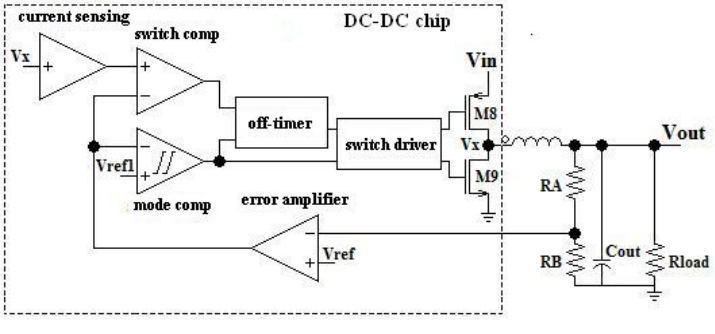


Fig. 5. System architecture of the proposed DC/DC BUCK converter

4.2 Chip-Level Simulation

Fig.6 shows the Chip-Level Simulation of the proposed BUCK converter with an 2.2V input and stable 1.8V output , respectively.

During 0-200uS(Fig.6.a), the system is powering up so there is no output current; when the load is relatively light(Iload =20mA), the chip operates with PFM mode that improve efficiency greatly, as show in Fig.6.b; when there is heavy load(500mA) the system switches into PWM automatically, shown in Fig .6.c.

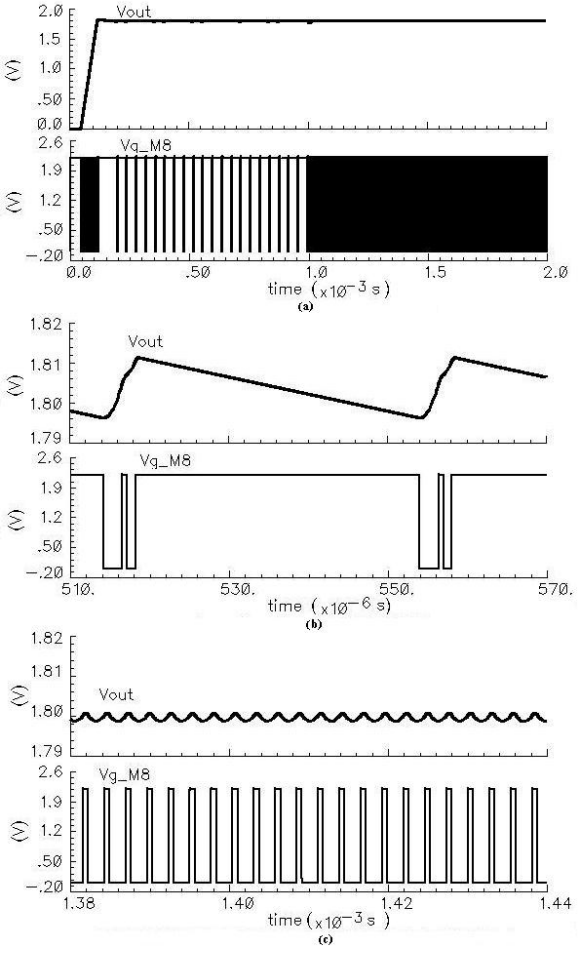


Fig. 6. (a)DC-DC converter of transient response,(b)transient response of PFM MODE, (c)transient response of PFM MODE

Fig.7 illustrates that proposed current-sensing circuit can be accurate sensed the POWER MOS current (Ipass) variation and for a 700mA peak switching current the sensed current is only 19uA which better than the conventional sensing technique.

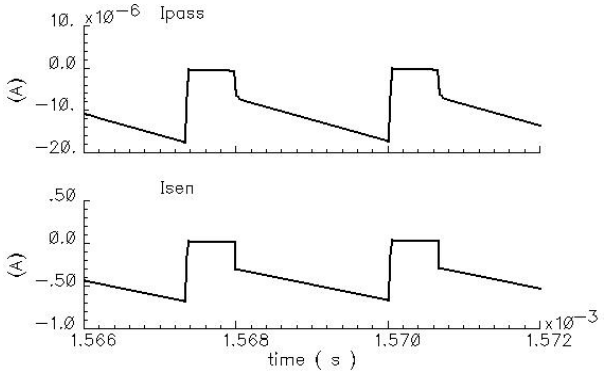


Fig. 7. Sensed current Isen versus Output Current Ipass

The chip-level simulation results are shown in Table.1

Table 1. Chip-level simulation results

Electrical Characteristics	Simulation Results	Units
Input Voltage Range	2.2-5.5	V
Quiescent Current	130	μA
Constant OFF-time	512	nS
Peak Switching Current	1.3	A
Load Regulation	7	mV/mA
Line Regulation	0.4	%/V
Efficiency	96	%

4.3 Physical Verification

The Current-mode DC/DC BUCK converter has been implemented with a 0.5μm 2P3M Mixed Signal CMOS process. Its Layout is shown in Fig. 8, the size of the whole chip is 1.1 mm^2 .The fabricated current-mode DC/DC BUCK converter has been tested and the overall performance is summarized in Table.2.

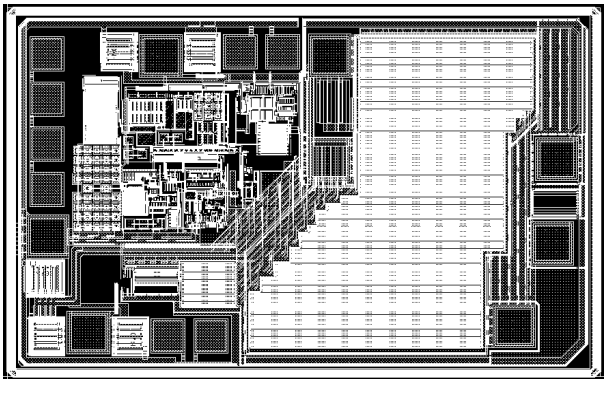


Fig. 8. Physical Layout of the Proposed BUCK Converter

Table 2. Overall Chip Performance

Electrical Characteristics	Measurement Results	Units
Input Voltage Range	2.3-5.5	V
Quiescent Current	114	μA
Constant OFF-time	530	nS
Peak Switching Current	1.1	A
Load Regulation	9	mV/mA
Line Regulation	0.6	%/V
Efficiency	93	%

5 Conclusion

In this paper, the design of monolithic current-mode DC/DC buck converter with a novel Low-voltage, High-efficiency on-chip current sensing circuit is presented, including the comparison analysis, circuit implementations, physical verification and measurement results. For a 700mA peak switching current the sensed current is only 19uA.The bandwidth of the current-sensing circuit is 12MHz with the input voltage of 2.2V. The measurement results show that the efficiency of the proposed converter is up to 93% for loading current is 500mA.

References

[1] Bonizzoni, E., Borghetti, F., Malcovati, P.: A 200mA 93%Peak Efficiency Single-Inductor Dual-Output DC-DC Buck Convert. In: 2007 IEEE International Solid-State Circuits Conference, USA, pp. 526–527 (2007)

[2] Givelin, P., Bafleur, M., Tournier, E., Laopoulos, T., Siskos, S.: Application of a CMOS current mode approach to on-chip current sensing in smart power circuits. IEEE Proc. -Circuits Drvires. Syst. 142(6) (December 1995)

[3] Lee, C.F., Mok, P.K.T.: A Momolitic Current-Mode CMOS DC-DC Converter With On-Chip Current-Sensing Technique. IEEE Journal of Solid-State Circuits 39(1), 3–14 (2004)

[4] Cheng, K.-H., Su, C.-W., Hsin-Hsin.: A High Accuracy and High-Efficiency on-Chip Current Sensing for Current-Mode Control CMOS DC-DC Buck Converter, pp. 458–461. IEEE (2008)

[5] Brown, M.: Practical Switching Power Supply Design. Academic, San Diego (1990)

Construction of Mamdani Fuzzy Classifier Based on Genetic Algorithm

Zhou Weihong* and Xiong Shunqing

School of Mathematics & Computer Science, Yunnan Nationalities University,
Kunming 650031, China
ynzwh@163.com

Abstract. Most of the fuzzy classifiers are created by fuzzy rules based on apriori knowledge and expert's knowledge, but in many applications, it's difficult to obtain fuzzy rules without apriori knowledge of the data. To solve this problem, a new way of creating Mamdani fuzzy classifier based on Mamdani fuzzy logical system is proposed in this paper, and the new fuzzy classifier is improved with the genetic algorithm further. The result of data simulation with Iris data shows the new Mamdani fuzzy classifier has minimum number of features, minimum number of fuzzy rules and better precision.

Keywords: fuzzy classifier, Mamdani fuzzy logical system, fuzzy reasoning, genetic algorithm.

1 Introduction

At most of the simple classification problem, fuzzy classifier is constructed usually using priori expert knowledge, but for many complex systems, it is difficult to obtain complete priori expert's knowledge, so that the study of fuzzy classification system base on data structure becoming a hot spots at this research area, mainly including fuzzy clustering [1], fuzzy neural network [2], and evolutionary fuzzy system [3] and so on. At these methods, a universal approximator fuzzy model features was used, the accuracy is the main object of the fuzzy model, but the explanation of fuzzy model has not been a good deal.

Compared to neural network model, Fuzzy model integrate the knowledge representation and reasoning mechanism with the priori expert experience and knowledge, consistent with people's habits of mind, its structure and membership function parameters have obvious semantic meaning, it can be easily understood its internal operation mechanism by studying the rules of fuzzy system, in conclusion, the explanation is the most prominent feature of a fuzzy model. How to automatically constructed the fuzzy systems with accuracy and explanatory from data analysis, is the key point of this research area.

The explanatory of Fuzzy classification system, so far has no clear definition, but is generally believed that the explanatory of fuzzy classification system is closely related with the number of the characteristics variables, the number of fuzzy rules,

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 583–590.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

and the characteristics of membership functions, and the fuzzy classification system with fewer number of feature variables, fewer number of fuzzy rules has better explanatory.

This article base on the work of the references[4], the characteristics variable of antecedent was selected and the domain of it was partitioned, further more, we combined it with Mamdani [5], the output was improved, as a result, a new fuzzy classifier was generated, and a data simulation was taken with Iris data, contrast to references[1], the result of new classifier has fewer number of variables and rules, and the results of classification is better.

2 An Improved Mamdanis Fuzzy Classifier Base on Mamdanis Fuzzy Logic System

Fuzzy reasoning is a process of mapping specific input space to a specific output space by fuzzy logic calculation, which involves membership functions, fuzzy logic operation, If-then rules and some basic theory. Mamdani and Sugeno-type fuzzy logic system is the two main type fuzzy logic system commonly used today, especially the Mamdani algorithm is the most commonly used algorithm. In this paper, a new fuzzy classifier base on Mamdani fuzzy logic system was constructed and described as following.

2.1 Typical Fuzzy Classification System

Fuzzy classifier consist of fuzzy system and classifier. For the classify problem with n-dimensional M class and N sample, which $x = (x_1, x_2, \cdots x_n)$ is the characteristic variable, $\{C_1, C_2, \cdots C_M\}$ is the output class, the typical rules of fuzzy classification divided into following three types:

1. The consequent of fuzzy rules is the specific classes.

R_i : If x_1 is $A(1,k)$, x_2 is $A(2,k)$, $\cdots$, x_n is $A(n,k)$, then class is C_i ... ... (1)

Here $A(1,k)$, $\cdots$, $A(n,k)$ is the membership function defined in the characteristic domain, which can to be the triangle, Gaussian-shaped, trapezoidal function etc.

2. The consequent of fuzzy rules is the specific classes with confidence.

R_i : If x_1 is $A(1,k)$, x_2 is $A(2,k)$, $\cdots$, x_n is $A(n,k)$, then class is C_i with w_i ... ... (2)

Here w_i is the confidence of i th rule.

3. The consequent of fuzzy rules is all possible classes with confidence.

R_i : If x_1 is $A(1,k)$, x_2 is $A(2,k)$, $\cdots$, x_n is $A(n,k)$, then class is $C_i(o)$ with $w_i(o)$ (3)

Here $C_i(o)$ is all possible classes, $w_i(o)$ is its corresponding confidence value.

2.2 Mamdani Fuzzy Logic System

1.The fuzzy rules which we used is show as formula (4), the consequent is a single-point membership functions of the output class.

R_i : If x_1 is $A(1,k)$, x_2 is $A(2,k),\cdots,x_n$ is $A(n,k)$, then class is $A(n+1,Ci)$ ... ... (4)

2.Which $A(n+1,Ci)$ is the corresponding membership function of output class, we usually take triangle, Gaussian, trapezoidal membership function, the single point support membership function is taken in our paper, because the result of fuzzy output is still in a continuous domain(2.3), so that the result of fuzzy reasoning need further process by a classifier to get a discrete categories.

For the K th membership functions of variables x_j, were used bilateral Gaussian membership function which bring forward in references[4].

$$A_{(j,k)} = \begin{cases} \exp\left[-\left(\dfrac{x_j - c_{(j,k)}}{\omega^l_{(j,k)}}\right)\right]^2 & x_j \le c_{(j,k)}. \\ \exp\left[-\left(\dfrac{x_j - c_{(j,k)}}{\omega^r_{(j,k)}}\right)\right]^2 & x_j > c_{(j,k)} \end{cases} \tag{5}$$

At equation (5), $c_{(i,k)}$, $w^l_{(j,k)}$, $w^r_{(j,k)}$ is the mid-point, left width and rightwidth of bilateral Gaussian membership function $A_{(j,k)}$ respectively.

We use a single point function for the output membership function:

$$A(n+1,Ci) = \begin{cases} 1 & y = Ci \\ 0 & y \ne Ci \end{cases} \tag{6}$$

We get the output of fuzzy system by "winner takes all" approach with sample x_k, the system output is the class membership function those have the rules with greatest incentive degree of the corresponding class:

$$x_k \in A(n+1,C_l), l = \arg(\max(\beta_i(x_k))) \quad 1 \le l \le M \tag{7}$$

At formula (7), where β_i is the incentive intensity of i th rule:

$$\beta_i(x_k) = \prod_{j=1}^{n} A_{ij}(x_k) \tag{8}$$

At formula (8), $A_{ij}(x_k)$ is the membership function for the j th variable of rule i.

2.3 Improved Mamdani Fuzzy Classifier

Fuzzy rules we discussed above constitute a fuzzy rule base, but the output of this Mamdani inference system ϖ is still a continuous domain, it is necessary to classify the final output of the Mamdani system by using a classifier . Because the number of

class is M , so the output domain is a continuous domain on $[1, M]$, then we get the classification output of the Mamdani system by dividing the domain into M blocks, each block is a class, as show at the following formula:

$$i = \text{int}(\frac{M}{M-1}(\varpi - 1)) \tag{9}$$

At formula (9), i represent the i th class. The flow chart of Classification is shown in Figure 1.

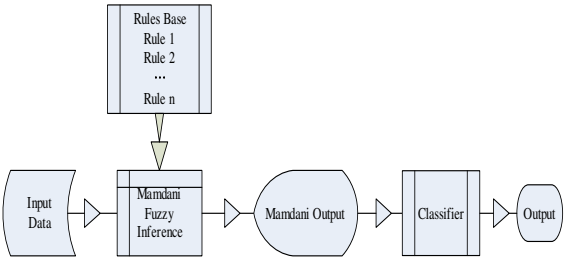


Fig. 1. The flow chart of the fuzzy classifier

The Mamdani fuzzy classifier fuzzy rules:

R_i : If x_1 is $A(1,k)$, x_2 is $A(2,k), \cdots, x_n$ is $A(n,k)$,the y is $A(n+1, C_i) \rightarrow y \ is \ C_i$ ... ... (10)

3 Optimal Mamdani Fuzzy Classifier by Using Genetic Algorithms

Genetic algorithm was used to obtain optimum fuzzy classifier in our paper, the population was constructed by an appropriate number fuzzy classifiers, the objective function is the weighted value with the samples number of wrong classification by classifier, the number of characteristic variable of classifier, and the number of fuzzy rules of classifier.

In order to get a appropriate population, and higher explanatory of fuzzy classifier, alleviate the curse of dimensionality, the grid method[4] was used to divided the domain of all variables, and multi-objective genetic algorithm was used to achieve the selection of characteristic variable and the partition of fuzzy domain. First input variable was partition by grid method and the antecedent was determined, then the membership function of the consequent was determined by "winner takes all" strategy, lastly, output population was get through classifier's training by genetic algorithm method.

3.1 Obtained the Antecedent of Fuzzy Rules by Coding Method

The code of population is consist of the part of characteristics variables and the part of partitioning the corresponding variable, the part of characteristic variables is the

control genes, the length of this part equal to the length of characteristics variables, each variable corresponds to one bit, if the variable is selected, then the corresponding bit is set to 1, and if the variable is not selected , the corresponding bit is set to 0.

First step, we get the domain of each variable by statistical method. We first divided the domain of each variable evenly by experience, so a candidates mid-point for partition was determined; second step, in order to realize the fuzzy partition, we determine the mid-point by genetic algorithm. The mid-point of each variable's domain is selected by fixed-length coding, and each mid-point corresponds to one bit, if the corresponding bit is set to 1, then select the mid-point, if the bit is set to 0, give up the mid-point; third step, determine left width and right width corresponding to the mid-point. The width values is relate to the overlap factors μ of the membership function [4], the left and right width $w^l_{(j,k)}$, $w^r_{(j,k)}$ expressed as follows:

$$w^l_{(j,k)} = [-\frac{(c_{(j,k)} - c_{(j,k-1)})^2}{\ln \mu}]^{\frac{1}{2}},\ w^r_{(j,k)} = [-\frac{(c_{(j,k+1)} - c_{(j,k)})^2}{\ln \mu}]^{\frac{1}{2}} \tag{11}$$

Here $c_{(j,k)}$ is the center of the K th membership function of the J th characteristic variables, formula(10) ensure if the value of K th membership function is equal to μ, it corresponds variable's value overlap with the center of $K-1$ and $K+1$ membership function. It make the membership function has higher explanatory.

3.2 Get Consequent of the Fuzzy Rules by Training Method

"Winner takes all" strategy was adopted to select consequent here, and a single point membership function was used as the membership function of the class of consequent. Base on determining the antecedent of the fuzzy rules, firstly, the maximum ownership value of the corresponding class was calculated with training samples of each class, excitation intensity was calculated according to formula(8), secondly, compare the largest ownership value of various classes, the class with largest ownership value become the class of consequent, thirdly, define the membership function of the rule consequent, at last, as a result, a fuzzy classifier is designed show at figure 1 in section 2.3.

3.3 Achieving Mamdani Fuzzy Classifier with Genetic Algorithm

The basic genetic algorithm was adopted at this section, and binary encoding was adopted, the genes consist of the control genes and the domain partition selection genes. The weight calculated by three objectives is used Fitness function.

$$F_1 = \omega_1 f_1 + \omega_2 N_f + \omega_3 N_r \tag{12}$$

At formula(12), f_1 is the number of samples wrongly classified, N_f is the number of the selected characteristic variables, N_r is the number of fuzzy rules, ω_1, ω_2, ω_3 is the weight of the objective function respectively.

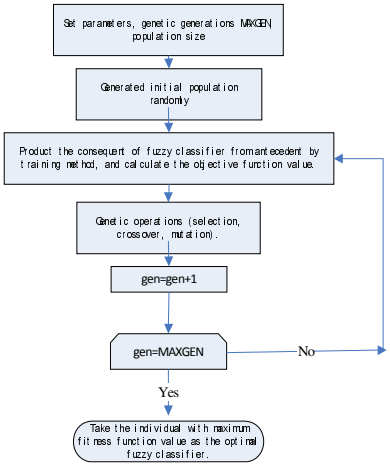


Fig. 2. Algorithm flow chart

Algorithm are described as follows:

Step1: Set parameters, genetic iterate generations MAXGEN, population size L.

Step2: Generated initial population randomly, the number of antecedent of classifier is L (gen = 1).

Step3: Get the consequent of fuzzy classifier from antecedent by training method, and get the objective function value by calculate.

Step4: Genetic operations (selection, crossover, mutation).

Step5: If gen = MXAGEN then go to Step6 else go to Step3.

Step6: Take the individual with maximum fitness function value as the optimal fuzzy classifier.

Algorithm flow chart is as follows Figure 2.

4 Data Simulation

Iris Database is a very typical classification data can be used as criteria for the evaluation of various classification algorithms. It created by famous British statistician R.A.Fisher, the Iris database is a four-dimensional data(pental length, pental width, sepal length, sepal width), total consist of 150 samples, three classes (1 - Iris-setosa ,2 - Iris-versicolor ,3 - Iris-virginica), each class has 50 samples. Class 1 can separate from other two class completely, and Class 2 and Class 3 has overlap.

Our experiment were implemented on Matlab7.0, genetic algorithm adopted the Genetic Algorithm Toolbox developed by Sheffield University. The parameters of Genetic Algorithm set as following: genetic iterate generation:100; the number of individuals of the population:20; crossover probability :0.9; mutation probability:0.7; the membership function overlap factor $\mu = 0.15$: the fitness function parameters: $\omega_1 = \omega_2 = \omega_3 = 1$.

The result of our fuzzy classifier base on Genetic algorithm, only one variable, three rules, as specified in Table 1:

Table 1. Fuzzy classifier rules

Rule	Characteristic variables x_4			Category No
	Mid-point	Left-width	Right-width	
1	0.5	0.2904	0.5808	1
2	1.3	0.5808	0.5808	2
3	2.1	0.5808	0.2904	3

The content of table 2 is the comparison result between our new Mamdani classifier and the Classifier in reference[4], as we can see from Table 2, the number of variables of classifier in reference[4] is always 2, the number of rules were 12, 3, 3, the number of samples classified correctly were 146, 145, 147, the number of samples classified wrongly is 4, 5, 3 respectively. Compare to the classifier in reference[4], Our Mamdani fuzzy classifier has the fewer number of variables, less then classifier in reference[4] by 1, the number of rules is the same, but the number of samples wrong classification is a little more than the classifier in reference[4], more then the preliminary classification system[4] by 2, more then the first step optimization system[4] by 1, more the second step optimal classification system[4] by 3, in conclusion, compared with reference[4], our Mamdani fuzzy classifier has better explanatory, but has a little weaker accuracy.

Table 2. Comparison of the Mamdani fuzzy classifier and the classifier in reference[4]

	Variable number	Number of rules	Number of Correct-classified	Number of Misclassified
Preliminary classification system	2	12	146	4
First step in optimizing classification system[4]	2	3	145	5
Second step optimized classification system[4]	2	3	147	3
Mamdani fuzzy classification system	1	3	144	6

5 Conclusion

In this paper, we achieved a new Mamdani fuzzy classifier based on a Mamdani fuzzy logic system without apriori knowledge. This new classifier is optimized by improving consequent of principle and membership function with the genetic algorithm. Data simulation shows that the new classifier has a good classification performance, its explanation is better than reference[4], but its Classification accuracy is a little weaker, our next work is to optimize the fuzzy classifier further in order to achieve better classification accuracy.

Acknowledgements. This paper is supported by National Natural Science Foundation(60963026) and Yunnan Provincial Department of Education Research Foundation (2010J071).

References

[1] Fan, Y., Ji, H., Zhang, H.: A new combination clustering algorithm based on fuzzy classification. Computer Applications 28(5), 1024–1027 (2008)
[2] Fuller, R.: Introduction to neuro-fuzzy systems. Physica-Verlag, New York (2000)
[3] Cordon, O., Gomide, F., Herrera, F., et al.: Hoffman, Ten years of genetic fuzzy systems: current framework and new trends. Fuzzy Sets and Systems 141, 5–31 (2004)
[4] Xing, Z., Hou, Y., Jia, L.: Fuzzy classification system based on Multi-objective genetic algorithm. Southeast University (Natural Science Edition) 36(5), 725–731 (2006)
[5] Wu, X., Lin, Z.: Matlab auxiliary fuzzy system design. Xidian University Press, Xi'an (2002)
[6] Lei, Y., Zhang, S., Li, X., Zhou, C.: Matlab genetic alorithm toolbox and application. Xidian University Press, Xi'an (2005)
[7] Wang, L.: A course in fuzzy systems and control. Tsinghua University Press, Beijing (2003)
[8] Wang, X., Cao, L.: Genetic Algorithm - Theory, Applications and Software. Xi'an Jiaotong University Press, Xi'an (2002)
[9] Xing, Y.: A classification algorithm based on genetic algorithms and fuzzy rules. Jishou University (Natural Science Edition) 27(2), 31–35 (2006)

Change Detection Using High Spatial Resolution Remotely Sensed Imagery

Zhang Ruihua[1] and Wu Jin[2]

[1] School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China
[2] Experiment Center, Air Force Radar Academy, Wuhan 430019, China
zrh0302@163.com

Abstract. This paper presents an evidence theory based change detection method capable of utilizing multiple image features.With a moving window, we first get the structural similarities of both time phase image visual features and construct the basic probability assignment function (BPAF) of D-S evidence theory. We then fuse all the evidence and get the changed image areas with decision rules. Comparative work on different experimental areas, combinations of change evidence and with other methods has been carried out. It shows that our method prevents effectively the detection errors from only utilizing single feature and thus improves the detection precision. Furthermore, since the image similarity is derived from image statistical features rather than original grey, texture and gradient features, this method is robust to low calibration precision.

Keywords: Evidence theory, change detection, high spatial resolution image, multi-feature, structural similarity.

1 Introduction

Remote sensing image change detection is to extract ground feature changes using two phase images of the same region by digital image processing and pattern recognitio n [1]. It has many applications including environment protection, agriculture resource investigation and water conservancy construction.The ground feature changes represent as the changes of image gray value, structure, shape and texture features, and they are the main clues for change detection. Since gray values are the most feasible image feature, they are widely used in change detection. A lot of methods have been proposed including band ratio, band difference, regression model, NDVI, and PLA [2]. Among these methods, the band difference method is the simplest, which first obtains a difference image by subtract one time phase image from the other, and distinguishes changed or unchanged pixels with some threshold.This kind of method has many shortcomings. Firstly, it is difficult to specify the threshold in many. $case$ [3]. Secondly, image features need to be precisely extracted from the same position of different time phase images, which puts forward rigid precision criterion to radiant calibration and image registration [4].

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 591–597.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

With the rapid development of spaceflight, sensors and computer technologies, the spatial resolution of remote sensing has improved significantly. High spatial resolution remotely sensed imagery is with voluminous data, more ground details,and more serious spectral confusion. It makes traditional change detection methods only relying on image spectrum not applicable [5]. In high spatial resolution remote sensing applications, it has already become a hot spot to develop new change detection methods based on steadier image visual features.

In this study, we propose a novel remote sensing image change detection method comprehensively considering edge,texture and gradient changes. We calculate the structural similarity of the three features of both time phase images, create the basic probability assignment function (BPAF) of evidence theory for each feature, and implement multi-evidence fusion.With decision rules, we get the changed image areas. Comparative work on different experimental areas, combinations of changed evidence and with other methods has been carried out.It shows that our methodprevents effectively the detection errors from utilizing single feature and thus improves the detection precision.

2 Method Principle and Steps

This method involves image visual feature extraction, feature histogram statistics, structural similaritycalculation of feature vectors, and multi-evidence fusion, etc.We first introduce the basic principle of evidence theory and structural similarity,and then give the method detailed steps.

2.1 Evidence Theory

D-S evidence theory is a mathematical tool for uncertainty modeling and reasoning [6]. It considers both the objectivity and subjectivityof evidence in probability reasoning, which is different to Bayesian theory. The probability of evidence theory is the belief to a proposition based on evidence. The theory implements induction and estimation based on multi- source information, and then gives a correct decision.

Given a non-empty set U, we call U a frame of discernment,which is composed a series of mutual exclusive and exhaustive elements. Given a proposition A inthe problem domain, it belongs to 2^U .Define BPAF:m: $2^U \rightarrow [0,1]$ in 2^U ,where $m(A)$ represents the belief exactly committed to the subset A of U. (1) If $A \subset U$, $m(A)$ denotes the determined belief to A;(2) if $A=U$, $m(A)$ denotes an uncertain assignment; (3) if $A \subseteq U$and $m(A)>0$, A is called a focal element of m. D-S evidence theory combines different evidence with orthogonal sum. Let $m_1, m_2, m_3 .... m_n$ be n BPAFs in 2^U , their orthogonal sum is denoted as

$$m = m_1 \oplus m_2 \oplus m_3 ... \oplus m_n \tag{1}$$

and is defined as

$$m(\phi) = 0 \tag{2}$$

D-S evidence theory is an effective tool for uncertainty reasoning.It has many successful applications in the fields of remote sensing, e.g., image classification, road extraction and shadow detection from multicolor airborne images [7]. In this study, it is used to combine edge, texture and gradient changes for multi-feature image change detection.

2.2 Structural Similarity

Structural similarity, which was first proposed by Wang et al.(2004), has already been used in image quality evaluation.The structural similarity of two vectors X and Y is defined as:

$$SSIM(X,Y) = [I(X,Y)]^{a} \bullet [c(X,Y)]^{\beta} \bullet [s(X,Y)]^{r} \tag{3}$$

where

$$I(X,Y) = \frac{2u_X u_Y + C_1}{u_X^2 + u_Y^2 + C_2}, C(X,Y) = \frac{2\sigma_X \sigma_Y + C_2}{\sigma_X^2 \sigma_Y^2 + C_2}, s(X,Y) = \frac{\sigma_{XY} + C_3}{\sigma_X \sigma_Y + C_3} \tag{4}$$

$u_X, u_Y, \sigma_X, \sigma_Y, \sigma_X^2, \sigma_Y^2$ and σ_{XY} are the mean, standard deviation,variance and co-variance respectively, α, β and γ are the weights, and C_1, C_2 and C_3 prevent zero division. When $\alpha=\beta=\gamma=1$, $C_3 = C_2/2$, Eq. (4) can be simplified as

$$SSIM = \frac{(2u_X u_Y + C_1)(2\sigma_{XY} + C_2)}{(u_X^2 + u_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \tag{5}$$

In our study, the BPAFs are derived from the structural similarities of image features. Generally speaking, the similarity of vector ***X*** and ***Y*** SIM(***X***, ***Y***) should satisfy these conditions: (1) Bondedness. For example, $0 \le |SIM(\boldsymbol{X}, \boldsymbol{Y})| \le 1$. (2) Symmetry.That is SIM(***X***, ***Y***) = SIM(***Y***, ***X***); (3) Single maximum. That is SIM(***X***, ***Y***) = 1 if ***X***=***Y***. Many commonly used similarity measures defined in vector space, including Euclidian, mahalanobis and Minkowski distances, do not satisfy bondedness. Histogram intersection does not satisfy symmetry. Correlation coefficient, as a commonly used measure, does not satisfy the single-maximum condition. As illustrated in Fig.1, the correlation coefficient of the two parallel but not very 'close' vectors reaches 1.0. Because structural similarity is derived from the mean, variance and covariance, and satisfies all the above needs,it represents similarity better than correlation coefficient. In this example, the structural similarity is only 0.64. For its advantages,structural similarity is used to measure vector similarity and create the BPAF.

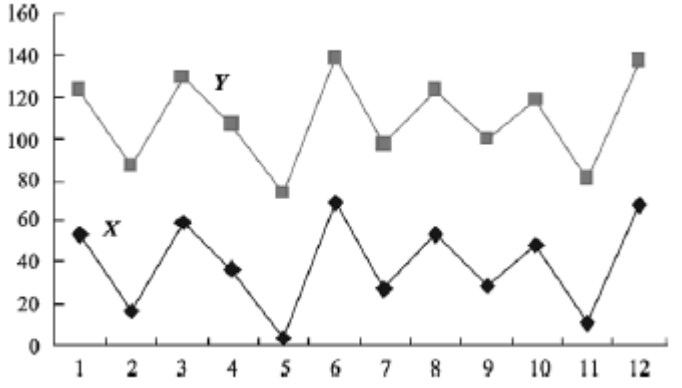


Fig. 1. Two high correlated vectors

2.3 Method Steps

We first implement image registration and radiant calibration for two time phase images. In our case, histogram matching is used to make pixel value distribution of an image close to the other, which causes fewer disturbances to the detection of image changes.

We extract visual features for both time phase images.GLCM contrast is used as image texture descriptor, Canny operator to extract image edges, and Sobel operator to extract image gradient.

With a moving window, we calculate the texture, gradient and edge similarities of the two time phase images. If the original visual features are used for similarity calculation, they are sensitive to the precision of image registration and noise. We thus construct derivate features with histogram statistics.Firstly, we quantize the gradient and texture features with Eq.(6):

$$h=(x-min)/M,M=(max-min)/L \quad (6)$$

where x is the original feature value, “max” and “min” are the maximum and minimum values in this moving window respectively,and L is the quantitative level. We then create the two time phase feature histograms and calculate their structural similarities S_1 and S_2 with Eq. (5). Since edges are binary formatted, we design the edge pattern distribution histogram (EPDH) to calculate the edge similarity. EPDH is a statistical relationship descriptor of edge distribution patterns and their occurrence frequency. As illustrated in Fig. 2, in a 2×2 sub-window, there are totally 14 edge istribution patterns. Traveling the edge map with these templates and count their matching times, we get an EPDH. For example, Fig. 3(a) is the edge map in a 9×9 window, and Fig. 3(b) is the corresponding EPDH. We can then get the edge structural similarity S_3 with Eq. (5).

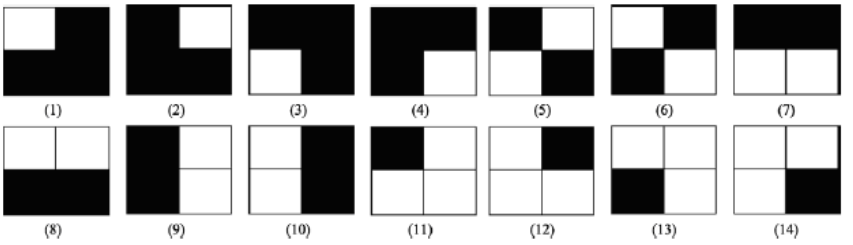


Fig. 2. Edge distribution patterns

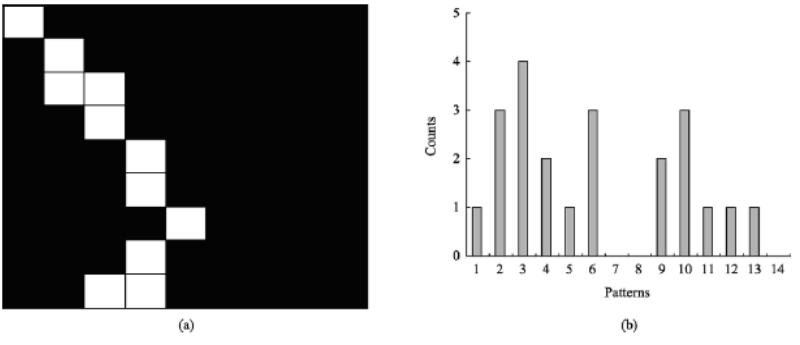


Fig. 3. Edge map and its distribution pattern histogran

We design the discernment frame $U=\{Y,N\}$, where Y represents the changed classes, and N the unchanged classes. The non-empty subsets of $2U$are $\{Y\}$, $\{N\}$, $\{Y,N\}$. Create the BPAFs according to the structural similarities of texture, gradient and edge with Eq. (7):

$$\begin{aligned} m_i(\{Y\}) &= (1.0 - S_I) * a_i \\ m_i(\{N\}) &= S_i * a_i \\ m_i(\{Y,N\}) &= 1.0 - a_i, i = 1,2,3 \end{aligned} \tag{7}$$

where a_i is the trust degree of evidence to the discernment frame. Finally, we implement evidence fusion with Eq. (3).Setting suitable BPAF thresholds for changed and unchanged classes, we can then get the changed image areas.

3 Experimental Analyses

Our experimental data are ALOS imagery received in November,2003 and December 2005 respectively, located in Jiangning,Nanjing, with spatial resolution 2.5m.As illustrated in Fig. 5(a) and Fig. 5(b), the first experimental area ranges from 31°39′49.84″N —31°40′30.01″N, 119°2′9.66″E—119°3′44.74″E,with 1012×477 pixels in size. The second experimental area ranges from 31°52′45.40″N—31°54′4.47″N,118°46′2.58″E—118°47′35.97″E, with 999×963 pixels in size.

In the fist experimental area, we wanted to find the best evidence fusion way by comparing the seven combinations of the evidence of gradient, edge and texture. We also wanted to validate evidence theory. The second experimental area was used to verify our conclusion from experimental area 1. We thus used uniform algorithm inputs in both cases. It was as follows: the moving window size was 9×9, gray level 32, window size 5×5, orientation 0°, and inter-pixel distance 1 pixel to calculate GLCM contrast; the standard variance was 0.6, ratio of low to high threshold 0.7, and ratio of pixels with values lower than high threshold was 80%. The quantitative level was 14 to quantize texture and gradient features; C_1 and C_2 were 0.3 and 0.6 to calculate structural similarity; a_1, a_2 and a_3 were 0.9, 0.9 and0.95 respectively for evidence fusion; the BPAF supporting image changes was larger than 0.25, or that supporting un-changes was less than 0.7.

3.1 Experiment

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

After image registration and radiant calibration, we tested the seven combination modes of evidence. Fig. 4(c) to Fig. 4(i) are the detection results; Fig. 4(j) is the detection result of similarity validation using gradient feature with similarity threshold 0.7, and Fig. 4(k) illustrates the result by combining edge and gradient with operator 'OR'.

From Fig. 4(c) and Fig. 4(d), it was found that the fusion of gradient, texture and edge, or fusion of edge and gradient issued better detection, with most changed areas found. But they were with different sensitivity to changes. For example,the threefeature-fusion method recognized the bottom right area in Fig.4(c) as a changed area which is with very tiny changes but the two-feature-fusion method omitted it.

3.2 Experimental Analyse

From Fig. 4(e) to Fig. 4(i), we found that the fusion of texture and edge is very sensitive to changes, which detected most changed areas but with a lot of misdetection. Fusion of texture and gradient, which is not very sensitive to changes, detected a lot of changed areas but with un-detection and a lot of misdetection (see an example in Fig. 4(f)). Single feature detection using edge or gradient issuedserious un-detection results. On the contrary,the method using single texture was with serious misdetection.As illustrated in Fig. 4(j), similarity validation method detected many changed areas but with serious misdetection.

We combined the detection results using edge and gradient features by operation 'OR', and compared it with that of the fusion of edge and gradient method. As Exemplified in the marked changed area in Fig. 4(k), the former was with serious un-detection results, while the two-feature fusion method issued correct detection. It proves that evidence theory is an effective method for uncertainty induction, not simple pileup of multievidence.

4 Conclusion

In this study, we propose a novel remote sensing image change detection method by fusing multi-features. We first get the structural similarities of both images and construct the BPAFs of D-S evidence theory. We then fuse all the evidence and get the changed areas. Comparative work shows that our method prevents effectively the detection errors from only utilizing single feature and thus improves the detection precision.Furthermore, since the image similarity is derived from image statistical features rather than original grey, texture and gradient features, this method is robust to low calibration precision. We also point out the openness of this method since we can easily add and combine different evidence to improve its detection precision and applicability.

References

[1] Bruzzone, L., Prieto, D.F.: Automatic analysis of the difference image for unsupervised change detection. IEEE Transactions on Geoscience and Remote Sensing 38(3), 1171–1182 (2000)
[2] Deng, W., Shao, X., Liu, H., Wan, G.: Discussion of remote sensing image classification. method based on evidence theory. Journal of Remote Sensing 11(4), 578–583 (2007)
[3] Jacobs, I.S., Bean, C.P.: Fine particles, thin films and exchange anisotropy. In: Rado, G.T., Suhl, H. (eds.) Magnetism, vol. III, pp. 271–350. Academic, New York (1963)
[4] Strunk Jr, W., White, E.: The elements of style, 3rd edn. Macmillan, NewYork (1980)
[5] Nicole, R.: Title of paper with only first word capitalized. J. Name Stand. Abbrev. (in press)
[6] Yorozu, Y., Hirano, M., Oka, K., Tagawa, Y.: Electron spectroscopy studies on magneto-optical media and plastic substrate interface. IEEE Transl. J. Magn. Japan 2, 740–741 (1987); Digests 9th Annual Conf. Magnetics Japan, p. 301 (1982)
[7] Young, M.: The Technical Writer's Handbook. University Science, Mill Valley (1989)

Optimization of BP Neural Network Classifier Using Genetic Algorithm

Zhou Weihong* and Xiong Shunqing

School of Mathematics & Computer Science, Yunnan Nationalities University,
Kunming 650031, China
ynzwh@163.com

Abstract. In this paper, a new BP neural network classifier was constructed and optimized by Genetic Algorithm, first, the BP neural network was improved by using genetic algorithm[2] to train the initial weights values of the BP neural network[3], second, a new classifier was constructed based on the new BP neural network optimized by Genetic Algorithm. Finally, data simulation experiment was taken and the result of data simulation with famous IRIS data shows that the new BP neural network classifier improved by the Genetic Algorithm has higher accuracy of classification and greater gradient of convergence than the BP Neural Network classifier which Proposed in literature [3].

Keywords: Genetic Algorithm, BP neural network, BP neural network classifier, Convergence, Weight.

1 Introduction

The classification can be used for picking up the model of important data and predict the trend of data, usually, BP neural Network, Rough Set, Fuzzy Set, Bayes Network, Genetic Algorithm were adopted to construct a classifier, the advantage of using BP neural Network for a data classification is simple and fictile. The input/output of BP neural network can be see as a kind of highly nonlinear mapping system, the input data is the data which needs to be classified, and the output data is the classes of input data, the process of BP neural network classifier is similar to the thinking habit of people. But the BP neural network is deficient in convergence slowly and easily plunging into the local optimum. Meanwhile, the Genetic Algorithm have the advantage of Robustness, global optimum and be suitable for parallel processing, which we can use to overcome the deficient of BP neural Network.

2 Improving BP Neural Network Algorithm by Genetic Algorithm

2.1 Introduction of BP Neural Network Algorithm and Genetic Algorithm

The BP(Back Propagation) network is a kind of multilayer feed forward neural network, which can realize arbitrary nonlinear mapping from input to output. Its study

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 599–605.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

rule is through feedback transmission to continuously adapt the weight and threshold value of the network with steepest descent method, to make the error square sum of network the least. According to the Kolmogorov theorem, a BP nerve network with totally 3 layers which can complete arbitrary mapping from n dimension to m dimension, but there exists weaknesses of slow in convergence rate, can not promise to convergence on a global minimum and the network structure is difficult to determine.

Genetic Algorithm[5] (GA) is a set of self-adaptive global-optimistic search algorithm which simulate the process of Biological evolution, which has the advantages of simple universal, does not depend on gradient information, Robustness, and be suitable for parallel processing. Genetic Algorithm (GA), as a set of global-optimistic search algorithm, can search in multiple regions of solution space at the same time, and are able to jump from local optimum and get the global optimum.

2.2 Improving BP Neural Network Algorithm by Genetic Algorithm

he chromosome, so the convergence rate of BP neural network was increased greatly. The improving of tIn this paper, the weight value[6][7] of BP neural network is improved by Genetic Algorithm**,** the chromosome of Genetic Algorithm was constructed by coding with the connection weights of BP neural network, then the natural evolution process was imitated, selection, crossover and mutation operation was carried on the chromosome, finally, the chromosome which represent the optimal solution of the problem was created and we can get the improved connection weights of the BP neural network by decoding with the algorithm is show as follow aspects: expression of the chromosome, definition of the objective function and the fitness function, the construction of the genetic operator.

1. Chromosome coding and the production of the initial weight population. Each chromosome of the initial weight population is produced by the normal regulations way of the BP nerve network which producing the initial weight, any set of integrity neural weight $W_i = \{W_{1,i}, W_{2,i}, W_{3,i}, b_{1,i}, b_{2,i}, b_{3,i}, i = 1,2,\cdots p\}$ equals to a chromosome, there are p chromosomes, the scale of the weight population is p.
2. The definition of the objective function and the fitness function. The objective function can be defined as the least weight of the search network error in all of the evolution generation, the corresponding expression formula is as follows:

$$E(W^*) \le E(W_i)_{gen}, i = 1,2,\cdots, p; gen = 1,2,\cdots, termgen \tag{1}$$

The fitness function of the Genetic Algorithm is take as the evolution target and it can only evolve towards the direction which cause the value of fitness function increasing, thus the suitable conversion between the fitness function and the objective function should be perform. Because of the error in the evolution network is nonzero constant, the reciprocal value of the target function is used as the fitness function, a lager coefficient M was introduce in order to make sure the value of the fitness function is not too small, the last form of the fitness function is:

$$p(W) = \frac{M}{E(W_i)}, i = 1, 2, \cdots, p \tag{2}$$

3. Selection operation, the selection operator in the toolbox of the Genetic Algorithm is chosen in the connection weight evolution.
4. Crossover operation, the crossover operator in the toolbox of the Genetic Algorithm is chosen, in the connection weight evolution.
5. Mutation operation, the mutation operator in the toolbox of the Genetic Algorithm is chosen in the connection weight evolution.

The process of the algorithm of improving the weights of the BP neural network by Genetic Algorithm are as follows:

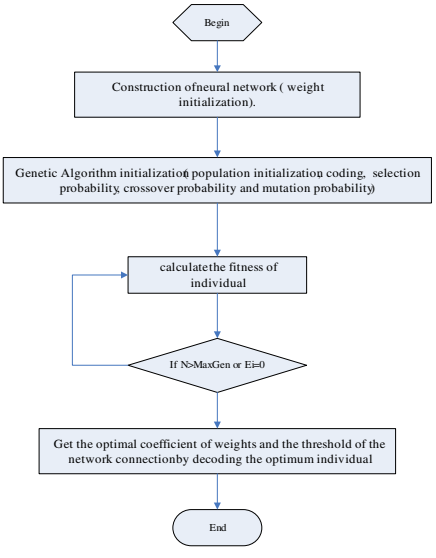


Fig. 1. The flow chart of the Algorithm

First step: construct the neural network (including the weight initialization).

Second step: Genetic Algorithm initialization and the population initialization (including the coding method, selection probability, crossover probability Pc and mutation probability Pm).

Third step: new individual was inserted to the population P, each individual evaluation function is calculated and sorted, the individual was chosen by the probability of the next formulate: f_i is the fitness of individual i, E_i is the square error between the output value of experiment:

Fourth step: judge whether the algorithm is completed. If we have got the satisfied individual or reached the biggest iteration times then end, otherwise goto the third step to get into the next iteration.

Fifth step: If we have reached the performance target which we have established, the optimal coefficient of weights and the threshold of the network connection can get by decoding the optimum individual in the final population. The flow chart is as shown in Figure 1.

3 Construction and Optimization of the BP neURAL Network Classifier

3.1 Construction of BP Neural Network Classifier

When the BP neural network is used as a classifier, it can be treated as two stages from the input, the first stage of the classifier is used as a matching degree calculator and the result of first stage will be sent to the second stage through m input lines parallel, the second stage of classifier is established on the foundation net framework of the third stage, each stage has a exportation through the second stage, and the output is only one "high" while others are "low". After a correct classification result is got, the output of classifier is feedback to the first stage of the classifier, and the weight is adapted by using the error-back-propagation. when the test sample is quite similar to the sample which is studied before, the classifier will respond exactly. The structure of the traditional BP nerve network classifier is as shown in Figure 2.

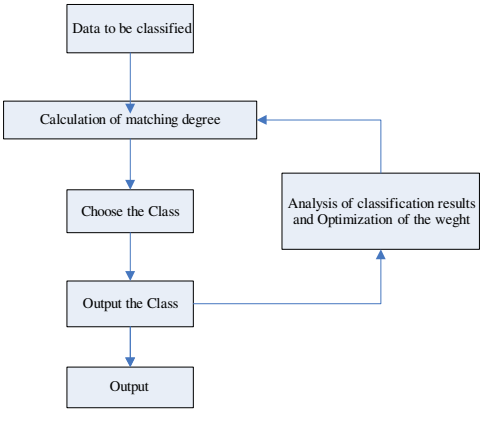


Fig. 2. The structure of BP neural network Classifier

3.2 Optimization of the BP Neural Network Classifier

As a classifier, the BP neural network has advantages of simple and fictile, other classifiers are more complex compared with the BP nerve network classifier, but the BP nerve network algorithm is based on the gradient, which is slow in convergence rate and easily plunging into the local optimum, Genetic Algorithm (GA) is one of the global-optimistic search algorithm, it can search in multiple solution space at the same time, and be able to jump from local optimum and get the global optimum, so we can

take advantage of global superiority and the parallel processing ability of Genetic Algorithm to overcome the deficient of the BP Neural Network.

Based on the 1.2 section of this paper, the process of BP neural network algorithm classifier improved by Genetic Algorithm is as follows:

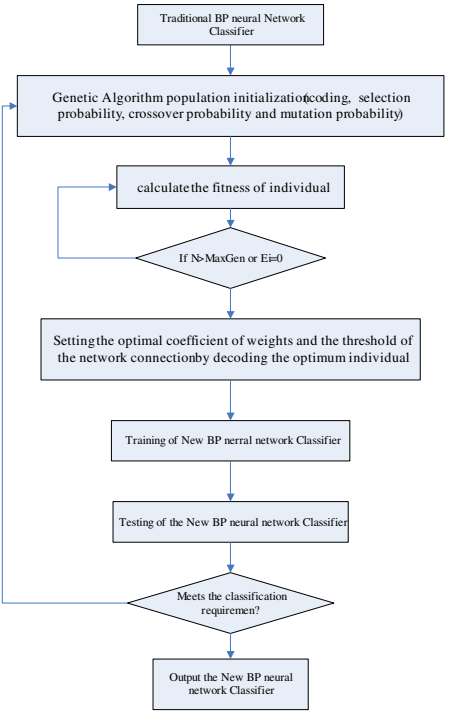


Fig. 3. The flow chart of BP neural network Classifier improved by Genetic Algorithm

Step 1: initialize the BP neural network.
Step 2: calculate the weight range of the BP neural network and the coding length of Genetic Algorithm;
Step 3: initialize the population of Genetic Algorithm.
Step 4: set the corresponding parameters, execute the Genetic Algorithm , choose individual with roulette policy, single point crossover, variation, evolution to produce new generation population, take the square error as the target function value, iterate until there is no generation to reach and get the evolution solution;
Step 5: divide the evolution solution to the weight of the BP neural network;
Step 6: get a more excellent BP nerve network classifier by training with the weight which is got before, otherwise goto step 2.

4 Data Simulation

4.1 Iris System

The Iris system is a common benchmark problem in classification and pattern recognition studies. It contains 150 measurements of four features (sepal length, sepal width, pental length, pental width) from each of three species (setosa, versicolor, virinica). The first class is separate from others clearly, while the second and third class are overlap slightly.

4.2 Experiments and Results

This paper did the experiment of classify the iris data by the classifier of improving BP neural network algorithm by Genetic Algorithm and the classifier of the BP nerve network in literature[2]**,** and carried on a comparison of classification effect between two algorithms**.**

All the program of experiment is run under the programming environment of Windows XP operate system and Matlab 2010a, and all the BP neural networks which are used are 4-25-3 BP neural network, the 150 Iris data which is used in the experiment are divide into 2 sets: the training set and the testing set, the training set contains 90 samples of data and the testing set contains 60 samples of data. The training data is different from the testing data, and if the training data is used as the testing data, the result shows that the classifier is an ideal classifier, the algorithm flow chart is shown as Figure 3.

Table 1. Comparison of the two BP neural network Classifier

Algorithm	BP neural Network	BP neural network optimized by GA
Training Accurate rate	100%	100%
Testing Accurate rate	96.67%	96.67%
Convergence step	469	125
Elapse Time	43	13
Error	0.0001	0.00001
Gradient of Algorithm	0.0000117	0.000150

The experiment result is shown as TABLE 1: Both of the accurate rate of the result of classifying ninety Iris training samples are 100% in the test experiment with the classifier of BP neural network algorithm improving by Genetic Algorithm and the traditional BP neural network classifier; and the accurate rate of classifying another 60 Iris data are all 96.67%, 2 samples was classified wrong with both of two algorithms; the steps of convergence at global optimum of traditional BP neural network classifier are 469, while the BP neural network algorithm classifier improving by Genetic Algorithm are 125, the new algorithm's convergence steps decrease 344 compared with the traditional BP neural network classifier; The elapse running time of new algorithm is also decrease 30 seconds compare with old

algorithm; and the error decrease 0.00009; The gradient of convergence of new algorithm is 0.0001338 greater than old algorithm, the result of the experiment shows that the new classifier improved by the Genetic Algorithm has higher accuracy of classification and needs less time to solve the problem.

5 Conclusion

In this paper, we discuss the optimization of the initial weight of the BP neural network algorithm by Genetic Algorithm, we have also constructed a new BP neural network classifier based on the new algorithm, furthermore, we simulated the new classifier with the Iris benchmark data, the result of simulation experiment shows that the new BP neural network classifier improved by the Genetic Algorithm has more excellent classification ability than the BP neural network classifier in literature[3]**.**

Acknowledgements. This paper is supported by National Natural Science Foundation(60963026) and Yunnan Provincial Department of Education Research Foundation (2010J071).

References

[1] Jiang, Z.: Introduction to Artificial Neural Networks. Higher Education Press, Beijing (July 2001)
[2] Chen, G., Wang, X., Zhuang, Z.: Genetic Algorithm and its application. Posts & Telecommunications Press, Beijing (2001)
[3] Sheng, Z.B.: nerve network principle and MATLAB imitation. Journal of Wei Nan Normal College 23(5), 65–67 (2008)
[4] Chen, F., Zhu, B., He, X.: Genetic Algorithm and nerve network II. Journal of Chemical Engineering College 47(4), 421–426 (1996)
[5] Lei, Y., Zhang, S., Li, X., Zhou, C.: Matlab genetic algorithm toolbox and application. Xidian University Press, Xi'an (2005)
[6] Tian, X., Song, T., Liu, Y.: Improving the structure and the parameter of the BP nerve network with Genetic Algorithm. Journal of Dalian University of Technology 2(6), 69–71 (2004)
[7] Luo, C., Liu, Z., Wang, C., Liu, Z.: Classification of remote Agriculture land cover sensing data based on BP neural network improved by Genetic algorithm. Journal of Agricultural Engineering 22(12), 133–137 (2006)
[8] Xing, Y.: A classification algorithm based on genetic algorithms and fuzzy rules. Jishou University (Natural Science Edition) 27(2), 31–35 (2006)
[9] Guo, C., Tang, H.W.: An improved evolutionary programming algorithm and its convergence. Journal of Numerical Mathematics 24(1), 51–56 (2002)
[10] Wu, J., Nong, J., Jing, L.: Genetic algorithm convergence analysis. Journal of Guangxi Teachers Education University (Natural Science Edition) 21(1), 48–52 (2004)

Output-Feedback Stabilization of Discrete-Time Systems with Multiple Time-Varying Delays

Li Zhang[1] and Fule Li[2]

[1] College of Life Science, Qingdao Agricultural University, Qingdao 266109, China
[2] School of Science and Information, Qingdao Agricultural University, Qingdao 266109, China
lfl2004666@126.com

Abstract. This paper is concerned with the problem of out-feedback stabilization for a class of discrete-time systems with multiple time-varying delays. All the time delays are assumed to be time-varying and bounded. An improved delay-dependent stability condition for the unforced system is proposed in terms of linear matrix inequality (LMI). Then, based on the obtained result, a delay-dependent result for the existence of a static output-feedback controller, which ensures the asymptotical stability of the system with controlled input and multiple delays, is derived.

Keywords: multiple delays, linear matrix inequality (LMI), out-feedback stabilization, Lyapunov function.

1 Introduction

The stability is an important property of any control system with time delays or without any time delays. But the time delays are often encountered in many dynamic systems, such as cold rolling mills, traffic control systems, water resources systems, networked control systems and so on. Generally speaking, time delays are the main source of poor performance and instability in real system. And the stability analysis of time delay systems is relatively more difficult than the systems without any delays. In studies of robust stability problems, linear matrix inequality (LMI) is the most efficient method in that it needs no tuning of parameters and can be efficiently solved numerically by software (matlab).

During the last decades, the problem of stability and stabilization for discrete time systems with delays has received considerable attention. Many results [1-3] have been obtained. The existing results can be classified into two types: delay-independent results [4-5] and delay-dependent results. Since time delay is not taken into consideration in the process of analyzing stability and designing controller, the delay-independent results are considered to be more conservative than delay-dependent ones. In recent years, many researchers are devoted to study the delay-dependent results. Especially, the behavior of real system can become more complicated when the states conclude multiple times delays. And, some delay-dependent stability results [6-8] are derived for systems with multiple delays.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 607–613.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Motivated by [1], in this note we will consider the problem of out-feedback stabilization for a class of discrete-time systems with multiple time-varying delays. Based on the Lyapunov-Krasovskii functional approach, a delay-dependent stability condition for the unforced discrete system is derived in the terms of a linear matrix inequality (LMI). Then the problems of stabilization by static output-feedback controllers are solved.

2 Problem Statement

Consider a class of discrete system with multiple delays

$$\Sigma:\begin{cases} x(k+1)=Ax(k)+\sum_{i=1}^{n}A_i x(k-\tau_i(k))+Bu(k) \\ y(k)=Cx(k)+\sum_{i=1}^{n}C_i x(k-\tau_i(k)) \\ x(k)=\phi(k), k=-\tau_M,-\tau_M+1,\cdots,0 \end{cases} \tag{1}$$

Where $x(k)\in R^n$ is the state vector; $y(k)\in R^m$ is the measurable output; $u(k)\in R^l$ is the controllable input. $A, A_1,\cdots,A_n, B, C, C_1,\cdots,C_n$ are system matrices with appropriate dimensions. $\tau_1(k),\cdots,\tau_n(k)$ are time-varying delays and satisfying

$$\tau_m \le \tau_{mi} \le \tau_i(k) \le \tau_{Mi} \le \tau_M, i=1,\cdots,n \tag{2}$$

where $\tau_m, \tau_{mi}, \tau_{Mi}, \tau_M$ are constant positive scalars. $\{\phi(k)\}$ is a given initial condition sequence.

The following lemma is needed for deriving our main results.

Lemma [1]. Assume that $a\in R^n, b\in R^m$ and $N\in R^{n\times m}$. Then, for any matrices $X\in R^{n\times m}, Y\in R^{n\times m}$ and $Z\in R^{n\times m}$ satisfying

$$\begin{bmatrix} X & Y \\ Y^T & Z \end{bmatrix} \ge 0$$

the following inequality holds

$$-2a^T Nb \le \begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} X & Y-N \\ Y^T-N^T & Z \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}.$$

3 Stability Analysis

First, we consider the problem of stability for discrete system described as the following unforced form

$$\Sigma_1 : \begin{cases} x(k+1) = Ax(k) + \sum_{i=1}^{n} A_i x(k-\tau_i(k)) \\ x(k) = \phi(k), k = -\tau_M, -\tau_M + 1, \cdots, 0 \end{cases} \tag{3}$$

The following theorem is derived.

Theorem 1. The unforced system Σ_1 is asymptotically stable if there exist matrices $P = P^T > 0, X > 0, Y > 0,$ and $Q_i^T = Q_i > 0, Z_i > 0 (i = 1, 2, \cdots, n)$ of appropriate dimensions such that

$$\begin{bmatrix} \Phi_{11} & \Phi_{12} \\ * & \Phi_{22} \end{bmatrix} < 0 \tag{4}$$

$$\begin{bmatrix} X & Y \\ Y^T & Z_i \end{bmatrix} \geq 0, i = 1, 2, \cdots, n \tag{5}$$

where

$$\Phi_{11} = \begin{bmatrix} -P & & & \\ & -\tau_M^{-1} Z_1 & & \\ & & \ddots & \\ & & & -\tau_M^{-1} Z_n \end{bmatrix}, \quad \Phi_{12} = \begin{bmatrix} P(A + \sum_{i=1}^{n} A_i) & PA_1 & \cdots & PA_n \\ Z_1(A-I) & Z_1 A_1 & \cdots & Z_1 A_n \\ \vdots & \vdots & \ddots & \vdots \\ Z_n(A-I) & Z_n A_1 & \cdots & Z_n A_n \end{bmatrix},$$

$$\Phi_{22} = \begin{bmatrix} \Gamma & -Y & -Y & \cdots & -Y \\ * & -Q_1 & 0 & \cdots & 0 \\ * & * & -Q_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & * & -Q_n \end{bmatrix}, \quad \Gamma = -P + n\tau_M X + (\tau_M - \tau_m + 1) \sum_{i=1}^{n} Q_i + nY + nY^T.$$

Proof: It is easy to see that

$$x(k - \tau_i(k)) = x(k) - \sum_{m=k-\tau_i(k)}^{k-1} \zeta(m) \tag{6}$$

where

$$\zeta(m) = x(m+1) - x(m) = (A - I)x(m) + \sum_{i=1}^{n} A_i x(m - \tau_i(m)).$$

Then the unforced discrete system can be transferred into the following system.

$$x(k+1) = (A + \sum_{i=1}^{n} A_i)x(k) - \sum_{i=1}^{n} \sum_{m=k-\tau_i(k)}^{k-1} A_i \zeta(m). \tag{7}$$

Choose a lyapunov functional candidate as

$$V(k) = V_1 + V_2 + V_3 + V_4,$$

where

$$V_1 = x^T(k)Px(k),\ V_2 = \sum_{i=1}^{n}\sum_{j=k-\tau_i(k)}^{k-1} x^T(j)Q_i x(j),$$

$$V_3 = \sum_{i=1}^{n}\sum_{j=-\tau_{Mi}+2}^{-\tau_{mi}+1}\sum_{m=k+j-1}^{k-1} x^T(m)Q_i x(m),\ V_4 = \sum_{i=1}^{n}\sum_{j=-\tau_M}^{-1}\sum_{m=k+j}^{k-1}\zeta^T(m)Z_i\zeta(m).$$

Then along the trajectory of system (7), we have

$$\Delta V_1 = x^T(k+1)Px(k+1) - x^T(k)Px(k) = x^T(k)[(A+\sum_{i=1}^{n}A_i)^T P(A+\sum_{i=1}^{n}A_i) - P]x(k)$$
$$+\sum_{i=1}^{n}\sum_{m=k-\tau_i(k)}^{k-1}[-2x^T(k)(A+\sum_{i=1}^{n}A_i)^T PA_i\zeta(m)] + [\sum_{i=1}^{n}\sum_{m=k-\tau_i(k)}^{k-1}A_i\zeta(m)]^T P[\sum_{i=1}^{n}\sum_{m=k-\tau_i(k)}^{k-1}A_i\zeta(m)] \quad (8)$$

By using lemma and (5), the following inequality can be derived

$$\sum_{i=1}^{n}\sum_{m=k-\tau_i(k)}^{k-1}[-2x^T(k)(A+\sum_{i=1}^{n}A_i)^T P\sum_{i=1}^{n}A_i\zeta(m) \le n\tau_M x^T(k)Xx(k) + \sum_{i=1}^{n}\sum_{m=k-\tau_M}^{k}\zeta^T(m)Z_i\zeta(m)$$
$$+\sum_{i=1}^{n}\sum_{m=k-\tau_i(k)}^{k-1}2x^T(k)[Y-(A+\sum_{i=1}^{n}A_i)PA_i]\zeta(m) \quad (9)$$

In addition, we have

$$\Delta V_2 = \sum_{i=1}^{n}\sum_{j=k-\tau_i(k+1)+1}^{k}x^T(j)Q_i x(j) - \sum_{i=1}^{n}\sum_{j=k-\tau_i(k)}^{k-1}x^T(j)Q_i x(j) = x^T(k)(\sum_{i=1}^{n}Q_i)x(k) - \sum_{i=1}^{n}x^T(k-\tau_i(k))Q_i x(k-\tau_i(k))$$
$$+\sum_{i=1}^{n}\sum_{j=k-\tau_i(k+1)+1}^{k-1}x^T(j)Q_i x(j) - \sum_{i=1}^{n}\sum_{j=k-\tau_i(k)+1}^{k-1}x^T(j)Q_i x(j). \quad (10)$$

Note that

$$\sum_{i=1}^{n}\sum_{j=k-\tau_i(k+1)+1}^{k-1}x^T(j)Q_i x(j) = \sum_{i=1}^{n}\sum_{j=k-\tau_m+1}^{k-1}x^T(j)Q_i x(j) + \sum_{i=1}^{n}\sum_{j=k-\tau_i(k+1)+1}^{k-\tau_m}x^T(j)Q_i x(j)$$
$$\le \sum_{i=1}^{n}\sum_{j=k-\tau_i(k)+1}^{k-1}x^T(j)Q_i x(j) + \sum_{i=1}^{n}\sum_{j=k-\tau_M+1}^{k-\tau_m}x^T(j)Q_i x(j),$$

Then, we have

$$\Delta V_2 = x^T(k)(\sum_{i=1}^{n}Q_i)x(k) - \sum_{i=1}^{n}x^T(k-\tau_i(k))Q_i x(k-\tau_i(k)) + \sum_{i=1}^{n}\sum_{j=k-\tau_M+1}^{k-\tau_m}x^T(j)Q_i x(j), \quad (11)$$

$$\Delta V_3 = (\tau_M - \tau_m)\sum_{i=1}^{n}x^T(k)Q_i x(k) - \sum_{i=1}^{n}\sum_{j=k-\tau_M+1}^{k-\tau_m}x^T(j)Q_i x(j), \quad (12)$$

$$\Delta V_4 = \sum_{i=1}^{n}\sum_{j=-\tau_M}^{-1}[\zeta^T(k)Z_i\zeta(k) - \zeta^T(k+j)Z_i\zeta(k+j)] = \tau_M\zeta^T(k)(\sum_{i=1}^{n}Z_i)\zeta(k) - \sum_{i=1}^{n}\sum_{m=k-\tau_M}^{k-1}\zeta^T(m)Z_i\zeta(m). \quad (13)$$

From (8), (11), (12) and (13), we can derive the following result

$$\Delta V = \Delta V_1 + \Delta V_2 + \Delta V_3 + \Delta V_4 = x^T(k)[(A+\sum_{i=1}^{n}A_i)^T P(A+\sum_{i=1}^{n}A_i) - P + n\tau_M X]x(k) - \sum_{i=1}^{n}x^T(k-\tau_i(k))Q_i x(k-\tau_i(k))$$

$$+(\tau_M - \tau_m)\sum_{i=1}^{n}x^T(k)Q_i x(k) + \tau_M \zeta^T(k)(\sum_{i=1}^{n}Z_i)\zeta(k) + 2x^T(k)[Y-(A+\sum_{i=1}^{n}A_i)^T P(\sum_{i=1}^{n}A_i)]\sum_{i=1}^{n}\sum_{m=k-\tau_i(k)}^{k-1}\zeta(m)$$

$$\le \eta^T(k)\Theta\eta(k) \qquad \text{where } \eta^T(k) = [x(k), x(k-\tau_1(k)), \cdots, x(k-\tau_n(k))],$$

$$\Theta = \begin{bmatrix} \Gamma + \tau_M(A-I)^T(\sum_{i=1}^{n}Z_i)(A-I) & -Y+(A+\sum_{i=1}^{n}A_i)^T PA_1 & -Y+(A+\sum_{i=1}^{n}A_i)^T PA_2 & \cdots & -Y+(A+\sum_{i=1}^{n}A_i)^T PA_n \\ +(A+\sum_{i=1}^{n}A_i)^T P(A+\sum_{i=1}^{n}A_i) & +\tau_M(A-I)^T(\sum_{i=1}^{n}Z_i)A_1 & +\tau_M(A-I)^T(\sum_{i=1}^{n}Z_i)A_2 & & +\tau_M(A-I)^T(\sum_{i=1}^{n}Z_i)A_n \\ & A_1^T PA_1 - Q_1 & & & \\ * & +\tau_M A_1^T(\sum_{i=1}^{n}Z_i)A_1 & A_1^T PA_2 + \tau_M A_1^T(\sum_{i=1}^{n}Z_i)A_2 & \cdots & A_1^T PA_n + \tau_M A_1^T(\sum_{i=1}^{n}Z_i)A_n \\ & & A_2^T PA_2 - Q_2 & & \\ * & * & +\tau_M A_2^T(\sum_{i=1}^{n}Z_i)A_2 & \cdots & A_2^T PA_n + \tau_M A_2^T(\sum_{i=1}^{n}Z_i)A_n \\ * & * & * & \ddots & \vdots \\ & & & & A_n^T PA_n - Q_n \\ * & * & * & * & +\tau_M A_n^T(\sum_{i=1}^{n}Z_i)A_n \end{bmatrix}$$

By Schur complement, (4) guarantees $\Delta V < 0$ for all nonzero $x(k)$. So the unforced system (3) with time-varying multiple delays is asymptotically stable. This completes the proof.

4 Static Out-Feedback Stabilization

In this section, we will design a controller for system (1). Then, we can construct a static out-feedback controller as the following form.

$$u(k) = Dy(k) \tag{14}$$

Substituting the static out-feedback control law into system (1) results in the following closed-loop system

$$x(k+1) = \bar{A}x(k) + \sum_{i=1}^{n}\bar{A}_i x(k-\tau_i(k)), \tag{15}$$

where

$$\bar{A} = BDC, \bar{A}_i = BDC_i, i = 1, 2, \cdots, n. \tag{16}$$

Theorem 2. For system (1), a stabilizing static out-feedback controller described by (13) exists if there exist matrices $P = P^T > 0, X > 0, Y > 0$, $Q_i^T = Q_i > 0$, $Z_i > 0$ and $l \times m$ matrix D satisfying

$$\begin{bmatrix} \bar{\bar{\Phi}}_{11} & \bar{\bar{\Phi}}_{12} \\ * & \Phi_{22} \end{bmatrix} < 0, \tag{17}$$

$$\begin{bmatrix} X & Y \\ Y^T & Z_i \end{bmatrix} \ge 0, i = 1, 2, \cdots, n \tag{18}$$

where

$$\bar{\bar{\Phi}}_{11}=\begin{bmatrix}-S & & & \\ & -\tau_M^{-1}S_1 & & \\ & & \ddots & \\ & & & -\tau_M^{-1}S_n\end{bmatrix}, \tag{19}$$

$$\bar{\bar{\Phi}}_{12}=\begin{bmatrix}A+BDC & A_1+BDC_1 & \cdots & A_n+BDC_n \\ A+BDC-I & A_1+BDC_1-I & \cdots & A_n+BDC_n-I \\ \vdots & \vdots & \ddots & \vdots \\ A+BDC-I & A_1+BDC_1-I & \cdots & A_n+BDC_n-I\end{bmatrix}, \tag{20}$$

$$PS=I, Z_iS_i=I, i=1,2,\cdots,n\,. \tag{21}$$

Proof: Notice Theorem 1, the closed system (15) is asymptotically stable if there exist matrices $P=P^T>0,\ \ X>0, Y>0,\ Q_i^T=Q_i>0,\ Z_i>0$ satisfying

$$\begin{bmatrix}\Phi_{11} & \bar{\Phi}_{12} \\ * & \Phi_{22}\end{bmatrix}<0\,, \tag{22}$$

where

$$\bar{\Phi}_{12}=\begin{bmatrix}P(\bar{A}+\sum_{i=1}^{n}\bar{A}_i) & P\bar{A}_1 & \cdots & P\bar{A}_n \\ Z_1(\bar{A}-I) & Z_1\bar{A}_1 & \cdots & Z_1\bar{A}_n \\ \vdots & \vdots & \ddots & \vdots \\ Z_n(\bar{A}-I) & Z_n\bar{A}_1 & \cdots & Z_n\bar{A}_n\end{bmatrix}.$$

We multiply both sides of (21) by $diag\{P^{-1}, Z_1^{-1}, Z_2^{-1}, \cdots, Z_n^{-1}, I, I, \cdots, I\}$.

Defining $PS=I, Z_iS_i=I, i=1,2,\cdots,n$, (22) can be converted into the form described by (17).

Remark: The obtained conditions are not strict LMI. So, a method of solving nonconvex feasibility problem should be developed. We omit it here.

5 Conclusion

In this note, the problem of out-feedback stabilization for discrete system with time-varying multiple delays is considered. By using the stability theory of Lyapunov and technique of LMI, Delay-dependent results are derived. The stable criterion is formed in the form of LMI. But the conditions in Theorem 2 are not strict LMI.

Acknowledgements. This work is supported by Shandong Provincial Natural Science Foundation (ZR2009AL012) and Foundation for Education Department of Shandong (J08LJ65).

References

[1] Gao, H., Lam, J., Wang, C., Wang, Y.: Delay-dependent output-feedback stabiliastion of descrete-time systems with time-varying state delay. IEE Prov.-Control Theory Appl. 151, 691–698 (2004)
[2] Wang, Z., Huang, B.: Robust H-infinity observer design of linear state delayed systems with parametric uncertainty: the discrete-time case. Automatica 35, 1161–1167 (1999)
[3] Palhares, R.M., de Souza, C.E., Peres, P.L.D.: Robust H-infinity filtering for uncertain discrete-time state-delayed systems. IEEE Trans. signal Process. 49, 1696–1703 (2001)
[4] Esfhani, S., Perersen, I.R.: An LMI approach to output-feedback guaranteed cost control for uncertain time-delay systems. Int. J. Robust Nonlinear Control 10, 157–174 (2000)
[5] Ge, J., Frank, P., Lin, C.: H-infinity control via output feedback for state delayed systems. Automatica 64, 1–7 (1996)
[6] Yan, H.C., Huang, X.H., Wang, M., Zhang, H.: New delay-dependent stability criteria of uncertain linear systems with multiple time-varying delays. Chaos, Solitions and Fractals 37, 157–165 (2008)
[7] Lu, C.Y., Tsai, J.S.H., Su, T.J.: On improved delay-dependent robust stability criteria for uncertain systems with multiple-state delays. IEEE Transaction on Circuits and Syatems I:Fundamental Theory and Applications 49, 253–256 (2002)
[8] Yue, D., Lam, J., Ho, D.W.C.: Reliable H-infinite control of uncertain descriptor systems with multiple delays. IEE Prov.-Control Theory Appl. 150, 557–564 (2003)

An Improved OTP Authentication System and Its Application in Mobile Commerce

Ye Xi

Wenzhou Medical College, Wenzhou 325035, China
yexi2010@gmail.com

Abstract. E-commerce based on Internet is gradually becoming a new mode for commercial activities and the greatest misgiving of enterprises to conduct e-commerce is "safety". One-time password (OTP) scheme can help to solve password problems such as password conjecture and wiretapping, which occurs by repeatedly using the same password. The S/KEY one-time password system is one of most popular authentication schemes, however it cannot withstand small integer attack. This work illustrates as well as discusses of the S/KEY based authentication schemes with their pros and cons. This paper suggests an improved one-time password system based on bidirectional virtual authorization in mobile application systems. Our proposed Scheme can defend copy attacks, replay attacks, integer attack and Brute-Force Attack.

Keywords: E-commerce, OTP, Authentication, S/KEY, Integer attack, Mobile.

1 Introduction

The increased availability of computers and high-speed networks could make e-commerce a reality. People hope to carry out network purchasing and transaction, making use of the convenience of network. However the Safety has become the bottleneck of e-commerce development currently [1]. For safe and reliable information exchange between parties in open type network like Internet, it needs a procedure to verify that the other party requesting access is actual authentic user. Authentication can be based upon what someone has (a smart card, token, or ID card), what someone knows (a password or PIN), what someone is (fingerprint), or any combination of these [2].

Among such authentication protocols, static username and password is the traditional and most widely deployed solution for user authentication in enterprises even today. However, static passwords can be vulnerable to the following circumstances [3]:

Static password intercepted once by intruders can be potentially replayed later to gain access to system.

Static passwords are more vulnerable to guess, and offline dictionary attacks, etc.

Strong password policy measures (password length, selection, re-cycling time, etc.) make users to write down their passwords and hence leading to more vulnerable systems.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 615–624.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Users often forget passwords and therefore a significant cost accrues to the enterprises from the help desk support required for user password management.

The S/KEY one-time password system is one of most popular authentication schemes, however it can not withstand small integer attack. In this paper, we present an improved S/KEY system based on bidirectional virtual authorization, which can defend copy attacks, replay attacks, integer attack and Brute-Force Attack.

The remainder of this paper is organized as follows: in Section 2, we discuss related work. In Section 3, we review the S/KEY authentication scheme. We then propose our improved OTP scheme in Section 4 and in Section 5, we discuss the Application of our scheme in Mobile Commerce. Finally, we discuss the strengths and limitations of our proposed scheme through security analysis in Section 6 and conclude in Section 7.

2 Related Work

Dynamic one-time passwords are considered as a secure replacement for static passwords that offer strong authentication. In 1981, Lamport [4] first proposed a well-known hash-based password authentication scheme. But the scheme suffers high hash overhead and password resetting problems. Since then, several schemes[5-9] have been proposed to address this problem for achieving more functionality and efficiency.

The basic thought of OTP (one-time password) authentication is that: add some uncertain factors to the SPP (Secure Pass Phrase) of the users to carry out Hash calculation when logins the system so that the passwords, namely the OTPs are different from each other for each time. For example, we adopt SHA1 as the uni-directionally hashing function, and random number or timestamp as uncertain factors, so OTP=SHA1 (Username + SPP+ random number/timestamp), in which '+' means concatenation operation. In this way, user's SPP is not transmitted on network directly, thus increasing the safety during the login [10]. After receiving the OTP, the server end will check in the same way to verify the legality of users [11-12]. There are three prominent methods of implementing OTP based authentication solutions in enterprises [3].

Time-Synchronous Technique- in this technique, both the client and server will have synchronous time clocks and use an algorithm that generate OTP from synchronous times and any other inputs (SPP).

Asynchronous Challenge-Response Technique- In this technique, the application present a dynamically generated unique challenge to the users when they try to login to server. Client software use some crypto primitive to generate a unique password from the challenge and any other information (SPP) provided.

Event- Synchronous Technique – in this technique, both the client and server will typically have an identical initial seed (counter Value). Client generates an OTP from the initial seed and any other input and updates the seed (increment/decrement the counter). Server also generates the password for that instance using the seed and other inputs. If both passwords match, the server authenticates the user and updates the seed (increment/decrement the counter).

3 Review and Analysis of S/KEY Scheme

S/KEY one-time password authentication Scheme, the method first proposed in the paper "The S/KEY one-time password system" [13], authenticates password by classifying into three stages, registration, generating (log-in) and authentication in software method not hardware method using token.

3.1 Registration Period

A user selects a fixed SPP(Secure Pass Phrase) and a seed randomly, and calculates a dynamic password sequence with N factors: H^1(SPP//Seed), H^2(SPP//Seed),…, H^{N-1} (SPP//Seed), H^N(SPP//Seed), here , SPP//Seed means the connection result of SPP and seed, H^N(SPP//Seed) means the information summary after Hash calculation of SPP//Seed for N times. The following data shall be transmitted to the system mainframe at server end through the secure channel: UserID, seed, factor number N, and the checking factor H^N(SPP//Seed) for checking the first password of user. The user keeps the SPP and in consideration of safety, the system shall be initialized if the Seq(sequence number) of Hash function is less than or equal to 5 during practical application.

3.2 Password Generation Period

The S/KEY password system generates a procedure to connect seed and SPP. The sequence of one-time passwords P_i is produced by applying the uni-directionally Hash function multiple times. This is, the first one-way password is produced by running the client's processed SPP and Seed through the hash function some specified number of times, N-1.

$$P_1=H^{N-1}(SPP//Seed)$$

The next one-way password is generated y running the user's password through the hash function only N-2 time.

$$P_2=H^{N-2}(SPP//Seed)$$

In general, the formula is:

$$P_i=H^{N-i}(SPP//Seed)$$

An eavesdropper who has monitored the use of the one-time password P_i will not be able to generate the next one in the sequence (P_{i+1}) because it would require inverting the hash function. Without knowing the SPP that was the starting point of the function iterations, this can not be done.

3.3 Checking Period

Checking procedures of OTP after being iterated for Seq(Seq=N-i) times are as follows:

1) User at client end sends a request to the mainframe for login;
2) Mainframe at server end sends challenge information to user, including Seq and seed.

3) Based on the secure hash function, user inputs SPP and calculate the $OTP=H^{Seq}(SPP//Seed)$ according the Seq and seed in challenge information, and sends it to the mainframe.
4) The mainframe summarizes the OTP using the same Hash function, if the calculation result is the same as previously stored $OTP'=H^{Seq+1}(SPP//Seed)$, the authentication succeeds, and the OTP is stored for use next time; otherwise, the authentication fails, and the system refuses to respond to user's requests.

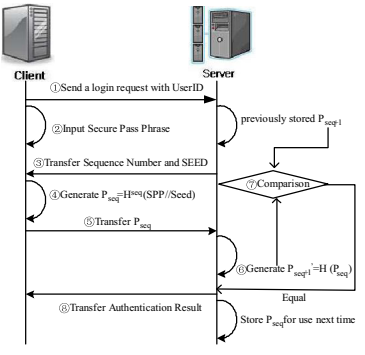


Fig. 1. Authentication procedure using S/KEY

Figure 1 shows the authentication procedure using S/KEY. From the above mentioned, we can know that the S/KEY plan has overcome the defects of traditional password identifying system, especially the replay attack problem that can not be solved by traditional system. For example, even if illegal listener-in obtains the password HSeq(SPP//Seed) after iterating for Seq times, he can not get the password HSeq-1(SPP//Seed) of the user next time, because H is a sage uni-directionally Hash function, thus withstanding replay attack effectively. In particular, S/KEY is a typical OTP System, in which SPP is neither stored in server or client computer (Users input SPP when they login), nor transmitted on the network, but only known to the user, thus the system is relatively safer.

3.4 Safety Loophole of S/KEY Scheme

Such S/KEY based authentication scheme can experience restriction of number of permissible authentications and has the limitation that system safety depends on one-way function. In addition, its Seed and iterated value are transmitted pellucidly, so small integer attack becomes a main safety loophole of the system. For example, when the hacker applies to the server for authentication, he imitates the user, intercepts the seed and iterated value from the server, alters the iterated value to a smaller one, and imitates the server to send the obtained seed and the iterated value to user. The user calculates the OTP, using the seed and smaller iterated value received from the hacker, sends it to the server, and the hacker intercepts the OTP again, calculates the OTP of larger iterated value with the help of known uni-directionally

hashing function, and obtains a succession of passwords of this user. Thus the hacker's imitation to legal users can not be perceived in a period of time, which is called small integer attack.

4 Proposed Scheme

Directed toward the loophole that S/KEY scheme can not withstand integer attack, in this section, we perform a comparison and analysis for three improved S/KEY authentication schemes which are: Yan's scheme, Kim's scheme and our proposed scheme.

4.1 Yan's Scheme

In 2008, Yan[14] proposed an improved S/KEY authentication scheme which can withstand the integer attack. In his theory, the following aspects have been improved: Keep the iterate time Seq at server end, as well as its backup at client end, making it alter with the server end. During the checking period, compare the iterated time Seq of the challenge information sent from server end with the stored iterated time Seq in the client end, if they are the same, calculate the OTP and send it to the server for checking, otherwise the system refuses to respond. This protects the users from small integer attack effectively by altering the iterated time Seq. When registering, the user has to initialize the iterated time Seq stored at client end to N-1, and descend 1 once login successes so that it alters with the Seq at server end.

Yan's proposal has the same problem that it is still similar with event- synchronous technique. As the Seq on client and server may drift (due to passwords generated by client but does not submitted, passwords submitted by client but does not reaches to server due to network failure, etc.), synchronizing the Seq value stored at client and server is a major challenge in this method.

4.2 Kim's Scheme

In the same year, Kim [15] proposed another improved S/KEY authentication scheme by using public key infrastructure. His scheme authenticates user through certificate, and since the password for the applicable session is encoded once again with the public key of service provider while the password for the applicable session is electronically signed by using the private key only known by user self, spoofing (integer) attack can be prevented beforehand. However, since it authenticates user and executes password authentication one more time, it has the shortcoming that authentication takes some more time. Therefore Kim's theory is suitable in the fields where security rather than time characteristics is emphasized such as internet banking, electronic payment and medical system. In addition, this technique is difficult to integrate with existing enterprise applications due to its complexity.

4.3 Our Proposed Scheme

In the S/KEY authentication scheme, seed and iterated value N are transmitted pellucidly, so the hacker can imitate the user, intercept the seed and iterated value

from the server , alter the iterated value to a smaller one, and imitate the server to send the obtain seed and the iterated value to user. Therefore, the root reason why the S/KEY authentication scheme can't withstand small integer attack is that the client can't authenticate the server. To solve this problem, we modify the checking procedures of S/KEY system as follows:

1) User at client end sends a request to the mainframe for login;
2) Mainframe at server end sends challenge information to user, including Seq, seed and previously stored OTP'. Here, OTP'= HSeq+1 (SPP//Seed).
3) User calculates the OTPSeq=HSeq (SPP//Seed) according the Seq and seed in challenge information and stored it firstly, and then user calculates OTPSeq+1= H1 (OTPSeq) =HSeq+1(SPP//Seed). If OTPSeq+1 = OTP', user authenticates the server and sends OTPSeq to the mainframe; Otherwise the system refuses to respond to the requests and prompts the message of "Under Small Integer Attack".
4) The mainframe summarizes the OTP using the same Hash function, if the calculation result is the same as previously stored OTP'=HSeq+1(SPP//Seed), the authentication succeeds, and the OTP is stored for use next time; otherwise, the authentication fails, and the system refuses to respond to user's requests.

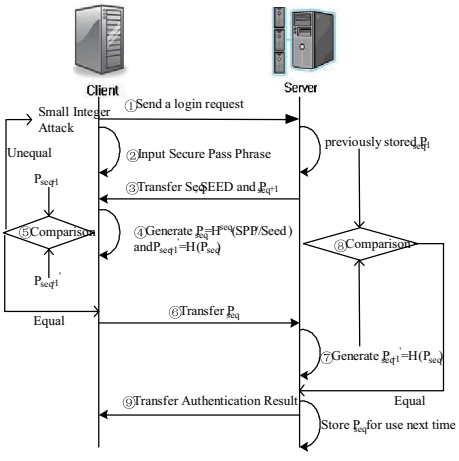


Fig. 2. Authentication procedure using our improved S/KEY scheme

Figure 2 shows the authentication procedure using our improved S/KEY scheme. When the hacker applies to the server for authentication, he imitates the user, intercepts the seed, iterated value(Seq)and previously stored OTP' from the server , alters the iterated value(Seq) to a smaller one(Seq-i), and imitates the server to send the obtained seed , Seq-i and OTP' to user. The user calculates $OTP^{Seq-i+1}=H^{Seq-i+1}$(SPP//Seed), using the seed and smaller iterated value (Seq-i) received from the hacker. Obviously, $OTP^{Seq-i+1}$ isn't the same as OTP', so the client end will know it

suffers small integer attack and refuse to respond to hacker's requests. A secure hash function is a function that is easy to computer in the forward direction, but computationally infeasible to invert, so the hacker can't calculate the correct $OTP^{Seq-i+1}$ according to OTP'.

5 The Application in Mobile Commerce

Figure 3 outlines the physical architecture of the improved system. In 2005, Vipul Goyal *et al.* [16] pointed two serious limitations of S/KEY scheme in mobile commerce:

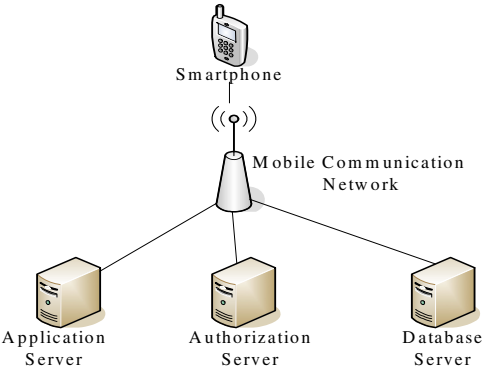


Fig. 3. The physical architecture of the improved system

1) As discussed before, N, the number of times a user can authenticate to the server, is finite (see 2 for why n cannot be made very large). Further, Alice is forced to choose a new password every time N reaches 1 and the system should be reinitialized. The old password cannot be reused again. This user unfriendly requirement may not be desirable in many environments.
2) The system is computationally intensive for the client especially when n is large. For example, with n=500, the client should compute 499 hash functions for authenticating first time, 498 for second time and so on. Hence with n=500, the client should compute about 250 hash functions per authentication on an average. So the scheme is unsuitable for mobile devices having low computational resources.

The high computation was a problem for mobile at that time, however with the rapid developments of smart phone technology, the computational ability of mobile device has improved significantly. In order to know the time a hash calculation would take, we run a program which computers 100000 hash functions (MD5) at a Desktop (Platform: Inter CPU T1300 1.66GHZ with 1GB memory) and the result is 10 second. Supposed that a user logins into system 10 times everyday and the system is only reinitialized once per year, then N=10*365=3650, the calculating time of N at desktop therefore is 10/100000*3650=0.365 s. The CPU frequency of latest smart phone is between 300MHZ and 1GHZ. Although the calculation ability of smart phone is

lower than desktop, providing that both of them have the same CPU frequency, the calculation time of smart phone should be still acceptable by user when N=3650.

6 Security Analysis

This section performs comparison and analysis for the items of traditional S/KEY authentication scheme, time synchronous authentication scheme, asynchronous challenge-response authentication scheme, event- synchronous authentication scheme, Yan's improved S/KEY authentication scheme, Kim's improved authentication scheme, our proposed scheme, integer attack , replay attack, time synchronization, sequence synchronization, computation cost , verification table and mutual authentication. The results of comparison and analysis on the proposed mechanism and existing algorithm are arranged and provided in Table 1 and 2, and it is notable that our proposed method is excellent in security aspect over other methods.

6.1 Adaptability for Small Integer Attack

Since S/KEY authentication scheme uses easily predictable N value and seed value, attacker can easily conjecture and can request authentication to service through this. However the improved scheme we presented in this paper can withstand the small integer attack very well because it realizes the bidirectional authentication between server and client end.

6.2 Adaptability for Reply Attack

Time using authentication scheme is exposed to reply attack intimidation by generating password by using the same representative value during certain period.

6.3 Time Synchronization

Authentication scheme that uses time requires time synchronized on certain reference between client and server. The method proposed in this paper doesn't require time synchronization.

6.4 Sequence Synchronization

In event- synchronous technique and Yan's improved S/KEY scheme, both the client and server will typically have an identical initial seed (counter Value), so they require sequence synchronization.

6.5 Computation Cost

Using public key algorithm will takes some more time to computer than using symmetry key algorithm and Hash algorithm, so the Kim's Scheme using public Key infrastructure takes more time than other schemes.

6.6 The Verification Table

Storing secret keys or passwords on a host increase its attractiveness as a target, and causes a breach of security to be more wide-spread. When a common password file is used for many machines, this risk becomes even greater. However, because the S/KEY techniques don't need to store SPP in server and client computer, no verification table will be required for these schemes.

6.7 Bidirectional Authentication

The traditional S/KEY scheme is one-way authentication, so it can't withstand the small integer attack. All improved S/KEY schemes in this paper are designed as bidirectional authentication and small integer attack can be prepared against.

Table 1. Comparison Evaluations

	traditional S/KEY	time synchronous	Challenge Response	Event synchronous	Yan's S/KEY	Kim's S/KEY	Proposed S/KEY
Adaptability for integer attack	LOW	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH
Adaptability for replay attack	HIGH	LOW	HIGH	HIGH	HIGH	HIGH	HIGH
Time synchronization	NO	YES	NO	NO	NO	NO	NO
Sequence synchronization	NO	NO	NO	YES	YES	NO	NO
Computation cost	LOW	VERY LOW	VERY LOW	VERY LOW	LOW	HIGH	LOW
Verification table	NO	YES	YES	YES	NO	YES	NO
Mutual authentication	NO	NO	NO	NO	YES	YES	YES

7 Conclusion

This paper analyzed the problem of vulnerability of different authentication schemes security, such as the traditional static password authentication scheme, S/KEY authentication scheme, Yan's improved S/KEY authentication and Kim's improved S/KEY authentication scheme. In an S/KEY system, SPP is neither stored in server nor client computer, nor transmitted on the network, but only known by the users; meanwhile OTP is transmitted for only one time on the network, thus the system is relatively safer. However the traditional S/KEY scheme is one-way authentication, so it can't withstand the small integer attack. As a result of our study, we proposed an improved one-time password scheme to solve this issue - a bidirectional authentication scheme which can prevent small integer attack in advance. The proposed scheme inherits the advantages of S/KEY OTP system, and

improves it in security. We also discussed the application of our scheme in Mobile Commerce. With the rapid developments of smart phone technology, the improved system model will more suitable for the authorization of mobile commerce.

Acknowledgement. This work was supported by "Wenzhou Science & Technology Bureau (Y20100301)."

References

[1] Qi, M.: E-business Practical Guide. Higher Education Press, Beijing (2000)
[2] Wu, T.: The Secure Remote Password Protocol. In: Internet Society Network and Distributed System Security Symposium, Sandiego (1998)
[3] Saxena, A.: Dynamic Authentication: Need than a Choice. In: 3rd International Conference on Communication Systems Software and Middleware and Workshops (2008)
[4] Lamport, L.: Password authentication with insecure communication. Communications of the ACM 24(11), 770–772 (1981)
[5] Chien, H.Y., Jan, J.K., Tseng, Y.M.: An efficient and practical solution to remote authentication:smart card. Computers and Security 21(4), 372–375 (2002)
[6] Hwang, M.S., Lee, C.C., Tang, Y.L.: A simple remote user authentication scheme. Mathematical and Computer Modelling 36, 103–107 (2002)
[7] Ku, W.-C., Chang, S.-T.: Impersonation attack on a dynamic id-based remote user authentication scheme using smart cards. IEICE Transactions on Communications E88-B (5), 2165–2167 (2005)
[8] Sun, H.M.: An efficient remote user authentication scheme using smart cards. IEEE Transactions on Consumer Electronics 46(4), 958–961 (2000)
[9] Sun, H.M.: Cryptanalysis of password authentication schemes with smart cards. In: Information Security Conference 2001, pp. 221–223 (May 2001)
[10] Hu, T., Liu, J., Chen, F., Sui, Z.: Principle and realization of OTP authentication system based on MD5. Information Technology 9, 140–142 (2005)
[11] Li, X., Tong, H.: Improvement of OTP Authentication. Microcomputer Applications 23, 52–54 (2007)
[12] Li, F., Liu, P., Ju, H.: Realization of a New Scheme of Improvement on OTP Authentication Technology. Computer Systems & Applications 4, 34–36 (2003)
[13] Haller, N.M.: The S/Key One-Time Password System. In: Proc. Internet Society Symposium on Network and Distributed System Security, pp.151–158 (1994)
[14] Li, Y.: Research on e-business identity authentication system based on improved one-time password. In: 2008 International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM (2008)
[15] Kim, H.-C., Lee, H.-W., Lee, K.-S., Jun, M.-S.: A Design of One-Time Password Mechanism using Public Key Infrastructure. In: Proceedings of 4th International Conference on Networked Computing and Advanced Information Management, NCM 2008, vol. 1, pp. 18–24 (2008)
[16] Goyal, V., Abraham, A., Sanyal, S., Sang, Y.H.: The N/R one time password system. In: International Conference on Information Technology: Coding and Computing (2005)

Research on ESC Hydraulic Control Unit Property and Pressure Estimation

Liu Wei, Ding Haitao*, and Guo Konghui

State Key Libratory of Automobile Dynamic Simulation, Jilin University, Changchun 130025, China
dht_jlu@126.com

Abstract. Vehicle mounted Electronic Stability Control (ESC) requires exact estimation of wheel brake cylinder pressure. A model of pressure transfer property is established according to the structure of ESC Hydraulic Control Unit (HCU). The property model of HCU is established for pressure estimation algorithm. The property is acquired with the ESC hardware in loop (HiL) test bed, to calibrate the model parameters. The algorithm is then used in the closed loop control of HCU. The test is carried out to verify the effect of pressure estimation method. The test result demonstrates that the algorithm can supply stable pressure information to the ESC system.

Keywords: Electronic Stability Control, Hydraulic Control Unit, Pressure estimation.

1 Introduction

ESC system plays an important role on research of vehicle active safety. ESC logic distributes the braking force to control the gesture of vehicle, in order to improve the handling performance and safety of the vehicle. ESC has been widely used in vehicle electronic control. FMVSS126 which published in April 2007 is the first regulation in the world about ESC, it regulates that, since Sep 1st 2011, all car sold in US with the load below 4,536Kg（10,000pounds） must mount the ESC system [1].

HCU is the actuator of ESC, the ESC ECU acquires vehicle states information through sensor and CAN bus, then distributes braking force according to control logic. The HCU controls four wheel cylinders pressure according to independent up-keep-down instructions of four wheel cylinders. Hence, the research on HCU dynamic property provides the possibility to wheel cylinder pressure estimation, which saves the mounting of wheel pressure sensor, saves the costs of manufacturing and improves the stability of the vehicle [2].

HCU model is established with some simplifications to describe the property of hydraulic dynamic characteristic. The ESC HiL is used to acquire the property of HCU hydraulic system, which is analyzed for identification of HCU model parameters. The pressure estimation algorithm based on the HCU model is applied in

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 625–631.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

the closed-loop control logic which is used for verification of pressure estimation algorithm [3].

2 ESC HiL Test-Bed and ESC HCU

2.1 ESC HiL Test-Bed

An ESC HiL test-bed is built for braking system research to avoid the complicate hydraulic system modeling. The states of hardware parts are measured by sensor and communicated with HOST PC and ECU on CAN bus.

The hardware of ESC HiL test-bed system includes braking system of vehicle with ESC HCU and pressure sensor mounted to measure pressure of main cylinder and wheel cylinder. The braking pedal is driven with braking-by-wire system to provide main cylinder pressure. The scheme of test-bed is demonstrated in Fig. 1.

In the braking-by-wire system, main cylinder is pushed with linear actuator which is driven by motor controlled by electronic brake pedal, to provide braking force. ESC ECU adjusts HCU to realize the required pressure of each wheel. The vehicle model is running in HOST PC to provide software environment for control logic. The monitor PC is used to modify and download control logic or pressure estimation algorithm into ESC ECU.

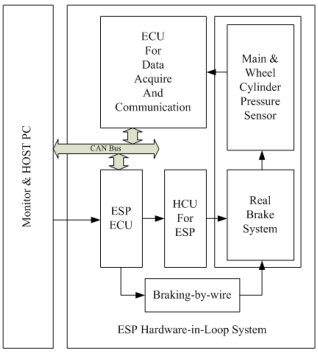


Fig. 1. The scheme for ESC HiL

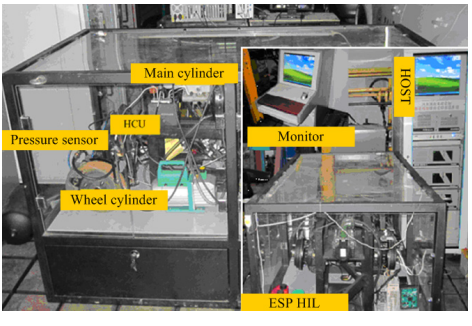


Fig. 2. Test bed for ESC HiL

Independent instructions of four wheels according to wheel cylinder required pressure is acquired with the calculation of veDYNA vehicle model running in simulink xpc-target, then communicated with ESC ECU on CAN bus, the ESC ECU executes to adjusts the wheel cylinder pressure and acquires the pressure sensor information which is send to HOST PC on CAN bus for analyzing.

As mentioned before, the test-bed is able to acquire the property of ESC HCU during pressure increase process and decrease process, which is used to identify the parameters of HCU model. Fig. 2 shows the structure and components of ESC HiL test-bed.

2.2 The Structure of ESC HCU

ESC HCU which is able to control four wheel cylinders pressure independently is the actuator of ESC, it connects two main cylinder and four wheel cylinder with an ingenuity designed hydraulic system, which distributes braking force to wheel cylinder based on main cylinder pressure (the reflection of driving purpose) and control logic.[5][6]

The hydraulic system is designed with cross-pipe which means that FR/RL and FL/RR wheel cylinders separately share the same pressure establishing route of ESC HCU. The hydraulic system scheme is shown with one wheel cylinder route in Fig.3.

In valve group 1 for pressure establishing, the main valve is always ON, the overflow valve by-pass constrains the maximum pressure which can be established. Valve group 2 for pressure establishing is always OFF. Valve group 3 for pressure maintaining is always ON, the one-way valve ensures that cylinder pressure is less than main pipe route pressure. Valve 4 for pressure draining is always OFF.

The motor drives the piston pump to pump the brake fluid in the accumulator, in order to ensure pressure draining rapidly. At the same time, it pumps the brake fluid in main cylinder to establish pressure in main pipe route for ESC mode.

ESC HCU is the actuator of ESC, it adjusts the wheel cylinder pressure according to ABS logic, when ESC logic works it establishes the pressure in main pipe route for wheel cylinder to create braking force actively.

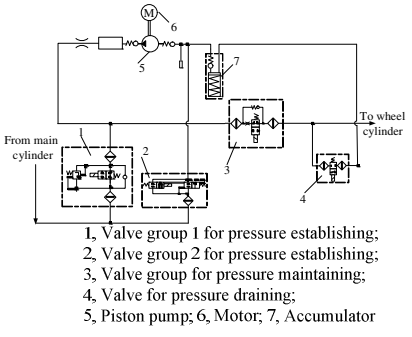


Fig. 3. The scheme of ESC HCU hydraulic system for one route

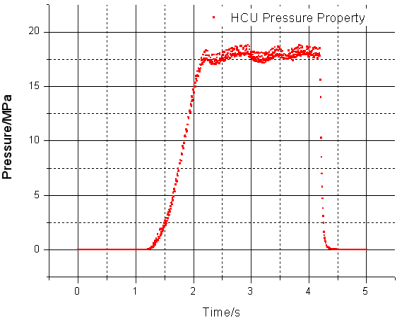


Fig. 4. The pressure property of ESC HCU

3 The Model of HCU Property

3.1 ESC HCU Model

According to the requirement of actuator in ESC system, ESC HCU is designed to be a hydraulic system with critical fabrication requirement, good response characteristic and stable dynamic characteristic. The HCU can be exactly modeled by reasonable test and calibration, the model is then used to describe the dynamic response of HCU, which is the base of cylinder pressure estimation [7].

The HCU property as shown in Fig. 4 includes: the delay of valve response, the transfer characteristic of hydraulic system, the pressure variation property of HCU in different pressure variation stage. The property is described by hydraulic dynamic model with parameters to be identified.

The modeling of HCU property is divided into two main processes, which are pressure increase process and pressure decrease process. The identification of HCU property also consists of two groups of parameters according to the HCU model.

We simplifying the system by neglecting the temperature impact of brake fluid viscosity, considering that brake fluid flow quantity is low during braking, the inner wall of hydraulic pipe is smooth. The model is established with the following process: ignoring pressure loss in pipes and parts, ignoring transient surge of brake fluid during valve changing, ignoring the elastic deformation of brake pipe and cylinder, ignoring the impact of temperature to brake fluid viscosity. These simplifications maintain the basic and main characteristic of HCU hydraulic system, at the same time, make the modeling process succinctly which ultimately simplifies the pressure estimation algorithm used in real time control logic.

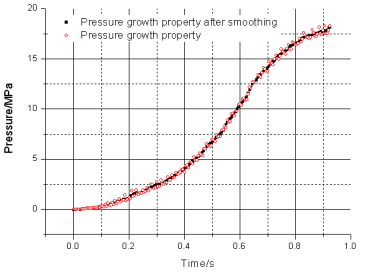


Fig. 5. The pressure increase property ESC HCU

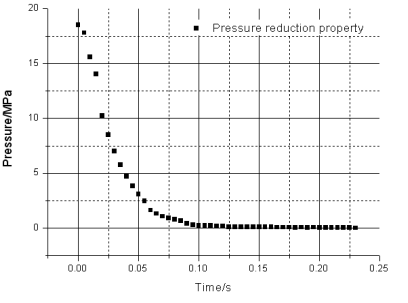


Fig. 6. The pressure decreaseproperty of ESC HCU

The pressure property after simplifying, Pm is pressure of main pipe route. Pw is pressure of wheel cylinder:

Pressure increase process:

$$\frac{dP_w}{dt} = \frac{1}{C_e R_e}(P_m - P_w)^{\phi} \tag{1}$$

- Ce: equivalent liquid capacity in pressure increase process;
- Re: equivalent liquid resistance in pressure increase process;
- Φ: throttling index in pressure increase process.

Pressure decrease process:

$$\frac{dP_w}{dt} = -\frac{1}{C_e R'_e}(P_w - P_r)^{\phi'} \tag{2}$$

- C'e: equivalent liquid capacity in pressure decrease process;
- R'e: equivalent liquid resistance in pressure decrease process;
- Φ': throttling index in pressure decrease process.

3.2 Parameters Identification of ESC HCU

The nonlinear relevant relationship between P_w and t is acquired by solving the states equations of the pressure increase and decrease processes. Through variable substitution, the nonlinear regression problem is then transferred to linear regression problem for system parameters identifying.

Pressure increase process:

$$(P_m - P_w(t))^{1-\phi} = (P_m - P_w(t_0))^{1-\phi} - \frac{1}{C_e R_e}(1-\phi)(t-t_0) \tag{3}$$

Pm = 18.32 is set by the overflow valve by-pass of valve group 1. According to the analysis of test data, we set Φ = 0.5 when correlation coefficient is 0.9963, identify the parameter $1/(C_eR_e)$ = 12.952 with least squares criterion. The inertial characteristic of brake system is also considered during modeling process.

Pressure decrease process:

$$(P_w(t) - P_\gamma)^{1-\phi'} = (P_w(t_0) - P_\gamma)^{1-\phi'} - \frac{1}{C_e R_e'}(1-\phi')(t-t_0) \tag{4}$$

Pγ is the pressure parameter of accumulator, we set Pγ = 0. According to the analysis of test data, we set Φ' = 0.9 when correlation coefficient is 0.9945, identify the parameter 1/(C'eR'e) = 43.14 with least squares criterion.

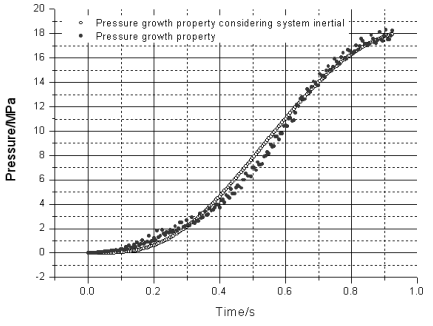


Fig. 7. The pressure increase property after parameter identifying

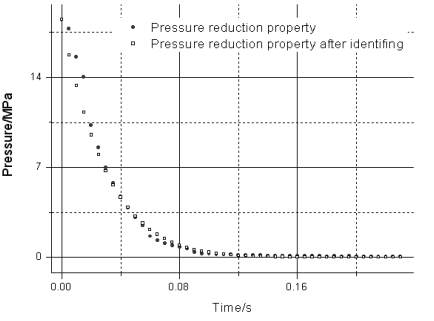


Fig. 8. The pressure decrease property after parameter identifing

After identification of parameters in each process, the complete HCU model is established for pressure estimation algorithm research.

4 Research on Pressure Estimation

4.1 Open-Loop Pressure Estimation

The performance of open-loop pressure estimation is guaranteed by the accuracy of HCU model. The algorithm for pressure estimation is also based on the HCU model. According to HCU property, the relationship between increment of cylinder pressure and cylinder pressure is acquired:

$$\Delta \mathrm{p} = \Delta \mathrm{p}(p_w) \tag{5}$$

In the control logic of HCU, the pressure of next control period is estimated by following discrete state function：

$$p_w(k+1) = p_w(k) + \Delta \mathrm{p}\left(p_w(k)\right) \tag{6}$$

4.2 The Closed-Loop Control Based on Pressure Estimation

The closed-loop control based on pressure estimation is achieved by the deference between estimated pressure and required pressure.

In ESC HiL system, HOST PC calculates the value of required pressure for each wheel by the veDYNA vehicle model and ESC logic, and sends required pressure to ECU, which uses the open-loop pressure estimation algorithm to decide the control instructions of each wheel cylinder.

Signals of pressure sensors are acquired by ECU and sent to HOST PC on CAN bus for data analysis.

4.3 Test Result

Two typical types of pressure control processes are applied with the ESC HiL test-bed to test the pressure estimation algorithm. One is that required pressure is square signal, the other is that required pressure is step up and step down signal.

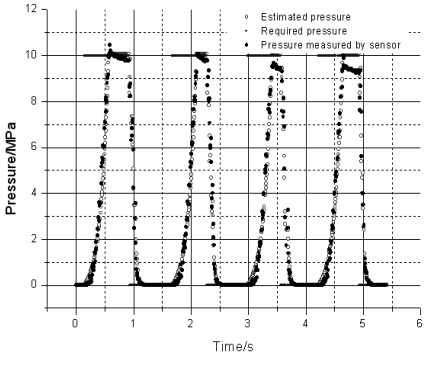


Fig. 9. The result when the required pressure is square signal

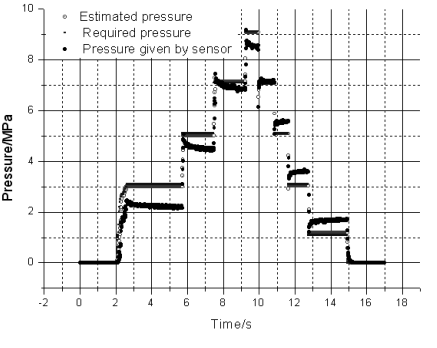


Fig. 10. The result when the required pressure is step signal

The required pressure is given by HOST PC, and is sent to ESC ECU, which controls the HCU with closed loop control based on pressure estimation. The wheel cylinder pressure created by HCU is measured by pressure sensor mounted in HiL. So the estimated pressure is acquired during the pressure control process, the required pressure is calculated by HOST PC, the wheel cylinder pressure measured by sensor is sent to ESC ECU. Through the CAN bus, the estimated pressure and pressure measured by sensor is also acquired by HOST PC for data analysis.

The maximum required pressure of square signal is 10MPa. As shown in Fig. 9, the estimated pressure follows the pressure given by sensor well, that means the closed-loop control logic based on open-loop pressure estimation algorithm describes the basic dynamic characteristic of hydraulic system.

In the step signal process, the required pressure increases step-by-step to maximum pressure , then decreases step-by-step to 0.The result is shown in Fig. 10.

5 Summary

According to the results which show the compare between estimated pressure and sensor tested pressure based on required pressure, the closed-loop pressure control logic based on the open-loop pressure estimation acquired by identifying the property of HCU is proved to be reliable, the open-loop pressure estimation provides the ESC with stable pressure state during the period when HCU works.

References

[1] Liu, W.: Research on Nonlinear State Estimation of Side-slip Angle for Vehicle Electronic Stability Program. CAE of Jilin University (2007)

[2] O'Dea, K.: Anti-Lock braking performance andhydraulic brake pressure estimation. SAE Paper 2005-01-1061

[3] Zhu, B.: Hydraulic brake pressure estimation for integrated chassis control. Journal of Jilin University (Engineering and Technology Edition) 39(sup. 1) (2009)

[4] Chen, C.-K., Dao, T.-K., Lo, M.-F.: A compensated-yaw-moment-based vehicle stability controller. In: Chinese Control and Decision Conference, CCDC 2008, July 2-4, pp. 892–897 (2008)

[5] Kusano, A., Kuno, T.: Hydraulic Brake System for Vehicles. Advics Co. Ltd., US Patent 6,709,072 (2004)

[6] Leppek, K.G., Hogan, M.A., Spadafora, P.J., Lee, A.J.: Antilock brake system with motor current controlof the pressure modulator. US Patent 5,102,207 (1990)

[7] Dawei, G.: Research on Hydraulic Modulator Modeling and Break Pressure Estimation for Vehicle Electronic Stability Program. CAE of Jilin University (2007)

The Emotional Robot Model Based on Endocrine System

Huang Guorui*

New Star Research Institute of Applied Technology, Hefei, 230031, China
huangdav@ustc.edu.cn

Abstract. Inspired by the regulation mechanism of endocrine system, we attempts to build an emotional robot model with the purpose to achieve the self-adaptation of robot behaviour in a complex and changeable environment. The model designs an internal environment and emotion- & motive-generating models based upon the endocrine regulation mechanism. The robot relies upon the dynamic stability of the internal environment to accomplish its self-organization to the external environment. Emotion-generating module is applied to in vivo and in vitro contingencies of various kinds; by releasing hormones, the emotion generated influences the robot's in vivo and in vitro perception, and then its behavioural choice. The motive-generating model is employed to maintain the relative stability of the internal environment, and the motive generated leads to the corresponding behavioural choice directly. In the end, the paper conducts detailedly the simulation experiment of robot planning, whose results prove the effectiveness of this emotional robot model.

Keywords: endocrine system, emotion, motive, robot control model.

1 Introduction

Drawing inspiration from work in Endocrinology and Neuroscience, we know that the organism's relatively stable *in vivo* environment is an enabling prerequisite for its adaption to the dynamic changing environment, as well as the basis for all its intelligent activities[1]. In recent years, researchers in artificial intelligence have also laid considerable stress on modelling the robot's internal state; however, instead of studying biological mechanisms to build effective intelligent models, in most cases, they have used the theory of finite state automata (FSA) in model building. Therefore, the results they have achieved are not ideal, and the robot's self-adaptive capacity is relatively poor. In maintaining the relative stability of the *in vivo* environment, the endocrine system plays a major role through opportune release of specific hormones. In a biological system, hormones can be seen as representing a low-bandwidth global communication protocol, which plays an irreplaceable role in regulating individual biological behaviours. The hormones released by the endocrine system provide a global state of the organism, which makes some behaviours easier to stimulate. And the implementation of these behaviours initiates the gradual transition of the organism's *in vivo* environment from instability to relative stability so that it can adapt to the dynamically changing environment.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 633–645.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

In the field of artificial intelligence, self-adaption of the robot's behaviour has been a long-standing subject, whereas biological information processing (BIP) a new research focus in recent years. How to combine the studies on the two subjects and devise a more effective self-adaptive robot model thus becomes a challenging task for researchers. Previous studies on BIP mechanisms has achieved fruitful results on the theory & models of the information processing mechanisms of the brain system, the genetic system and the immune system, as well as on their application[2-4]. The endocrine system is an important part of the BIP system, and the BIP mechanism via which it functions, because of its great theoretical significance and academic value to the development of new intelligent computation technology, is attracting more and more attention from researchers[5-6]. As a matter of fact, biological behaviour is regulated not only by the nervous system, the genetic system and the immune system, but also by the endocrine system.

Based upon these understandings, this paper draws upon the distributed regulation mechanism of the endocrine system to propose a new self-adaptive robot model. In this model, first, an emotion is generated via the emotion-generating subsystem according to the information perceived concerning the internal and external environments; then, the emotion stimulates the endocrine subsystem to generate corresponding hormones, which regulate the parameters of the internal environment and maintain its stability; and in response to the changes of the parameters and the external stimuli, the motive-generating subsystem computes the parameters concerning the motive; and in the end, according to the results of the computation, the behaviour-choosing subsystem chooses corresponding behaviours. To verify the effectiveness of the model, we have conducted a simulation experiment of robot motion planning. The results show that this model can effectively accomplish the robot behavioural decision-making and that it has a strong self-adaptive capacity.

2 A Brief Description of the Biological Endocrine System

The biological endocrine system is composed of endocrine glands, endocrine cells, and the hormones they release. In response to appropriate stimuli, some endocrine cells of the organism may generate specific hormones of certain amounts, which work with the nervous system to influence the organism's behaviour and then to maintain the relative stability of its in vivo environment. The endocrine system is a typical closed-loop control system. Thanks to the aid of the closed control loop, the central nervous system may receive various kinds of information (audio, light, temperature, smell, etc.) from the external environment; through the hypothalamus, it links the endocrine system to the external environment, thereby forming an open loop. It is the close cooperation between the endocrine system and the nervous system that enables the organism to adapt better to environmental changes.

3 Emotional Robot Model

3.1 The Structure of Emotional Robot Model

The emotional robot is made up of seven modules, i.e. internal sensors, external sensors, the emotion-generating module, the endocrine module, the motive-generating

module, the behaviour-choosing module, and the behaviour-implementing module. The relations between the robot's component modules are as follows: the internal and external sensors transmit the information perceived to the emotion generator, which generate a corresponding emotion; simulated by the emotion, the endocrine generator produces certain hormones, which in turn regulate the perception of the internal and external sensors. The motive generating module generates the current motive chiefly according to the perception of the external and internal sensors, which influence the strength of the motive. The behaviour choosing module chooses the behaviour that best fulfils the strongest motive. The behaviour implementing module implements the behaviour chosen, whose consequences change the internal and external environments. The following sections of this paper will describe the major components of the modules in detail.

3.2 Settings of the Internal Environment

The settings of the internal environment are not fixed. Generally speaking, it should be configured according to different control needs. In the experiment of robot planning described in the 4th section of this paper, as the expected objective is to enable the robot to survive longer in the dynamically changing environment, the settings are all concerned with the biological parameters related to its survival, which may be classified into five types i.e. blood sugar, water, energy, adrenalin and pain.

Every parameter of the internal environment has an ideal value, which represents the ideal state of the robot's internal environment. And every parameter has a range of variation. Its fluctuation cannot exceed the range, otherwise the robot will die. In addition, the robot also has a special parameter: organic matter, which increases after an intake of food and decreases after it is attacked. When the level of the organic matter drops to 0, the robot will die. Initially, all the parameters are set to their ideal values.

3.3 Emotion-Generating Module

3.3.1 Introduction of Artificial Emotion

Due to the extremely complicated nature of its working mechanism, our understanding of emotion is far from thorough. Generally, it is believed that, of the complex emotions of various kinds, some basic ones are regarded as fundamental to the organism's self-adaption. Drawing upon some biological theories and some existing models of artificial emotion, we have introduced a group of basic emotions such as fear, anger, sadness, joy, and disgust to construct our affective models.

The introduction of the emotional mechanism into the robot model in this paper may be based upon the following reasons: firstly, studies in neuroendocrinology has shown that the common regulatory effect of emotion and the endocrine system is the prerequisite for higher organisms to make right behavioural decisions; secondly, emotion may make timely response to changes in the internal and external environment and transmit messages concerning the changes to other modules in an efficient and rapid way; thirdly, the introduction of the emotional mechanism may avoid the model-building complexity when the endocrine module receives and handles what has been perceived from the external and internal environments, and at

the same time, ensures better proximity to the real biological information handling mechanism; last but not least, the emotional mechanism conducts preliminary processing of what has been perceived from the external and internal environments, and synthesizes the complicated information of various kinds so that the information can be directly used by other modules--therefore, it has greater flexibility and adaptability.

3.3.2 Introduction of Artificial Emotion

The cause of every emotion is different. Generally speaking, the causes can be classified into the following types: 1) externally perceived information—when the robot perceives the presence of an enemy or more around him, its emotion generating module will generate the sentiment of 'fear', whose intensity is associated with the number and distance of the enemies; 2) internally perceived information —when the internal sensor perceives that some parameters of the internal environment (pain, adrenalin, energy, etc.) have persistently stayed at a high level, the emotion generating module will generate the sentiment of 'anger', whose intensity increases progressively to the size and lasting time of the value of the parameter; 3) the outcome of the robot's own behaviour--if its own behaviour meets the desired objective, the sentiment of 'joy' will be generated, whose intensity is in direct proportion to the proximity to the desired objective; and 4) behaviours of other organisms—for example, when the robot is attacked by other organisms, the emotion generating module will generate the sentiments of ‘anger’ and ‘fear’, whose intensity is in direction proportion to the pain caused by the attack.

With the passage of time, the intensity of emotion will weaken gradually. Let’s use parameter $1-\rho$ to indicate the weakening degree of emotion. With every time step, the intensity of the robot's emotions will be adjusted according to formula (1).

$$\begin{aligned} &\tau_{ei}(t+1) = \rho * \tau_{ei}(t) + \Delta\tau_{ei} \qquad \rho \in (0,1) \\ &\Delta\tau_{ei} = \sum_{k=1}^{n} \Delta\tau_{ei}^{k} \end{aligned} \tag{1}$$

$\tau_{ei}(t)$ stands for the intensity of emotion i at moment t. $\Delta\tau_{ei}^{k}$ stands for the intensity of emotion i generated by the kth predisposing factor. $\Delta\tau_{ei}$ stands for the sum of the intensities of emotions generated by all the predisposing factors. The equation used to calculate $\Delta\tau_{ei}^{k}$ varies according to the difference of emotion i and predisposing factor k. Now, let’s take the sentiment of 'joy' for example to illustrate this briefly.

Let’s use $e2$ to symbolize the sentiment of ‘joy’, i.e. i=2. There are two stimulating factors for $e2$: 1) the behaviour achieves the desired objective; and 2) most of the internal parameters stay at ideal levels persistently. The criterion or criteria to judge if the desired objective is achieved rest upon whether the motive formerly stimulated has lowered in intensity. If it has, then the motive has been fulfilled. The fulfilment of the motive is in direct proportion to the decrease in stimulating intensity. When judging whether the internal environment remain at ideal level, we classify the parameters of the internal environment into four groups, corresponding respectively with the four levels i.e. low, ideal, high and common. If an internal parameter stays at the ideal level persistently, emotion $e2$ will be generated.

Now let's take $\Delta\tau_{e2}^2$ for example and discuss how to calculate the intensity of the newly generated emotion or sentiment.

The procedures may be described as follows:

(1) Judge whether each internal parameter j is maintained at an ideal level, if formula (2) is used to calculate stimulus intensity $Stimus_j$ corresponding with internal parameter j. Otherwise $Stimus_j$ is equal to 0.

(2) Employ formula (3) to calculate $Stimus$, or the sum of stimulus intensities corresponding with all the parameters, and count the number of the stimulus intensities that are larger than 0, i.e. num.

(3) Employ formula (4) to calculate $\deg ree$, or the degree of the internal environment's deviation from the ideal state.

(4) Calculate $\Delta\tau_{e2}^2$ with formula (5).

$$Stimus_j = \begin{cases} 0 & Step_j < MINSTEP \\ \sqrt{\dfrac{Step_j - MINSTEP}{MAXSTEP - MINSTEP}} & MINSTEP \le Step_j \le MAXSTEP \\ 1 & Step_j > MAXSTEP \end{cases} \tag{2}$$

$$Stimus = \sum_{j=1}^{PARAINUM} Stimus_j \tag{3}$$

$$\deg ree = stimus * \frac{num}{PARAINUM^2} \tag{4}$$

$$\Delta\tau_{e2}^2 = \begin{cases} MAXEC * \deg ree * \sqrt{1 - \tau_{e2}(t)/MAXEV} & \text{if } \tau_{e2}(t) \le MAXEV \\ 0 & \text{else} \end{cases} \tag{5}$$

$Step_j$ in formula (2) stands for the number of time steps throughout which internal parameter j stays at the ideal level. And $MAXSTEP$ and $MINSTEP$ symbolize respectively the maximum and minimum of $Step_j$. In formulae (3) and (4), $PARAINUM$ stands for the number of the internal parameters, and num for the number of those internal parameters whose intensities, $Stimus_j$, are above 0. In formula (5), $MAXEC$ symbolizes the maximum intensity of an emotion generated by a certain emotion-generating factor, whereas $MAXEV$ the maximum value of emotional intensity.

What is described above is the detailed calculation procedure of generating the sentiment of 'joy' according to the stability of the internal environment. As the procedures for other sentiments are similar to this one, they will not be described here because of the limited space.

3.4 Endocrine Module

3.4.1 Model Building for Hormone Generation

The main function of the endocrine system is to produce hormones in response to sentiments or emotions. And every sentiment corresponds different hormone. The amount of the hormone(s) produced as a result of the stimulation by a certain sentiment (stimulus) is related to the intensity of the sentiment, and to the density of the hormone(s) in question—the denser they are, the less they are produced in response to the same external stimulation.

Let's use parameter $1-\sigma$ to denote the degree of hormone disappearance. With every time step, the densities of all the hormones hl in the robot's endocrine system may be updated according to formula (6):

$$\begin{aligned} &C_{hl}(\mathrm{t}+1) = \sigma * C_{hl}(\mathrm{t}) + \Delta C_{hl} \qquad \sigma \in (0,1) \\ &\Delta C_{hl} = \sum_{i=1}^{\mathrm{m}} \Delta C_{hl}^{i} \end{aligned} \tag{6}$$

$C_{hl}(t)$ denotes the density of hormone l at moment t, ΔC_{hl} the density of the hormone newly generated in this cycle in response to the stimulation of all sentiments associated with hormone l, and ΔC_{hl}^{i} the density of the hormone l produced in this cycle due to sentiment i. The formula used to calculate ΔC_{hl}^{i} may be expressed as:

$$\Delta C_{hl}^{i} = \frac{\tau_{ei}(t+1)}{MAXEV} * \alpha_{il} * MAXHC * \sqrt{1 - C_{hl}(t) / MAXHV} \tag{7}$$

In formula (7), $MAXHC$ refers to the maximum variation of hormone density in a single cycle, $MAXHV$ to the maximum hormone density, and α_{il} to the degree of association between sentiment i and hormone l. From this formula, it can be seen that the density of the hormone(s) generated by a sentiment, one the one hand, increases as the intensity of the sentiment increases, and one the other, decreases as the density of the hormone increases.

3.4.2 Model Building for Hormone Regulation

Hormone regulation is the major means by which the high organisms adapt themselves to the external environment and maintain the stability of the internal environment. It is generally known that the endocrine system exerts distributed regulation by producing hormones. Drawing upon this mechanism, this paper factors in the following three aspects when it attempts to build the model for hormonal regulation.

(1) Regulate the perception of internal sensor(s)

The function of the hormone is singular. Every hormone has effect selectively on some physical parameters, which correspond with certain hormone acceptors. When the internal environment is configured above, a total of five internal parameters are involved. Assume that, for every internal parameter, there exists a pair of independent hormones, which correspond respectively with the increase and decrease of the

parameter perceived by the internal sensor. In other words, the hormone can regulate the sensitivity of the inner sensor's perception of the inner environment.

The regulatory effect of hormones on the internal sensor manifests itself in a certain amount of increase or decrease in what the internal sensor perceives. The amount of increase or decrease grows progressively as the hormone in question becomes denser. This process may be expressed in the following formula.

$$\Delta S_{ipn}^{l} = \beta_{\ln} * \sqrt{C_{hl}(t+1)/MAXHV} * MAXIPC \tag{8}$$

In formula (8), ΔS_{ipn}^{l} stands for the amount of hormone *l*-induced increase or decrease in what has been perceived by the internal sensor denoted as n; $\beta_{\ln}$ for the regulatory relationship between hormone l and internal sensor n, whose value is either 1 or -1. If $\beta_{\ln}$ =1, then the outcome of the sensor's perception has witnessed a certain amount of positive gain; when $\beta_{\ln}$ =-1, it has suffered a certain amount of loss. $MAXIPC$ refers to the maximum amount of variation caused to what the sensor perceives in a single cycle.

(2) directly changes the values of internal parameters

Apart from regulating the inner sensor's perception, hormones also modify the values of the parameters in the internal environment. The amount of modification may be calculated with the following formula i.e. formula (9).

$$\Delta S_{ipn}^{l} = \beta_{\ln} * \sqrt{C_{hl}(t+1)/MAXHV} * MAXIPC * Scale \tag{9}$$

In formula (9), $Scale$ represent a constant larger than 0 but smaller than 1. From the formula it can be seen that the modifications of the internal parameters induced by hormones are of no effect to the current cycle, because formula (9) uses $C_{hl}(t+1)$ rather than $C_{hl}(t)$, which is more conducive to the stability of the internal environment.

(3) Regulate the perception of the external sensor

Only hormone GABA is capable of regulating the external sensor; that is to say, the external sensor is subject solely to GABA. Stimulated by the sentiment of sadness or boredom, the endocrine system will produce GABA. As its density increases, the robot's vigilance will reduce, and as a sign of it, the variance of errors in external perception will be on the rise; otherwise, the robot's vigilance will increase, and the variance of error in external perception will be on the decline. The relationship between the variance of error in external perception and the density of GABA can be described in the following formula.

$$D(X) = MIND * (1 + dist / MAXDIST + (C_{h10}(t+1)/MAXHV)^2) \tag{10}$$

In the formula, X stands for the external sensor's error, a random variable which is assumed in this paper to be subject to a normal distribution; and $D(X)$ for the variance of X. $MIND$ represents the minimum value of the variance of error, $dist$ the external sensor's perceptible range, $MAXDIST$ the maximum perceptible range, and $C_{h10}(t+1)$ the density of GABA. It can be derived from formula (10) that as the density of GABA rises, $D(X)$ will increase and $D(X)$ will decrease contrarily.

3.5 Motive-Generating Module

3.5.1 Introduction of Motive

Viewing from the perspective of biology, the stability of the organism's internal environment underlies all intelligent activities, as well as its survival and growth. The state of the motive reflects that of the internal environment. Therefore, using the motive as the basis of behaviour choice can effectively achieves the objective of maintaining the robot's internal environment and adapting to the external environment. The motive is chiefly determined by the deviation or difference between the internal parameter perceived and its ideal value; and it is susceptible to the information concerning the external environment. If the robot's water content is lower than its ideal value, then the motive of 'thirst' arises. If the robot senses the existence of water nearby, the motive will be further strengthened. We have set altogether five motives: thirst, hunger, fatigue, self-defence, and aggressiveness.

In this paper, the robot always chooses the behaviour that best fulfils its currently strongest motive in order that the internal parameter that deviates from the ideal value can be restored and the internal environment can thus be maintained stable.

3.5.2 Construction of the Motive Model

The motive can be seen as underlying all processes that maintain the in vivo balance and keep all the internal parameters within certain ranges. As internal parameters are constantly changing, the strength of the motive is mainly determined by the fluctuation of the internal values. Therefore, when we calculate the strengths of the motives in the current cycle, we need not take into consideration the strengths of the motives in the previous cycle. This has differentiated our motive model from other motive models[7] built on principles of cognitive science. Whereas these models have considerable 'memory' as to its strength, our model puts memory on the level of physical parameter.

In the cycles, the strength of the robot's motive p can be calculated with formula (11).

$$\psi_{mp}(t+1) = \sum_{n=1}^{PARAINUM} \psi_{mp}^{n}(t+1) + \sum_{r=1}^{PARAONUM} \psi_{mp}^{r}(t+1) \tag{11}$$

In the formula (11), $PARAONUM$ stands for the number of the types of external sensors. The first part of the equation represents the influence of the internally perceived information on the strength of the motive, whereas the second denotes the influence of the externally perceived information on the strength of the motive. With regard to the influence of the internal perception, we should mainly consider the deviations of the internal parameters from their ideal values and the associations between these parameters and the motive. As for the influence of the external perception, we should add a constant to the strength of the corresponding motive. Being restricted by space, we will leave out the formulae calculating $\psi_{mp}^{n}(t+1)$ and $\psi_{mp}^{r}(t+1)$.

After we derive the strengths of the motives, it is likely that multiple motives are activated simultaneously. However, only the strongest will be responded to by the robot. The robot will adjust its behaviour to fulfil the motive.

4 Simulation Experiment and Its Result

This paper builds a two-dimensional grid-like virtual environment, creates two types of artificial life i.e. robot and enemy thorough programming, and conducts simulation experiments. Now, we will introduce the design of the experiments and conduct some tentative analysis of the experiment results.

4.1 Settings of the Virtual Environment

The robot and its enemies live in a two-dimensional world of 30×30 squares. Every square may be either occupied by an object or empty. With reference to the internal environment set up above, we have set a total of 5 different types of objects, which are food, water, rock, robot and enemy. Every object has four exterior attributes--they are luminance, hardness, smell, and organic matter. These attributes may be recognised by the robot or its enemies; and by these attributes, they may make judgements concerning the type of the object in each square.

4.1.1 Enemies

Enemies may be seen as a low organism. The sole purpose of their introduction is to add some dynamic to the chequered environment. They have only one physical parameter—pain—and they are immortal. Their behavioural pattern is simple. Under normal circumstances (in which they feel no pain), they always move in a direction they randomly choose and attempt to devour whatever they encounter. When the object they encounter is the robot, the robot will feel that it is under attack, that its organic matter is reducing, and that its pain is intensifying. And at the same time, the enemy itself also feels slight pain. When the object the enemy runs into happens to be another enemy, both the parties will feel pain, and the pain felt by the one under attack will be more intense. When the object is a rock, the enemy will feel pain, and the pain it feels will be stronger than when it attempts to devour the robot or another enemy. When the enemy is attacked by another enemy or the robot, it will feel pain too, and the pain it feels will be stronger than when it is the attacker. If the enemy's pain parameter is larger than 0, it will choose the behaviour 'escape', i.e. retreat to the opposite direction. The farther the retreat, the less pain it will feel. And it will retreat until it feels no pain. In order to avoid being trapped between two objects, the second time it has to escape, the enemy will attempt to move sideways.

4.1.2 Robot

Now let's build emotional robot model proposed above. In specific, this model-building process includes setting the internal environment, emotion generation, hormone product-ion, hormone regulation, and motive generation. In the following part, let's discuss the robot's internal and external perception, behaviour settings, and the algorithm of its behaviour planning.

(1) Internal sensors

Internal sensors can collect information concerning the robot's internal environment. Every internal parameter corresponds with an internal sensor. What internal sensors perceive is regulated by the hormones generated and the degree of the regulation can be calculated with formulae (8) and (9). Therefore, the information got by the internal sensors concerning the internal environment is not completely consistent with the real situation. There exists some deviation. The hormonal regulation of the perception of the internal sensors is a major part of the endocrine regulation mechanism.

Besides, the internal sensors have a 'memory' capacity, which enables them to remember the general picture of the internal parameters within a certain time span. We group the values of the internal parameters into three ranges: high, low and ideal. The robot is capable of remembering the number of time steps some internal parameters stays in a certain range.

(2) External sensors

As configured, the robot has three types of external sensors--visual, olfactory and tactile. These external sensors can sense the four external attributes of the object in any square within a certain range, and with the method of pattern recognition, judge what it is in the square.

The method of pattern recognition adopted in this paper involves comparing values perceived and the standard values in every pattern and selecting those with the most similarity as the results. The robot's external sensors have a certain degree of perception errors, which are roughly subject to the normal distribution with 0 at the centre. The variance of this normal distribution grows as the perceived distance increases, and decreases as the robot's vigilance strengthens. The robot's vigilance is regulated by hormone GABA. And the degree of its regulation can be calculated with formula (10).

(3) Settings of robot behaviours

The robot may conduct six behaviours: drinking water, taking food, rest, attack, escape, and strolling. The correspondence of the behaviours with the current motive, and the consequences and prerequisites of the behaviours, is shown in Table 1.

The algorithm of robot behaviour planning proposed in this paper may be described as follows:

Step 1: Set the default values of emotional intensities and hormone densities, t=0.

Step 2: Set the stimulated degrees of all the motives to 0.

Step 3: Probe the internal and external environments, and calculate $\tau_{ei}(t+1)$, or the intensity of emotion i at moment t+1 according to formula (1).

Step 4: Stimulated by the emotions or sentiments, the endocrine module produces various hormones; then calculate C_{hl}, or the density of hormone l at moment t+1.

Step 5: The hormones produced regulate the results perceived in Step 3; then calculate the robot's eventual internal and external perception at moment t+1 according to formulae (8), (9) and (10).

Step 6: Calculated the stimulated degrees of the robot's motives according to formula (11) and select the one with the greatest degree.

Step 7: Conduct behaviour choice and fulfil the motive selected.

Step 8: Implement the behaviour selected, which influences the robot's internal and external environments.

Step 9: If all the internal parameters are within the prescribed ranges, jump to Step 2; otherwise, output the number of time steps taken by the robot, and end the simulation.

Table 1. Settings of robot behaviours

Behaviour	Currently stimulated motive	Consequence(s) of behaviour	Prerequisite of behaviour
Drink water	thirst	increases water content	water in at least one of the four neighbouring squares
Take food	hunger	increases blood sugar content	food in at least one of the four neighbouring squares
Rest	fatigue	increases energy	rock in at least one of the four neighbouring squares
Attack	aggressiveness	ncrease pain, reduces water content, blood sugar, energy and adrenalin	at least one enemy in the four neighbouring squares
Escape	self-defence	relieve pain; reduces water content, blood sugar, energy and adrenalin	Of the four neighbouring squares, there is at least one empty
Strolling	when the conditions for the currently selected behaviour cannot be fulfilled	reduces water content, blood sugar, energy and adrenalin	Of the four neighbouring squares there is at least one empty

4.2 Experiment Results and Discussion

We conduct four independent groups of experiments, each of which is carried out ten times. Then we get the means of the ten times. The final results are shown in Table 2. The common settings for all the groups of experiments are: 1 robot, 10 enemies, 100 shares of water, 100 shares of food, and 100 rocks. The initial positions of the 311 objects are generated at random. Each object occupies a grid square; but there should not be more than one object in one square. Both the robot and the enemies can move in the grid world, whereas the positions of the rocks remain unchanged throughout the experiments. The robot and every enemy are materialised with one thread. The planned objective for the robot is to move in the grid world for as long as possible until it dies. The result of every experiment is the number of the time steps taken by the robot before it dies.

The results of the experiments above show that the emotional robot model is effective; the robot based upon this model can survive longer in the dynamically changing environment. It is also found in the experiments that, in different emotion states, the robot’s behavioural strategies undergo corresponding changes. When the

robot spots an enemy nearby, whether it will choose to escape or to attack is directly influenced by its current sentiments. It is this high emotional-regulating capacity that enables the robot not only to maintain the dynamic stability of its internal environment, but also to respond to contingencies in the internal and external environments in a timely way.

Table 2. Simulation results

Experiment sequence number	Whether emotions are generated	whether food and water consumed are regenerated	maximum number of time steps	average number of time steps
1	Yes	Yes	2358	1986
2	Yes	No	1454	827
3	No	Yes	589	465
4	No	No	373	209

5 Conclusion

Drawing upon the regulating mechanism of the organism's endocrine system for the internal environment and nervous system, this paper proposes an emotional robot model. In this model, first, an emotion is generated via the emotion-generating module according to the information perceived concerning the internal and external environments; then, the emotion stimulates the endocrine module to generate corresponding hormones, which regulate the parameters of the internal environment and maintain its stability; and in response to the changes of the parameters and the external stimuli, the motive-generating module computes the values related to the motive; and in the end, according to the results of the computation, the behaviour-choosing module chooses corresponding behaviours. The results show that this model can effectively accomplish the robot behavioural decision-making and that it has a strong self-adaptive capacity.

References

[1] Shen, W.M., Salemi, B., Will, P.: Hormone-inspired adaptive communication and distributed control for CONRO Self-reconfigurable robots. IEEE Transaction on Robotics and Automation 18(5), 700–712 (2002)

[2] Ma, C., Yue, X.: Intelligent Model of Urban Road Tunnel Ventilation System Based on Multi-Level Neural Network. In: Pacific-Asia Conference on Circuits, Communications and Systems, pp. 636–639 (2009)

[3] Yang, Y., Yu, X.: Cooperative Coevolutionary Genetic Algorithm for Digital IIR Filter Design. IEEE Trans. on Industrial Electronics 54, 1311–1318 (2007)

[4] Dasgupta, D.: Advances in artificial immune systems. IEEE Trans. on Computational Intelligence Magazine 1, 40–49 (2006)

[5] Timmis, J., Neal, M., Thorniley, J.: An Adaptive Neuro-Endocrine System for Robotic Systems. In: IEEE Workshop on Robotic Intelligence in Informationally Structured Space, pp. 129–136 (2009)
[6] Mendao, M.: A Neuro-Endocrine Control Architecture Applied to Mobile Robotics. PhD Thesis, University of Kent, Canterbury, UK (2007)
[7] Ganapathy, V., Yun, S.C., Lui, W.L.D.: Utilization of Webots and Khepera II as a platform for Neural Q-Learning controllers. In: IEEE Symposium on Industrial Electronics & Applications, pp. 783–788 (2009)

Research on the New Channel Estimation Algorithm for MIMO-OFDM System Based on Training Sequence

Zhigang Yang

School of Computer, Chongqing University of Arts and Sciences, YongChuan, ChongQing 402160, Chian
ayzg163@163.com

Abstract. Supporting high mobility will be an important character of the future wireless communication systems. It challenges the channel estimation technique a lot and there are two tough problems in front of us, they are multi-path fading channel and bandwidth efficiency. Orthogonal Frequency Division Multiplexing (OFDM) technique changed the frequency selective multi-path fading channel into flat fading channel in frequency domain which effectively mitigates the effects of multi-path propagation and, hence, increase data rate. In this paper, we discuss the principle of MIMO-OFDM systems based on the introduction of fading channels. Furthermore, the channel estimation technique in MIMO-OFDM system is investigated, with the emphasis on training-assisted channel estimation method. Firstly, we summarize and analyses the exiting channel estimation methods in MIMO-OFDM system. Secondly, we proposed a new channel estimation method for MIMO-OFDM system based on training sequence. A novel pilot pattern with higher frequency efficiency is established and we use a reasonable manner to insert pilot so that the system greatly reduced the training cycle, increased data rates. Finally, in order to track the time-varying channel, an improved direct decision method is further adopted. Theoretical analysis and numerical results show that the new channel estimation scheme can offer a good performance and a high ability to track the time-varying channel.

Keywords: channel estimation, MIMO-OFDM, training sequence, new algorithm.

1 Introduction

In recent years, wireless communication technology has made high-speed development and it provides a relatively low cost, high-quality services. With the development of wireless communication technology, new service generates accordingly. In the near future, wireless personal communications and multimedia communications will become the new requirements. To achieve these goals, the need for future wireless communications is to provide a very high transfer rate, and high-quality wide range of services including data, voice, pictures and even video.

Orthogonal Frequency Division Multiplexing (OFDM) technique changed the frequency selective multi-path fading channel into flat fading channel in frequency

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 647–655.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

domain which effectively mitigates the effects of multi-path propagation and, hence, increase data rate. In additional, Multiple Input Multiple Output (MIMO) technique can generate independently parallel multi-channel data stream in spatial at the same time, which directly and effectively increase the transmission efficiency. Therefore, combining MIMO and OFDM is believed to have the ability to achieve two purposes, the one is high transmission speed, the other is strongly reliability. Accurate channel estimation is crucial for the performance of wireless communication systems.

The channel estimation methods can be divided into three categories, they are non-blind channel estimation, blind channel estimation and semi-blind channel estimation. The non-blind channel estimation generally is applicable to the continuous transmission or burst transmission system. The algorithm based on pilot symbols or trained sequence, the receiver achieves the initial estimate. When the system sends the useful information, the algorithm will use the results of the initial decision to update and complete the real-time channel estimation. The biggest different between non-blind and blind channel estimation is that the blind method is based on the limited characters of the transmitted information symbols and their statistical trait. So this kind of algorithm has the most frequency efficiency. Semi-blind channel estimation solves the problems of spectrum waste from none-blind channel estimation and high complication of blind channel estimation methods. Hence the semi-blind channel estimation algorithm for MMO-OFDM system is regarded to be a promising way for channel estimation.

In this paper, we discuss the principle of MIMO-OFDM systems based on the introduction of fading channels. Furthermore, the channel estimation technique in MIMO-OFDM system is investigated, with the emphasis on training-assisted channel estimation method. Firstly, we summarize and analyses the exiting channel estimation methods in MIMO-OFDM system. Secondly, we proposed a new channel estimation method for MIMO-OFDM system based on training sequence. A novel pilot pattern with higher frequency efficiency is established and we use a reasonable manner to insert pilot so that the system greatly reduced the training cycle, increased data rates. Finally, in order to track the time-varying channel, an improved direct decision method is further adopted. Theoretical analysis and numerical results show that the new channel estimation scheme can offer a good performance and a high ability to track the time-varying channel. The reminder of this paper is organized as follows: the next section summarize the related work on MIMO-OFDM system. In section 3, we present a new channel estimation method for MIMO-OFDM system based on training sequence. In section 4, an improved direct decision method is further adopted and we make the theoretical analysis and simulation experiment according to the above proposed algorithm. Finally, we draw our conclusion on section 5.

2 The Related Work

MIMO and OFDM for its effective anti-fading characteristics and high spectral efficiency by a widespread are concerned by many experts [1][2]. The combination of the two constitutes the MIMO-OFDM system, the technical complement each other so that wireless channel high-speed data transmission to achieve the most promising solutions, and have a very extensive research development prospects [3]. With the

continuous development of wireless communications, broadband multimedia services, the high-speed transmission, against frequency selective fading channel, and to resolve multi-path channel inter-symbol interference caused by the increasing demand, OFDM technology came into being [4]. OFDM in the frequency domain the channel is divided into a number of orthogonal sub-channels, sub-channel overlap between the spectrum, and reducing the sub-letter inter channel interference (ICI), to improve the spectrum efficiency [5]. Also, because each sub-channel signal bandwidth is much smaller than channel bandwidth, while the overall channel is not flat, that has a frequency selective, each sub-channel is relatively flat, and each sub-channel can be equivalent to a narrowband channel, thus greatly reducing the inter-symbol interference (ISI) [6].

MIMO technology can not only increase exponentially fading channels the system capacity, and if further with the combination of channel coding technology, can greatly improve the communication system performance. Space-time coding technique is developed on this basis a new coding and signal processing technology that will channel coding techniques combined with array processing technology to significantly improve the wireless communication system capacity and transmission rate, as wireless channel bandwidth problem solving provides a new way [7][8].

Some space-time coding theory was largely based on flat fading channel, while in practice, most wireless communication environments are fast fading case, the channel non-flat, then the system performance will drop dramatically, which makes space-time codes in the future of broadband wireless communications are extremely limited [9]. Since the merger of the OFDM technology, frequency selective fading channel into several parallel flat fading sub-channels, this system not only brings with space-time coding diversity gain and coding gain, while OFDM receiver equalizer have both the advantages of simple structure [10].

The OFDM system channel estimation has been widely studied. It often can be divided into three methods including based on training sequence method, based on pilot symbols method and blind estimation methods [11]. The study shows that these algorithms in OFDM systems can achieve better performance, but they do not apply to MIMO-OFDM system. This is because the MIMO-OFDM system uses multiple transmitting and receiving antennas, which receive signals from multiple transmit antennas send a signal fading and additive noise superimposed, the use of such algorithms to the channel estimate, for a given transmitting and receiving antenna pairs, signals from other antennas shall interfere with the signal 0dB noise power less than often, so a great deal of estimation error, resulting in a sharp decline in system performance, MIMO-OFDM systems channel estimation is a challenging and meaningful research. This article will focus on MIMO-OFDM system channel estimation study and training sequence based channel estimation algorithm as a focus of the study.

3 A New Channel Estimation Algorithm For MIMO-OFDM System Based on Training Sequence

3.1 Proposed Algorithm I

Channel estimation is essential a detection problem. Given a carrier signal X, assuming that the attacker knows the probability density function P_0 and the channel

capacity probability density function P_1, then the problem becomes into binary hypothesis testing, that is:

$$\begin{cases} H_0 : X \sim P_0 \\ H_1 : X \sim P_1 \end{cases}$$

This detection problem can be Bayesian estimation, minimum and maximize detection, Neyman-Pearson test etc. If the channel capacity has the existence of possible multiple distributions, then the test will be more complicated. Related entropy (Kullback-Leibler Divergence) is often used to discriminate the probability distribution. Defined as formula:

$$D(P_0 \| P_1) \triangleq \int_x P_0(x) \log \frac{P_0(x)}{P_1(x)} d_x$$

It is associated with the probability of error detection, the value of $D(P_0 \| P_1)$ is the greater, the smaller the probability of error. Thus, usually using information theory to define the security of MIMO-OFDM systems (detection capability) often takes the method. Only if:

$$D(P_0 \| P_1) = 0, i.e., P_0 \equiv P_1$$

Systems have an excellent safety performance. Under ideal conditions, following embedding of the message into the cover signal, the resulting channel capacity is required to have exactly the same probability distribution as the cover signal, and then no statistical test can reliably detect the presence of the hidden message. We refer to such MIMO-OFDM schemes as perfect secure. But in reality, MIMO-OFDM system and channel estimation are frequently faced with a dilemma that P_0 is unknown. Therefore, the MIMO-OFDM personnel may not be able to design a beautiful embedding $f(\bullet)$ function to get the perfect security features; accordingly channel estimation-analyst can not get a reliability detection performance.

In channel estimation test of this hypothesis, there are usually two types of errors: false alarm rate and false negative rate. The corresponding error probability is expressed as α and β. Usually these two types of errors are difficult to be estimated. The threshold method was given as:

$$\alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta} \le D(P_1 \| P_0)$$

This type of errors includes two types of threshold hypothesis testing related to the probability distribution of entropy: greater related entropy means the stronger detection capability. In order to make MIMO-OFDM system security, it needs to reduce the related entropy, or even make it to be zero to obtain a perfect secure MIMO-OFDM system. On the contrary, in order to design good channel estimation algorithm, we need to look for characteristics of the probability distribution of channel capacity. While on the signals (images, audio, etc.) modelling, recent studies have made great progress, but on a unified model of the signal has not been

established. However, given the situation that contains two kinds of channels; this problem can be solved through supervised learning approach. Therefore, MIMO-OFDM faces enormous challenges in order to avoid changing the statistical features of cover signal when secret message has been embedded; conversely, channel estimation is to seek the statistical difference between the feature vectors caused by information hiding.

In this framework I, we aim to analyze the difference of statistical distribution in wavelet domain; the whole idea takes supervised learning. After extracting the feature vector, classifier is used to the distinguish channels. The overall framework is as follows in Fig.1:

Let X and Y be the cover audio and video channel, respectively. Before the statistical analysis, the audio channel is divided into blocks with same length M. Assume that these three kinds of audio channel with same sample length L, then the total number k of segments is $\lfloor L/M \rfloor$. Then in training phase, we can get frame signals set $(X_1(n), X_2(n), \ldots, X_k(n))$, $(Y_1(n), Y_2(n), \ldots, Y_k(n))$, where $1 \le n \le M$. After the DWT operation on each frame signal, the detail sub-bands are extracted. Let $(XS_1(n), XS_2(n), \ldots, XS_k(n))$, where $1 \le n \le M/2$, represent the sub-band signal of cover audio and video channel, respectively. Then we choose the appropriate probability density function to model the statistical distribution of each sub-band and extract effective statistical features. The feature set V_X, V_Y for cover audio and video channel is $(XV_1, XV_2, \ldots, XV_k)$ and $(YV_1, YV_2, \ldots, YV_k)$. Finally both the feature sets are transported into classifier as to get training model. In the testing phase, the test audio file T will go through the same procedures of segmentation, DWT module, PDF module until the feature sets V_T is achieved, and then it is used to judge the test file type with training model. The core steps are PDF model and feature extraction which will be discussed in detail in the following.

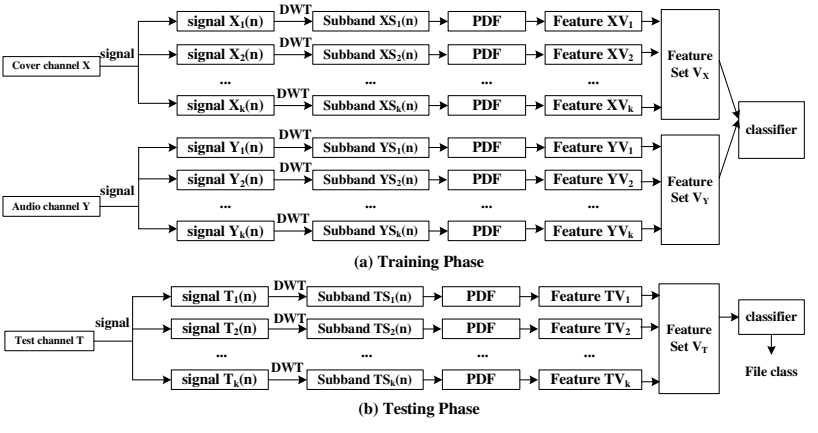


Fig. 1. Framework of proposed algorithm I

3.2 Proposed Algorithm II

The concept of using distortion measure to classify cover audio and audio channel has been introduced by Hamza Ozer, I. Avcibas . Their method attempts to find good features from the standard quality metrics which are designed to evaluate the perceptual and objective quality performance of audio. As the primary motivation for developing these quality metrics was for purposes other than system capability, the capability of distinguishing changes in quality due to embedding may be limited. Liu proposed the distortion metric based on Hausdorff distance. The distortion measurement was obtained at various wavelet decomposition levels from which the higher-order statistics was extracted as features for a classifier to determine the presence to hidden information. About the key issue of decomposition level and moment order which effect the performance directly, the research takes the empirical choosing, not giving the theoretic analysis in detail. In this algorithm, we propose an audio channel scheme that measures audio distortion using PDF distance metric. Given an audio object which could potentially be an audio channel object, we consider its de-noised version as an estimate of the cover-object. After appropriate segmentation, we apply wavelet decomposition to both and to generate wavelet coefficients at detail level of resolution. Next we use GMM and GGD to model the wavelet coefficients and calculate the probability density. Lastly distance metric is used as features to test the similarities of PDF between the audio channel and their de-noised versions.

Let and be the de-noised versions of cover audio and video channel capacity, respectively. Similar with the proposed algorithm I, we also do the procedures of segmentation, DWT and PDF model. Then we calculate the distance of PDF as feature set, denoted by $(\mathrm{D}_{x1},\ldots,\mathrm{D}_{xk})$ and $(\mathrm{D}_{y1},\ldots,\mathrm{D}_{yk})$ .Then the feature set V_X and V_Y are utilized to train the classifier. In the testing phase, the same feature extraction method is applied to the test audio T. The resulted feature set V_T is then used to distinguish audio channel type from the training model.

Here, we consider the ideal situation of wavelet coefficients. Let $X(n)$ denote a cover object and $Y(n)$ be its stego-version. Let $D(n)$ be the de-noised version of cover signal. As a good estimation of the cover channel, $D(n)$ is assumed to be Gaussian distribution, that is $D \sim N(\mu,\sigma_d^2)$ .The noise channel $S(n)$ and embedding data $W(n)$ are independent and identically distributed (i.i.d) Gaussian distribution, $S \sim N(0,\sigma_s^2)$, $W \sim N(0,\sigma_w^2)$ .Then the audio channel estimation can be expressed as $Y(n)=X(n)+W(n)$ with the additive model. So the clean audio and video channel estimation satisfy $X \sim N(\mu,\sigma_d^2+\sigma_s^2)$, $Y \sim N(\mu,\sigma_d^2+\sigma_s^2+\sigma_w^2)$ respectively.

A common way to measure the distance between two PDFs $p(x)$ and $p^{'}(x)$ often takes the Kullback-Liebler divergence. Noticing that $KL(p \| p^{'})$ is not necessarily equal to $KL(p^{'} \| p)$. Thus, it is more reasonable to use a symmetric version of the Kullback-Liebler divergence:

$SKL(p,p') = KL(p \| p') + KL(p' \| p)$ When both $p(x)$ and $p'(x)$ are Gaussian distributions, we can obtain:

$$SKL(p,p') = \sigma_p^2 / \sigma_{p'}^2 + \sigma_{p'}^2 / \sigma_p^2 + (\mu_p - \mu_{p'})^2 (1/\sigma_p^2 + 1/\sigma_{p'}^2)$$

So, we can get the KL distance between clean audio and its de-noised version, audio channel and its capacity estimation version respectively.

$$SKL(X,D) = (\sigma_d^2 + \sigma_s^2)/\sigma_d^2 + \sigma_d^2/(\sigma_d^2 + \sigma_s^2)$$

$$SKL(Y,D) = (\sigma_d^2 + \sigma_s^2 + \sigma_w^2)/\sigma_d^2 + \sigma_d^2/(\sigma_d^2 + \sigma_s^2 + \sigma_w^2)$$

So,

$$SKL(Y,D) - SKL(X,D) = \frac{\sigma_d^2\sigma_s^2\sigma_w^2 + (\sigma_w^2\sigma_s^2 + \sigma_w^4)(\sigma_d^2 + \sigma_s^2)}{\sigma_d^2(\sigma_d^2 + \sigma_s^2 + \sigma_w^2)(\sigma_d^2 + \sigma_s^2)}$$

It is very clear that $SKL(Y,D) > SKL(X,D)$, which verify the feasibility of our methods.

4 Improved Direct Decision Method and Simulation Experiment

As a dollar evolution of Gaussian probability density function, Gaussian mixture model (GMM) can be approximation of any probability density distribution of arbitrary shape, which is widely used in channel estimation. This paper also uses it to model the wavelet coefficients. In general, the Gaussian mixture distribution model can be the following using limited form of distribution and said:

$$f_k(x) = \sum_{j=1}^{k} \pi_j \phi(x, \theta_j)$$

Here, $\phi(x,\theta_j)$ is the j^{th} component of GMM model, θ_j is the vector of the mixture parameters which consists of weight π_j, mean μ_j, variance σ_j^2. The weight π_j must satisfy:

$$\pi_1 + \ldots + \pi_k = 1. \ \pi_j \geq 0$$

In this paper, after the details of DWT decomposition, we also use greedy EM algorithm to GMM modelling. Here, the sub-frame length is set to 1024 samples and the number of Gaussian components is set to 3. The largest number of iterations takes 20 times. After GMM training, three Gaussian components of parameters are $\pi_1 = 0.3859$, $\mu_1 = 0.0055$, $\sigma_1^2 = 3.3005e-006$; $\pi_2 = 0.3721$, $\mu_2 = -0.0054$, $\sigma_2^2 = 3.7409e-006$; $\pi_3 = 0.2690$, $\mu_3 = 0.0001$, $\sigma_3^2 = 1.2304e-005$.

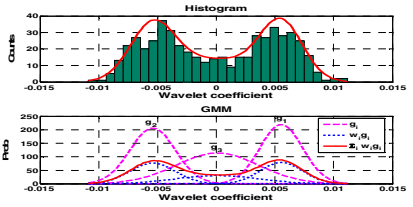


Fig. 2. Histogram of wavelet coefficients and GMM model

Figure 2 shows the histogram and simulated GMM model of wavelet coefficient. g_1, g_2, g_3 denotes the probability of GMM component with the weight w_1, w_2, w_3, respectively. $\sum_i w_i g_i$ is the weighted sum of component's probability. We can see that Gaussian mixture model can simulate the distribution of wavelet coefficients (the red line display it). Here we calculate the probability density function of GMM model with various embedding strengths α .

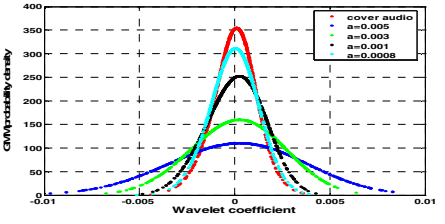


Fig. 3. Probability density function of GMM model

5 Conclusion

With the 4G era draws near, MIMO-OFDM technology is being more and more attentions. Many experts believe that this technology will become the core of next generation mobile communication systems technology. MIMO technology and the combination of OFDM technology can complement each other so as to get a good anti-fading characteristics and high spectral efficiency. OFDM technology in the frequency domain channel will be divided into several sub-channel, spectrum overlap and improve the utilization of the spectrum, but also to selective fading into flat fading in nature. MIMO technology can not only exponentially increase the system capacity fading channel but also can greatly improve the transmission efficiency of wireless communication systems, wireless channel bandwidth to resolve the issue of providing a new way.

Acknowledgements. This work is supported by Chongqing University of Arts and Sciences scientific research funds for talents introduction. The authors would like to thank Feng He professor of School of Communication and Information Engineering, Chongqing University of Posts and Telecommunication, for his thoughtful comments. This work supported by grant from the Central University Basic Research Operating

Expense-Young teachers funded projects. The project number: 3101014. This work also supported by grant from National Natural Science Foundation of China. The project number: 61170023.

References

[1] Piechocki, R.J., Nix, A.R., McGeehan, J.P.: Joint Semi-blind Detection and Channel Estimation in Space-Frequency Trellis Coded. MIMO-OFDM. In: Personal Mobile Communications Conference, vol. (22-25), pp. 101–105 (April 2003)

[2] Baek, M.-S., Kim, M.-J., You, Y.-H., Song, H.-K.: Semi-Blind Channel Estimation and PAR Reduction for MIMO-OFDM System with Multiple Antennas. IEEE Transactions on Broadcasting 50(4), 414–424 (2004)

[3] Yue, J., Kim, K.J., Reid, T., Gibson, J.D.: Joint Semi-Blind Channel Estimation and Data Detection for MIMO-OFDM Systems. In: IEEE 6th CAS Symp. on Emerging Technologies: Mobile and Wireless Comm., pp. 709–712 (2004)

[4] Suh, C., Hwang, C.-S., Choi, H.: Preamble design for channel estimation in MIMO-OFDM systems. In: IEEE Global Telecommunications Conference, GLOBECOM 2003, December 1-5, vol. 1, pp. 317–321 (2003)

[5] Li, Y.: Pilot-symbol-aided channel estimation for OFDM in wireless systems. IEEE Transactions on Vehicular Technology 49(4), 1207–1215 (2000)

[6] Gong, Y., Letaief, K.B.: Low Complexity Channel Estimation for Space-time Coded Wideband OFDM Systems. IEEE Trans. on Wireless Commune. 2(9), 876–882 (2003)

[7] Li, Y.G., Seshadri, N., Ariyavisitakul, S.: Channel Estimation for OFDM Systems with Transmitter Diversity in Mobile Wireless channels. IEEE J. Select Areas Commune. 17(3), 461–471 (1999)

[8] Moon, J.K., Choi, S.I.: Performance of channel Estimation for OFDM Systems in a Multipath Fading Channels. IEEE Trans. on Consumer Electronics 46(1), 161–170 (2000)

[9] Li, Y(G.), Winters, J.H., Sollenberger, N.R.: MIMO-OFDM for wireless communications: Signal detection with enhanced channel estimation. IEEE Trans., Commun. 50, 1471–1477 (2002)

[10] Li, Y., Seshadri, N., Ariyavisitakul, S.: Channel Estimation for OFDM Systems with Transmitter Diversity in Mobile Wireless Channels. IEEE Journal on Selected Areas in Communications 17(3), 461–471 (1999)

[11] Foschini, G.J., Gans, M.J.: On limits of wireless communications in a fading environment when using multiple antennas. Wireless Personal Communications (6) (1998)

The Research and Implementation of Technology of Generating Test Paper Based on Genetic Algorithm

Shao Ming-zhu*, Li Wei-feng , and Du Jing-yi

Dept. of Computer Science, Henan Mechanic and Electric Engineering College
Xinxiang, China
shaomingzhu79@163.com

Abstract. Automatic test paper generation is an important part of examination system. How to improve the test paper generating algorithm and the quality of test paper generating has become an important issue at present. Based on the researches on code tactic, adaptive function, controlling parameters, an improved genetic algorithm is used in the paper. Experiment shows that the improved genetic algorithm could compose test paper more efficiently than traditional algorithms.

Keywords: Automatic Test Paper Generation, Algorithm of Generating Paperr, Genetic Algorithm, Self-Adaptive Genetic Algorithm.

1 Introduction

With the quick development of computer technology and network technology, the way of test has made great change. Using resources of network, the test cycle time can be reduced and the efficiency of test work can be improved. To ensure the quality of test paper and enhance the scientificity and objectivity of the test work, it is vital to design logical automatic algorithm of test paper generation.

2 The Technology of Automatic Test Paper Generation

Automatic test paper generation [1] is to generate test paper taking the questions out of the questions pool by the system according to the requirements given by the teacher, such as question types, the difficulty and other requirements. The problem of automatic test paper generation is an objective optimizing problem with multiple restraints. The methods used commonly presently are mainly random sampling, backtracking and generic algorithm.

Generic algorithm[2] is an algorithm simulating natural selection and natural generic mechanism, which first generates initial population, then makes the individuals in it crossover and mutate to realize the recomposing of individuals' structure, and last, choosing and copying excellent individuals according to the

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 657–663.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

predetermined evaluating functions to compose the new generation, and thus repeating goes on, until the global optimal solution is got at last. Generic algorithm is characteristic of inherent parallelism, global optimality and quick convergence. Its seeking begins with many feasible solutions, then does iteration to bring new solutions using some rules and until gets the optimal results. It can solve the problems which need great computation work. That all makes it fit for solving the problem of the automatic test paper generation from questions pool.

2.1 Algorithm of Generating Test Paper Modeling

First, constitute the corresponding state space for each control index which is necessary in the course of generating test paper. We consider n-dimensional vector to decide a question (question value a_1, difficulty a_2, knowledge point a_3, cognizing level a_4, question type a_5, time appraisal a_6 … a_i), in which a_i corresponds to i index. Then, to determine a test paper need to determine an m*n matrix. In it, m represents the number of the questions in the test paper.

$$s_g = \begin{bmatrix} a_{11}a_{12} \ldots\ldots a_{1n} \\ a_{21}a_{22} \ldots\ldots a_{2n} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{m1}a_{m2} \ldots\ldots a_{mn} \end{bmatrix}$$

One certain row of S represents several attributes of one question, all in binary; whereas each line represents values of one attribute of all the questions in the test paper. In the generic algorithm of generating test paper designed in this paper, one individual is one question in a test paper, and so-called N-specie is an aggregate composed of N individuals, which is called population for short. N is the scale of the population, called N-population space. So called matrix is a group individuals (a_1, a_2... a_i), in which $a_i \in S$ (i=1, 2), and all the collection of the population is the matrix space.

The objective matrix should meet the following restraints condition:

① total score restraint of the test paper:

$$\sum_{i=1}^{m} a_{i1} = totalscore$$

② difficulty restraint of the test paper:

$$\sum_{i=1}^{m} a_{i1}a_{i2} / 100 = ND$$

(the average difficulty of the whole test paper)

The difficulty of a question is determined by the following expression:

Difficulty (d) = 1-(average score/full mark of the question)

Level of difficulty: very easy (d<0.2); easy (0.2-0.4); mean (0.4-0.6); difficult (0.6-0.8); very difficult (0.8-1.0)

Average difficulty of the test paper:

$$ND = \frac{\sum_{i=1}^{m} p_i d_i}{\sum_{i=1}^{m} p_i}, \quad (i=1, 2, 3\ldots\ldots m)$$

In it, m is the number of the questions in the test paper; p_i、 d_i respectively represent the score value and the difficulty of question i. The average difficulty of the test paper to generate is given.

③ the score value of each knowledge point:

$$\sum_{i=1}^{m} c_{1i} a_{i1} = Z_h$$

In it,

$$C_{1i} = \begin{cases} 1, a_{i3} = h \\ 0, a_{i3} \neq h \end{cases}$$

Z_h is the score value of knowledge point h, h is the knowledge point restraint. Commonly, we hope the test paper could cover all the knowledge ever taught, so each knowledge point (can be divided by cognition level or section) should be covered and the score value of each knowledge point should be in direct proportion to the class hours for it.

④ score value of questions on cognition level k

$$\sum_{i=1}^{m} C_{2i} a_{i1} = P_k$$

In it,

$$C_{2i} = \begin{cases} 1, a_{i4} = k \\ 0, a_{i4} \neq k \end{cases}$$

P_k is the score value of all the questions on cognition level k. Cognition level can be given the following value: knowing, understanding, mastering and practicing, and so on. And the score rate of each level should has is given by user.

⑤ score value of question type d;

$$\sum_{i=1}^{m} C_{3i} a_{i1} = R_d$$

In it,

$$C_{3i} = \begin{cases} 1, a_{i5}=k \\ 0, a_{i5}\neq k \end{cases}$$

R_d is the score value of all the questions which belong to type d, in which d is the question type restraint. Question type could be: multiple-choice, completion, T or F and short answer. And the score of each type is given by user.

⑥Estimated Time needed to finish question i (t_i), Estimated Time to finish the test paper (T)

$$T=\sum_{i=1}^{m} t_i \quad (i=1, 2 \ldots m),$$

In it, m is the number of the questions contained in the test paper, t_i is the estimated time needed to finish question i. (Time for the test paper is given by user.) That is the restraint of test time requirement.

To meet some special requirements, besides the six monomial models above, other indexes can be appointed such as Expected Value, degree of differentiation, etc. Students Scores roughly shows normal distribution, so Expected Value is equivalent to the average student score and the system makes the score value of questions of different difficulty in the test paper follow normal distribution too, whereby to control the average score is to change the difficulty of the test paper, and to generate test paper according to the expected value given by user is as well. However, in terms of the writer's experience, it is better not to have too many restraints, because too many indexes will increase difficulty and decrease efficiency. Therefore, the course of generating test paper is a course to get an optimal solution with multiple objectives and the solution is not unique.

2.2 Design the Algorithm

The basic process of genetic algorithm is: to ascertain code scheme, to initialize population, to ascertain adaptive function, to ascertain selection tactics, to choose the controlling parameters, to design the genetic operator, to ascertain the terminated rule of algorithm.

(1) Code scheme

The paper adopts binary code and independent code tactic. To be specific, it is to code for each question type and then use traditional code tactic to cope with each question type, meanwhile code of the group and other group is independent. Hereinto each group represents one question type, and the length of each group is determined by the number of the questions of this type in the question pool, the code length is determined by the number of the questions in the question pool.

The paper designs a six-dimension vector to ascertain a question (score value, degree of cognition, difficulty, knowledge point, type, answer time), and these attributes are symbolized and represented in binary code.

(2) Ascertain adaptive function

After coding, we need to ascertain the adaptive function of the test question. The adaptive function is used to describe the adaptive degree of each individual. Introducing adaptive function is to evaluate and compare each individual, basing on which assessing the superiority of each individual in terms of adaptive degree. Two principles should be conformed to when we convert objective function into adaptive function: one is that the value of adaptive function must be greater than or equal to zero; the other is that the varying trend of the objective function while optimizing should be conformity with the varying trend of the adaptive function while population evolving.

The adaptive function of the problem here is defined as:

$$F = \sum_{i=1}^{n} w_i f_i \,, \quad \text{i=1…n, considering a n-direction vector}$$

In it, w_i is corresponding to the importance weight of factor i, and $w_i > 0$. As a result of independent coding, it is generated in initial population. Type of question, number of questions and score value of question can meet demand all. Therefore, only point of knowledge, difficulty and teaching requirement should be considered. f_i is the absolute value of the error between test paper generating factor i and the user's demand. F is total of all the absolute value of the error between test paper generating factor and the user's demand.

(3) Design the generic operator

Generic algorithm is composed of three generic operators: selection, crossover and mutation operator[3-5]. The methods or the tactics of generic operator are different as for the different problems to solve. Considering coding tactic and initial population restraints, this system improves the generic operator in the following ways:

Selection operator uses probability computing method, which is relative to the adaptive degree value. The concrete steps as following: first, calculate the total of the adaptive degree value of each individual; then, calculate the proportion of each individual's adaptive degree value and use it as respective selection operator, by which calculate whether one individual will be chosen.

(4) Ascertain the terminated rule of algorithm

The simplest standard is stopping after a fixed number times of cycles. The fixed number relies on the complication of the model, which is decided by the number of the questions in the pool and the number of the restraints, so called i and k. This system will terminate after ik cycles.

The other terminated standard is based on the convergence of the population. If the bigger adaptive value between that of the former and the latter generation is smaller than one given value, we consider that the population is in stable condition, and the population mutation tend to converge, and the course of population mutation can be terminated.

(5) Ascertain the controlling parameters

The scale of the population affects the final outcome of the generic optimizing and the executive efficiency of the generic algorithm. If N is small, the convergence is fast but it is difficult to find out the optimal solution. If N is big, it tends to find out optimal solution but it has a high computing complication and a slow convergence

speed. Commonly, N is from 10 to 160. The crossover probability and mutation probability have great importance to generic algorithm. The bigger the crossover probability (Pc) and mutation probability (Pm), the more powerful the ability of the algorithm generates new individuals is. The smaller the crossover probability (Pc) and mutation probability (Pm), the faster the convergence is, but the premature may occur. Currently, we choose Pc=0.6-1.0, Pm=0.005-0.01.

The paper uses the self-adaptive idea, adjusting Pc and Pm during the course of computing, which vary as the individual adaptive value varies. Use the following calculating formulas:

$$Pc=0.8exp(-f/f_{avg}).....f\geq f_{avg} \quad (1)$$

$$Pc=0.8.....f<f_{avg} \quad (2)$$

$$Pm=0.01exp(-f'/f_{avg})/f_{max}.....f'\geq f_{avg} \quad (3)$$

$$Pm=0.01.....f'<f_{avg} \quad (4)$$

In it, f_{avg} is the average of the population adaptive value; f is the bigger adaptive value of two crossover individuals; f'' is the adaptive value of mutating individual.

3 Algorithm Test

To validate the feasibility and the validity of the improved generic algorithm in generating test paper automatic, make a generating-test-paper experiment, which used the curriculum of Computer Network as example.

The question pool contains 600 questions; the number of each type is 150, including multiple choice, choice (mark all correct), blanket-filling and short answer. In order to make each attribute of test question reasonable distributing, the value of each attribute is generated by random function.

Ascertain the parameter of generic algorithm: the scale of initial population is 80; crossover probability Pc=0.8, mutation probability Pm=0.1; the maximum number of cycles D=400; the total score is 100; the estimated time is 120 minutes.

We adopt three algorithms to solve this problem, and the experiment data as shown in table 1.

Table 1. Diagram 1 comparison of three algorithms to solve the problem of generating test paper

algorithm	Times of experiments	Failure times	Average iteration times	Average time
Random iteration method	40	2	85	60 seconds
Simple generic algorithm	40	0	60	38 seconds
Improved generic algorithm	40	0	50	36 seconds

The above data shows that when the question pool is large and the number of questions is very large, random iteration method is slow and the success rate is low; while generic algorithm improves the speed of generating test paper greatly and advances the success rate of generating test paper.

With the comparison of the two generic algorithms, we see that the course of simple generic algorithm is instable and the convergence success rate is low but it converges fast; whereas the course of improved generic algorithm is more stable and can ensure the individual's diversity however the speed doesn't show its advantage.

In all, considering the enlargement of question pool in future, the improved generic algorithm will accomplish its task to generate test paper better certainly and its advantage will stand out.

4 Conclusion

The paper makes improvement on traditional generic algorithm in terms of coding tactics, adaptive function, controlling parameter, and adopts a self-adaptive generic algorithm, which enhances the efficiency and quality of generating test paper. Its application in the curriculum of Computer Network indicates that it is reasonable and efficient.

Acknowledgements. Thanks are due to Li Wei-feng for assistance with the experiments and to Du Jing-yi for valuable discussion.

References

[1] Wang, Y.-R., Zhang, Z., Shi, Y.-X., Yao, R.: Theoretical Model and Method of Intelligent Test Paper Generation System from Item Pool. Journal of University of Electronic Science and Technology of China 35, 363–366 (2006)

[2] Ma, S.-P., Zhu, X.-Y.: Artificial Intelligence. Qinghua University Publishing house, Beijing (2004)

[3] Wu, M.-J., Huang, Y.-B., Liu, Z.-Y.: Research on Auto-composing Test Paper System Based on Improved Genetic Algorithm. Computer Applications and Software, 192–193 (December 2007)

[4] Galante, M.: Genetic algorithms as an approach to optimize real-world trusses. International Journal for Numerical Methods in Engineering 39(1), 361–382 (1996)

[5] Leung, Y.W., Wang, Y.P.: Multi-objective programming using uniform design and genetic algorithm. IEEE Trans. Systems, Man and Cybernetics, Part C: Applications and Review 30(3), 293–304 (2000)

Intelligent Control Technology for Frequency Conversion Pump in Air Conditioning Systems

Shizhong Yang[1,2,*] and Qingchang Ren[1]

[1] Xi'an University of Architecture and Technology, Xi'an, 710055, China
[2] Automation Engineering College of Qingdao Technological University, Qingdao 266520, China
yszxlj@126.com

Abstract. Pump is main equipment in air conditioning systems and used widely, but its energy consumption is great. In order to improve energy efficiency, using pump controlled by frequency-conversion technology is an effective method. In this paper, the working principle and control method of variable frequency pump is described in detail, and gives a comparison between variable frequency control and traditional control about energy saving effect. Because traditional PID controller can be hard to get good control performance in pump control, this paper presents a control method using fuzzy logic controller to achieve frequency conversion and constant pressure water supply. The controlled variable is obtained by fuzzy control table. The system has energy-saving, security, and stable water pressure, etc. Simulation results show that the proposed controller not only have good robustness, but also have good dynamic and static performance.

Keywords: Pump, frequency conversion, energy-saving, fuzzy control.

1 Introduction

Pump is main power equipment in air conditioning systems, but it is a large of energy-consuming equipment. Generally, pump power consumption accounts is about 70% of the total power consumption in water supply system , and its operation cost ranks first in water supply system; pump power consumption in air-conditioning pump accounts for about 18% of total electricity consumption, while the power consumption of air-conditioning accounts for about 40% to 60% in the building power consumption. In order to meet the requirements of water consumption and water pressure in practical application, the working conditions pump needs to be adjusted. Using the traditional adjustment way of mechanical valve to change the pump's working conditions (pressure and flow size), the energy loss is bigger. So a new, remarkable energy saving methods need be adopted, namely inverter adjustment. According to pump frequency characteristics, this paper constitutes a PID fuzzy controller of the automatic control system. Error which is existed between the pipe network pressure and given pressure and error rate is as fuzzy controller inputs.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 665–669.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Control variable is obtained by fuzzy control table, to achieve frequency conversion and constant pressure water supply automatic control. System is simulated with the MATLAB, the effect of fuzzy controller is analyzed from robustness and stability.

2 Control Method

2.1 Control Strategy

When pump of air conditioning systems is operating under industrial frequency power, the size of the flow only is adjusted by mechanical valves. In recent years, for energy saving and satisfy load requirements, frequency-inverter regulation technology is universally used to adjust the pump motor speed, which changes the flow of water. In a certain range, satisfactory effect of energy saving and adjustment can be got by frequency-inverter regulation technology [1]-[3].

In air conditioning systems, pump control usually uses differential pressure control method. Pump's working characteristics (flow, pressure changes) are closely related with the pipeline characteristic. While the pipeline resistance increases, the pump flow decreases and the head of delivery increases. The changes of flow cause the differential pressure changes of return pipe. The differential pressure signal inputs to frequency controller, and compares with the set value to control the pump speed.

The role of pump differential pressure system is to control and maintain terminal pressure in a set range, and to ensure the dynamic balance of water supply system. Differential pressure control system establishes differential pressure sensors in the export of frequency conversion pump and the end of the appropriate location. The pressure differences between water supply pipes are measured using sensors. Using pressure difference and security pre-set, according to the appropriate algorithm for processing, a speed-controlled output signal is got. By the interface with the inverter, the inverter achieves the pump speed regulation. Pump control principle is shown in Fig.1:

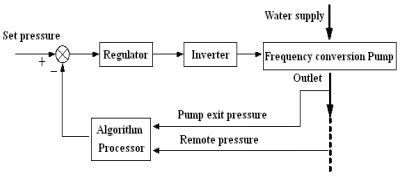


Fig. 1. Converter control schematic

The pump motor is the output link, which speed is controlled by the inverter to achieve variable flow constant pressure control. The converter receives the controller's signal for speed control of pumps. Controller synthesizes the given signal and feedback signal, through the regulator output operating frequency command to the inverter. Pressure sensors detect the effluent water pressure of pipe network and convert it into an acceptable analog signal for controller to adjust.

Control system working principle: frequency control is ultimately achieved by adjusting the pump speed, and pump is the execution unit of water supply. Pump head flow characteristics shown in Fig.2.

The horizontal axis is the pump flow rate Q, and the vertical axis is the pump head H. The relationship between pump head and water pressure is linear, therefore, it can be approximately expressed as the water pressure P. EA is the constant pressure line, n_1, n_2 and n_3 are flow-pressure characteristics under different speeds. It can be seen in the n1 speed, if the flow reduced to Q_C from Q_A by controlling the valve opening, the pressure will increase along the n1 curve to the D point. Clearly, while reducing traffic, the pressure is improved (DC segment is pressure rise). If the speed decreases from the n_1 to n_3, the flow rate decreases along the constant pressure line from Q_A to Q_C when the pressure has not changed. It can be seen in a certain range that the flow can be adjusted by changing the speed under the premise of maintaining a constant water pressure and losses don't exist because of the rising pressure. This characteristic shows that adjusting the pump speed can change water flow, stabilize the pressure at constant pressure line and achieve differential pressure control.

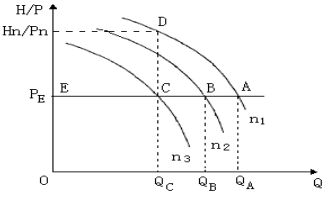


Fig. 2. Pump head flow characteristics

2.2 Fuzzy Control Technology

Pump control system is nonlinear, time-varying, hysteretic, at the same time it can be affected by external noise, load disturbances and other uncertainty factors. It is difficult to establish a precise mathematical model, so the control accuracy is not high. Because traditional proportion integral derivative (PID) control is easy to form system oscillation, its control effect may be somewhat fluctuations, and in wide range. It improves the system's energy consumption. If the system has general quality requirements of hydraulic precision and its mathematical model is difficult to determine, the fuzzy control is a practical and effective method. It can eliminate the theory of limit cycle oscillations [4]-[6]. The control principle is to measures the differential pressure of export and terminal of frequency pump through pressure sensor, and compare the stress value and the internal set pressure value, so as to get deviation signals. The fuzzy logic controller (FLC) outputs control signals, and relay the signals to the frequency converter. The frequency converter adjusts the asynchronous motor speed, in order to adjust the hydraulic pressure online.

Fuzzy logic controller's design thought is as follows: According to the fuzzy reasoning, the differential pressure error (e) and the change of (ec) are inputs; the frequency converter control variable u_f is output. Fuzzy relations of deviation are found to correctly build the fuzzy control rules. Through constantly detecting

conversion rate and calculating error, the system use fuzzy rule to infer and make decision, and adjust online through output variables of Fuzzy controller [7],[8].

All membership functions are defined as triangular partitions with seven segments from -3 to 3.Each domain interval is divided into 7 equal regions, denoted by NB(negative big), NM(negative medium), NS(negative small), ZO(zero),PS(positive small), PM(positive medium) and PB(positive big) and assigns each region a fuzzy membership function.

The control rule is got by the learning design experience of conventional fuzzy logic controller. The rule base of the FLC is shown in Table 1.

Table 1. Fuzzy rule table of output u_f

u_f		e						
		NB	NM	NS	ZO	PS	PM	PB
ec	NB	NB	NB	NB	NM	NS	NS	ZO
	NM	NB	NM	NM	NM	NS	ZO	PS
	NS	NB	NM	NS	NS	ZO	PS	PM
	ZO	NB	NM	NS	ZO	PS	PM	PB
	PS	NM	NS	ZO	PS	PS	PM	PB
	PM	NS	ZO	PS	PM	PM	PM	PB
	PB	ZO	PS	PS	PM	PB	PB	PB

From fuzzy control rule table, the method of selecting the controlled variable is: When the water pressure deviation is bigger, controlled variable is large in order to eliminate bias; when the water pressure deviation is smaller, controlled variable reduce correspondingly in order to stabilize the system.

3 Simulation and Analysis

By comparing the traditional PID controller with FLC controller, the steady and transient characteristics are studied. The input is step signal, the simulation results are shown in Fig.3.

It can be seen from the simulation figure that the FLC controller has smaller overshoot and shorter adjustment time, and its dynamic response and steady-state characteristics are superior to those of conventional PID control. When the response of FLC controller is near to the balance and has tended to deviate, the speed of curve's change becomes slower. After several times, the response of system becomes stable at the balance at last. In the FLC controller, control can be adjusted automatically according to the change of process. Thus, this kind of controller has better adaptive ability, self-adapting ability and good robustness.

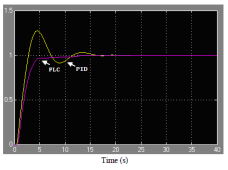


Fig. 3. Comparison of control response under two controllers

4 Conclusion

Pump with variable frequency control is a promising technology with broad application prospects in air conditioning systems. The advantages of frequency control technology embodied in two aspects: one is that the system has a good stability, high reliability; the other is that the system has significant energy saving, high efficiency. Because the delayed impact of the system, conventional PID control performance is not ideal, and controlling volatile is extent wide. Using fuzzy control algorithm realize frequency conversion, it can effectively solve nonlinear, time-varying and the accurate control of non-mathematical model. It is simple in design, convenient to implement, optimized in control, strong in robustness and good in real-time. Compared with the traditional pump control method, this control method has the incomparable superiorities in the economy, stability, reliability and automation level, and can certainly bring good economic and social benefits.

Acknowledgment. This work is supported by the Foundation of Educational Commission of Shaanxi Province of China (No.11JK0906) and the Youth Foundation of Xi'an University of Architecture and Technology (No. QN0931).

References

[1] Wang, X., Li, Q.: An Application of Variable Frequency Speed Control Technology to the Pump Energy Saving. Journal of Henan Institute of Engineering 22(3), 20–22 (2010) (in Chinese)
[2] Jackson, S., Liggett, F.: How to match ac motors and variable-speed drives. Plant Engineering, Barrington 56(7), 52–56 (2002)
[3] Ma, G., Tan, G., Chen, C.: Energy Saving Discussion For Speed Regulation Using Frequency Conversion for Pumps. Journal of Sichuan University of Science and Technology 22(3), 5–7 (2003) (in Chinese)
[4] Kukolj, D.D., Kuzmanovic, S.B., Levi, E.: Design of a PID-like compound fuzzy logic controller. Eng. Appl. Artif. 14, 785–803 (2001)
[5] Carvajal, J., Chen, G., Ogmen, H.: Fuzzy PID controller, design, performance valuation and stability analysis. Inform. Sci. 123, 249–270 (2000)
[6] Wai, R.J., Kuo, M.A., Lee, J.D.: Cascade direct adaptive fuzzy control design for a nonlinear two-axis inverted-pendulum servomechanism. IEEE Transactions Systems, Man and Cybernetics 38(2), 439–454 (2008)
[7] Liu, D., Wei, D., Liu, X.: Temperature fuzzy PID control for variable air volume system. Journal of Civil, Architectural Environmental Engineering 31(4), 98–102 (2009) (in Chinese)
[8] Adams, M.J.: Genetic multi-stage fuzzy PID controller with fuzzy switch. In: Proceeding of the IEEE International Conference on Systems, Man and Cybernetics, pp. 2239–2244 (2001)

The Application and Development of Photoelectric Sensor

Zheng Chunjiao

Electrical Engineering Institute, Liaoning University of Technology,
Jinzhou, China 121001
zcj5894@yahoo.com.cn

Abstract. At present, the application of photoelectric sensors are more and more extensive, it also promotes the development of photoelectric sensors. Photoelectric sensor has simple structure and diversity. It has high precision, fast response, non-contact and other advantages. In this paper, we analyze the principle of photoelectric sensors, introduce the classification of photoelectric, and then highlight introduce the application of photoelectric sensors and the use of the principle of photoelectric sensors, analyze the current and future development of photoelectric sensors.

Keywords: Photoelectric sensor, the application of photoelectric sensor, the development of photoelectric sensor.

1 Introduction

Photoelectric sensor is the sensor to use the electronic and optical element as the detection component. Photoelectric detection has high precision, fast response, non-contact advantages and so on. The sensor has simple structure, flexible and diverse forms. Therefore, the photoelectric sensor is widely used in the field of control and testing. It can be used to detect the non-electricity which can cause changes in the amount of light, such as light intensity, radiation temperature, gas composition. It can also use light transmission, occlusion, reflection, interference and others to measure a variety of physical quantities, such as object size, displacement, velocity, temperature, etc.. So it is a important and sensitive device which has a very wide application. When using the photoelectric sensor, it doesn't directly contact with the measured object and the beam quality is nearly zero, there is no friction in the measurement and almost no pressure on the measured object. Therefore, photoelectric sensors has obvious advantages than other sensors in many applications. However, its drawback is that optical devices and electronic devices are more expensive in some applications, and environmental conditions require higher on the measurement. In recent years, new optoelectronic devices are been emerging, particularly the birth of CCD image sensor, that creates a new field for the further application of photoelectric sensors.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 671–677.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 The Principle of Photoelectric Sensor

Photoelectric sensors use photoelectric elements as sensor conversion devices. The principle of photoelectric sensor is to reflect the measured objects changes by light signal, and then convert the light signal into electrical signals by optoelectronic components. Usually the photoelectric sensor is composed by light source, optical access and optical components. The working process of photoelectric sensor is shown in Figure 1.

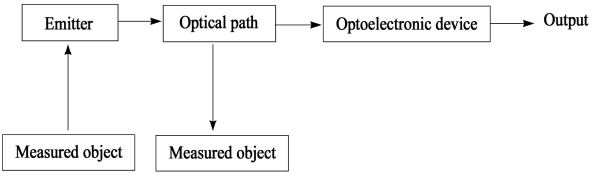


Fig. 1. The working process of photoelectric sensor

The role of optoelectronic devices is to convert the optical signal into electrical signal, which is based on the photoelectric effect. Photoelectric effect is a physical phenomenon that the light shines on certain substances and causes the material to change in the electrical characteristics. It can be divided into external and internal photoelectric effect.

External photoelectric effect is the physical phenomena that the object's electron escape surface of the object and emits outside under the influence of light. Photon is in the form of quantum "particles" to describe visible light waves. Photon energy is hv, h is the Planck constant, v is the optical frequency. Photon flux corresponds to light intensity. External photoelectric effect is described by the Einstein equation:

$$hv=1/2*mv0^2$$

Where m is electron quality, v0 is electronic escape velocity. When the photon energy is equal to or greater than the work function, the external photoelectric effect can be generated. So each object has a corresponding effect on the photoelectric threshold frequency of light, known as the red limit of frequency. For more than the red limit of the incident light frequency, exogenous photocurrent is proportional to light intensity.

When light shines on the object, it can cause the resistivity to change or generate photo-emf which is called the internal photoelectric effect. The internal photoelectric effect is divided into photovoltaic effect and photoconductivity effect. Photovoltaic effect is the phenomenon that makes objects have a certain direction of the force under the influence of light. Optoelectronic devices based on the effects are photosensitive diodes, transistors, photovoltaic cells; photoconductive effect is that under the influence of light, electrons absorb the photon energy from the bonding state of transition to a free state, which leads to material changes in conductivity. Optoelectronic devices based on this effects are photosensitive resistor and so on.

3 The Classification of Photoelectric Sensor

Photoelectric sensor classification is varied. Photoelectric sensing devices based on external photoelectric effect are photoelectric tubes and photomultiplier tubes. Photoelectric sensors based on photoconductive effects are photoresistors. Photoelectric sensors based on barrier effects are based on photodiodes and phototransistors. Photoelectric sensors based on reverse side of the photoelectric effect are reverse photodiodes. Photoelectric sensors can be classified into analog optical sensors and digital optical sensors by signal forms. Photoelectric sensors also include fiber optic sensors, solid image sensors and so on. Figure 2 shows common photoelectric sensors.

Fig. 2. Common photoelectric sensors

4 The Application of Photoelectric Sensor

With the development of technology, the use of photoelectric sensors is more and more. Photoelectric sensors have a growing number of applications in everyday life, industrial production and information technology.

4.1 The Application that Light Source Is the Measured Object

The unit of light illumination E is lx (lux), which is one of the units commonly used in photometry, it means the degree of exposure of the illuminated physical object, it can be used to measure the illuminance meter.

1)The infrared radiation thermometer

Infrared radiation thermometer has many applications in non-contact temperature measurement. Infrared radiation thermometer can be used for both high temperature measurement and below freezing temperature measurement, this is also the trend of radiation thermometer. Commercially available infrared thermometer temperature is range from -30 ℃ ~ 3000 ℃, the middle is divided into several different specifications, you can select the appropriate model according to the need. Figure 3 shows the shape of the infrared radiation thermometer.

2) Pyroelectric sensor

Pyroelectric infrared sensor can detect infrared light transmitted by person or animal and output corresponding electric signal. Pyroelectric infrared detection

Fig. 3. The shape of the infrared radiation thermometer

devices have been widely used in infrared spectroscopy, infrared remote sensing and radiation detectors. It is also used in human detection, alarm, which can produce far-infrared radiation, such as security doors, hotel lobby automatic doors, automatic light control. For example: it will automatically turn off air conditioners, drinking fountains when no one is in the room; TV can judge that no one is watching or person has to sleep and automatically shutdown the circuit. Figure 4 is the shape of the pyroelectric sensor.

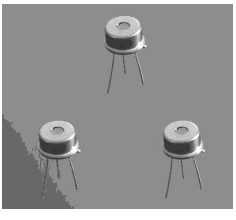


Fig. 4. The shape of the pyroelectric sensor

4.2 The Application That the Measured Object Absorbs Light Energy

In this application, the measured object can absorb light energy, according to the weakening degree of light energy to achieve the target analyte measurement.

1) Photoelectric turbidimeter

Photoelectric turbidimeter is based on the principle that the transmission attenuation and scattering attenuation of infrared light are relate to the suspended sediment concentration to achieve the concentration of sludge and suspended solids measurements. Different types of sludge and suspended matter have different attenuation to infrared light, we can adjust the transmission frequency and use different algorithms to meet different applications.

2) Smoke alarm

No smoke, a photosensitive element receives a constant infrared LED transmitter. In the event of fire, smoke comes into the testing room and blocks some of the infrared light, the phototransistor output signal is weakened, it is judging by the threshold circuit, an alarm signal is occured. Figure 5 is smoke alarms.

Fig. 5. Smoke alarms

4.3 The Application That the Measured Object Reflects Flux

This sensor uses the principle of diffuse reflection. In such sensors, emitter and receiver mounted on the same device. The light emitted by the emitter is reflected by the target object, the reflected light is in all directions, part of the reflected light reflects into the receiver, so that can detect the target object.

1) Reflective smoke alarm

In the absence of smoke, due to infrared tube is perpendicular to the interior and painted black smoke absorption material, so infrared LED infrared light can not reach the infrared phototransistor. When the smoke comes into the smoke chamber, solid particles of the smoke produce diffuse reflection towards infrared light, so that part of the infrared light reaches the phototransistor and light current output.

2) Photoelectric tachometer

Photoelectric tachometer is reflective photoelectric sensor, which can be tens of millimeters away from the measured object outside and measure the speed without contact.

3) Color sensors

Color sensors are used to detect a specific color or spot on the object, it measures the color by comparing with the zone without color, rather than direct measures color.

4.4 The Application That Measured Object Blocks the Light

The shading photoelectric sensor needs two separate cases, light-emitting device is installed in a chassis, receiver is installed in another chassis. The light emited from emitter injects to the receiver, when the target object blocks the light, the receiver's output will change. In the photoelectric detection technology, beam block is the most efficient and can detect the largest range.

The photoelectric detector of strip deviation. When the strip goes wrong road, the edge often collides with the sending machine, resulting in waste. When the strip is in the correct position (center position), the amplifier output voltage is zero; when the strip goes left-side, the shading area is reduced, the output voltage reflects the deviation of strip direction and size. The photoelectric detector of strip deviation is shown in Figure 6.

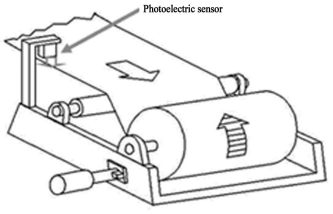


Fig. 6. The photoelectric detector of strip deviation

5 The Development of Photoelectric Sensor

Photosensitive semiconductor devices rapidly develop with the development of semiconductor technology in the 60's. During this period, a variety of optical materials have been fully studied and widely used. People have developed a variety of optoelectronic devices suitable for different occasions by the research about the principle of the photoelectric effect and device. Photoelectric sensor manufacturing process also achieves a high level with the development of film technology, surface technology and large scale integrated circuit technology, and the cost of the product is greatly reduced. In the past few decades, infrared sensors and other optical sensing devices access to a wide range of applications in the field of aerospace and defense. In the next few years, the main sensor technology will be breakthroughs in the sensitivity, resolution and overall performance. Photoelectric sensor applications has been extended to the textile, paper making, printing, medical, and environmental protection. The study of traditional areas has a new development, such as infrared detection, radiation measurement, optical fiber communication, automatic control and so on.

Infrared sensor and low light level sensor are two of the most widely used photoelectric sensors. The resolution of infrared sensors is increasing, and the reliability is improved with reducing the demand for cooling. Low light level sensors rely on the traditional image intensifier to improve, and the sensor's analog signals convert to digital signals. The development trend of the two technologies is the integration of a variety of multispectral sensors technology that can maximize the performance of individual technologies, more and more have applied to handheld devices and unmanned vehicles and other fields.

Combination of several different photoelectric sensors' strong points has many advantages, but only it is used in an appropriate forum that can play these advantages. For example, low light level sensors are more suitable for observing a wide area at night, but are insufficient to detect static person hidden in the woods. In this case, if we fuse long-wave infrared camera with low light level sensors, the staff after trees will be visible. If we fuse another short-wave infrared sensor, we can even identify the person's facial features after trees.

In addition, another major trend of the photoelectric sensor is to increase the number of pixels focal plane arrays.

References

[1] Lv, X.: The research of automatic monitoring system of Precision Planter. Dissertations (2007)
[2] Li, Q.: Pyroelectric infrared sensor. Electric Time (2004)
[3] Huang, J.: The automatic monitoring system of the sewage treatment plant based on DCS. Dissertations (2007)
[4] Li, X.: Auto-matching design and implementation of three parts of air conditioning compressor. Dissertation (2008)
[5] Zhang, G.: Photoelectric sensor applications and development trends. Science and Technology Information (academic research) 15 (2008)
[6] Wang, F., Fang, J., Zhang, X.: The research of greenhouse environment monitoring system based on wireless sensor network technology. Automation Technology and Application 10 (2009)

Online Random Seeking in FLV Video Based on File Cutting

Zhihui Du[1,*], Wenjun Fan[2], Yunpeng Chai[3], and Tianle Zhang[1]

[1] Tsinghua National Laboratory for Information Science and Technology
Department of Computer Science and Technology, Tsinghua University, 100084, Beijing, China
[2] School of Software Engineering, Beijing Jiaotong University, 100044, China
[3] School of Information, Renmin University of China, 100872, Beijing, China
duzh@tsinghua.edu.cn, 08122361@bjtu.edu.cn

Abstract. To support online watching any part of the videos smoothly on the streaming media cluster architecture, a file cutting based random seeking technology in FLV video is proposed in this paper. Three important aspects in design and implementation of this technology, the policy of player access to the back-end server, the FLV video file cutting algorithm, and the file deleting strategy, are described in detail. The method has been employed in the Outstanding Courses Integration System of China and the simulation results help us to find the performance bottleneck to be broken in the future work.

Keywords: streaming media, FLV cut, random seek.

1 Introduction

With the development of server, terminal playback equipment and computer network technology, people have an increasingly demand associations with streaming media servers. So streaming media server needs to improve its performance significantly. Streaming media server cluster has gradually become the mainstream of high-end streaming media server [1-2] because of its high degree of scalability, fault tolerance and resource sharing efficiency. Meanwhile, when watching video programs on the internet, users will often want to view those parts that have not been downloaded by dragging the process bar of the video player. Consequently, streaming media server cluster has to provide a mechanism which can make user's random seeking possible.

FLV, which is currently used by a number of next-generation video sharing website, is the fastest growing, most widely adopted video formats, and is developed on the basis of compression algorithm of Sorenson companies. FLV format not only can easily import Flash with fast speed, but also can support copyright protection.

FLV is a new video format which arises from Flash MX's emergence. It is based on the widely used Flash Player platform and it can integrate the video into the Flash animation. In other words, site visitors can look as long as Flash animations, naturally

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 679–688.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

able to see FLV format video without having to install an additional plug-ins player. FLV video for video transmission has brought great convenience.

FLV has very low hard disk occupancy. A one minute FLV video with high quality is only 1MB in volume. In same period, the movie in regular video format will occupy 2/3 storage more than the movie in FLV format. Meanwhile, FLV video has pretty low CPU occupancy and high quality. These properties make FLV a very widely adopted video format on the internet. Abundance and variety in resource of FLV format is another important reason that makes FLV becoming the unified online video format. At the present day, there are two kinds of main FLV providing web sites on internet. One is some specialized video sharing web sites. Such as the American site YouTube, and Chinese site youku. The other is the video blog parts of some Portals. They provide their own video channel. For example, video blog of sina uses videos in FLV format. Besides, Baidu released a video searching service and most of the results are in FLV format.

Our system uses FLV format videos to provide services as well. This paper relies on streaming media subsystem which belongs to National Outstanding Courses Integration project. It gives a design and implementation of how to cut FLV videos in a specific way according to user requirement.

This paper is organized as follows: In section 2, we introduce the architecture of the cluster streaming media server. In section 3, we analyze the implementation of online random seeking. Section 4 explains the policy of the player access. Section 5 depicts the FLV cutting algorithm. In section 6, we describe the file deleting strategy. The simulation results are given in section 7. Finally, section 8 gives the conclusion and discussion of future work.

2 The Architecture of the Cluster Streaming Media Server

The streaming media sub-system of National Outstanding Courses Integration projects is the services support system, which provides all kinds of course video and audio for the national outstanding courses integration system, and provide users on-demand video and audio content, and DRM encryption and authentication services. The sub-system is a cluster streaming media server, which includes three kinds of nodes: schedule server, streaming media server and storage server, as shown in Fig 1, and three kinds of modules: scheduling module, cache module and storage module. In this sub-system, we use HTTP protocol to provide the streaming media service.

The scheduling module provides Lighttpd [6] service software which can support HTTP service. The lighttpd service is deployed on the scheduling server node. Its main function is to choose the best streaming server node to provide service with better load-balancing. The cache module is deployed on the streaming server node. And the cache module's function is mainly provided by Squid [10]. It also has Lighttpd service and it employs FASTCGI [7] technology to supply better HTTP service. Storage module is deployed on SAN [11] which consists of disk array and fiber channel.

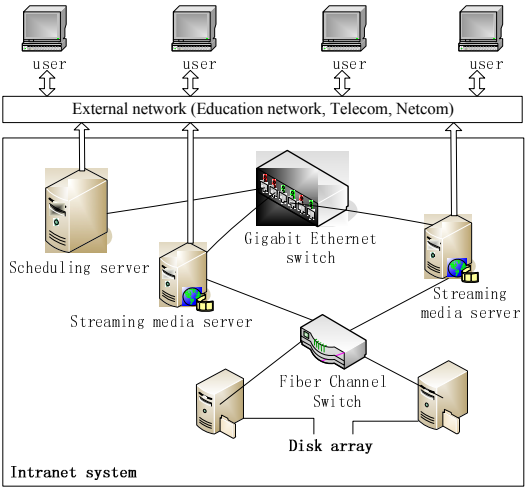


Fig. 1. Architecture of the streaming media sub-system

As the popularity of streaming media content follows the Zipf [8] distribution, 80% of users will focus on watching 20% of the contents. So the cache module in the storage node can improve the overall system performance [8, 9].

3 Implementation of Online Random Seeking

Currently, to meet users' random seeking requirement during watching videos online, mainstream video web sites like YouTube and youku use a popular online random seeking technology which is based on FLV media stream accessing. This technology cannot take advantage of cached parts in large memory. So this paper proposed a technology based on file cutting to implement online FLV video random seeking.

3.1 Implementation Based on FLV Media Stream Accessing

The media stream accessing based random seeking is specialized by FLV streaming with Lighttpd [3] FLV video. If any part of one key frame is added with "FLV\x1\x1\0\0\0\x9\0\0\0\x9" as public head symbol information, it can be read and played by any FLV supported media player. By using GET method of HTTP, we can ask for a specific video segment from the server. The URL is like this:

http://youserver.com/flv/abcdefg.flv?start=12345.

Here "start = xxxx" should be the key frame position of the FLV media file, otherwise it can't be played.

The implementation based on Media stream accessing will set all the server pressure to media stream storage nodes. This approach could be well supported if there were sufficient nodes to share pressure from the server.

The above approach would have some shortcomings if used in this system. Firstly, we could not provide enough media stream storage nodes like a commercial video website. Storage nodes occupancy will be considered as a restrictive condition in the system. Secondly, in order to improve service quality, we use squid buffer memory service in this system. However, squid can only support HTTP based buffer memory, but not media stream.

3.2 Implementation Based on FLV File Partitions Accessing

Based on the approach of file cutting online random seeking in FLV video, we cuts video files into parts according to key frames, and then stores a copy of each part into a streaming media node. When a GET request send by users arrives, system will find every part from the specific one to the end, and send back to the client. The client then could get what he wants to watch. In addition, the squid buffer memory of streaming media nodes supports file buffers which based on HTTP. Meanwhile video file parts could be transferred by HTTP. Thereby this overcomes the shortages of implementation based on FLV Media stream accessing. The implementation is shown in Fig 2.

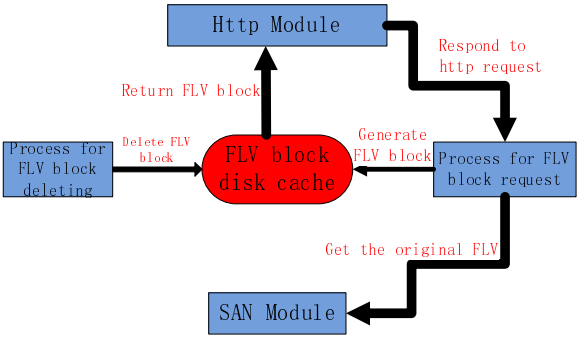


Fig. 2. Implementation of FLV block accessing

4 Policy of Player Access to the Back-End Server

It is transparent for the players of clients to seek in FLV video file supported by the server. Whenever a video player sends requests according to legal GET method to server, it will get corresponding video parts.

When a video file has been cut into parts, the duration of original won't be changed, but the size will. Therefore, before playing, players need to know how many parts the FLV has, the URL of each part, the size and duration of the original FLV file. Some information could be got by reading FLV file's metadata, including key frames, size and duration. Hence, players could do random seeking by two steps. The procedures are as follows:

Step 1 :
Player request data:
http://server.com/server?media/flv/20090521/1.flv
Server response data:
<root time="360" byte="654321">
<flv url="http://server.com/server?fname=flv_1.flv"/>
<flv url="http://server.com/server?fname=flv_2.flv"/>
…
<flv url="http://server.com/server?fname=flv_n.flv"/>
</root>
Step 2 :
Player request data:
http://server.com/server?fname=flv_2.flv
Server response data:
flv_2.flv

5 FLV Video File Cutting Algorithm

This system cuts video files into 2 minutes parts, therefore, flv_1.flv represents details from 0 sec to 120 sec.

As long as server receives a request like "http://server.com/ server?fname=flv_2.flv", it will figure out that users want to get the second part of original video file, which lasts from 120 sec to 240 sec. The cutting core algorithm will identify the start key frame and the end key frame which are nearest to the start time and end time respectively. The pseudo-codes are as follows:

```
0:  def round_to_keyframe (cut_time):
1:      best_time = -1;
2:      best_diff = 1_000_000_000;
3:      while keyframe_times:
4:          diff_time=abs(cut_time – keyframe_time);
5:          if  diff_time<best_diff:
6:          best_diff=diff_time;
7:          best_time=keyframe_time;
8:      return  best_time < 0?cut_time:best_time;
```

Afterwards, according to start key frame and end key frame, we use Perl's FLV::Cut module [4] to cut FLV file and get flv_2.flv part then save it into the local disk.

6 Partition File Deleting Strategy

Squid buffer service of buffer nodes stores those partition parts with high access rate, in this way the hit rate raises and the pressure of the storage nodes is reduced. But along with hit amount raises, streaming media nodes will have to save more and more file parts. When the disk runs out of space, new partition parts would not be able to be generated because of short of disk space. As the consequence, client's random

seeking will fail for it couldn't get any file part. So we need a strategy to delete the existing file parts.

Squid itself has cache exchange strategy. We can use squid's logs to change out cache's file partitions and delete them simultaneously from streaming media nodes. However, the amount of file partition parts stored in streaming media nodes is much bigger than that in squid's buffer. If we execute like that, some parts that never been accessed will stay in the disk forever as trashes and will never be deleted. So this strategy won't work. To overcome this shortcoming, this paper uses LRFU exchange strategy [5] to delete file partition parts.

LRFU uses LRU and LFU coordinately according to weight function. This algorithm gives attention to both access recent information and frequency information. Its principle is the latest accessing part has the highest weight, and the earlier, the lower. Then we calculate CRF with weight function, and delete file partition parts according to CRF. CRF's computing formula is like below:

$$C_{t_{base}}(b) = \sum_{i=1}^{k} F(t_{base} - t_{bi}) \cdot$$

C_{tbase} is the CRF value of part b in the base time. $F(x)$ is the weight function, $\{t_{b1}, t_{b2}, t_{b3}, ...t_{bk}\}$ is the accessed time of part b, $t_{b1} \prec t_{b2} \prec ... \prec t_{base}$

$$F(x) = \left(\frac{1}{p}\right)^{\lambda x}, (p \geq 2).$$

The characteristics of the weight function:

$$C_{t_{b_k}}(b) = F(0) + F(\delta) C_{t_{b_{k-1}}}(b), \delta = t_{b_k} - t_{b_{k-1}}$$

In special, we could calculate present CRF according to the last accessing time and CRF from last accessing time.

In addition, the weight function has another important feature as follows:

(1) If λ is 1, it is same with the LRU algorithm. Replacing the block according to the CRF is same with the LRU replacement algorithm.
(2) If the F(x) = C, C is a constant. When λ is 0 and F(x) is a constant, then it is same with the LRU algorithm.

While λ is gradually changed from 0 to 1, LRFU exchange strategy will transit from LFU to LRU. In this system we set λ=0.5 and p=2.

According to the access log files, we can get the accessing time and frequency of every file part. Then we can calculate CRF and order file parts from top to end by CRF value. When disk's occupancy is higher than 70%, the partition file deleting procedure will be started.

7 Experimental Results

The parameters of the streaming media server cluster node hardware in our test bed environment are shown in Table 1.

In the streaming media subsystem, scheduling server is currently only responsible for redirecting the user's request to the right streaming media node, requiring the least network bandwidth, disk reading and writing capability, and lowest CPU resources. Therefore, the main test is for the streaming media node. We use 150Kbps as small bit-rate and 1Mbps as large bit-rate, two kinds of video to do the testing to identify the system bottlenecks.

Table 1. Configration of the Cluster Vector

	schedule server	streaming media server
CPU	4core 3.0GHz	4core 3.0GHz
Memory	16G	16G
Disk	SCSI,60G	SAN disk arrays connected
NIC	6X1Gbps	2X1Gbps
Operating system	Linux	Linux
Interconnect switches	24-port Gigabit Ethernet switch	

7.1 Limit Test

In limit test, we use both the focus-on request and random demand. We gradually increase the amount of simulated users for streaming media server, until arrive the limit of some performance aspects. In the experiment, we mainly test the system utilization of CPU, the network bandwidth, and the reading and writing performance of disk.

In Fig 3, the CPU load and network bandwidth test results have shown that small bit-rate video takes more CPU resources in the same request mode. Because small bit-rate video files in the same bandwidth needs more CPU resources to process the request. CPU load is not only directly proportional to network bandwidth, but also relevant to bit rate. The smaller the bit rate is, the more CPU resources will be requested.

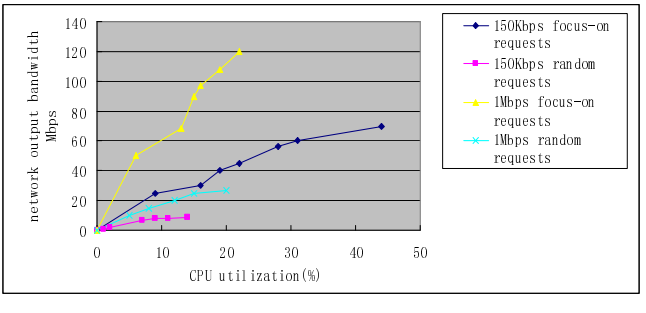


Fig. 3. CPU and network test

Under the same conditions, random requests need more CPU resources than focus-on requests. Because the random request need to consume more CPU resources to handle the disk reading and writing operations. In the random requests test, the maximum CPU occupancy rate is 20%, while the maximum network bandwidth is not more than 30Mbps. That CPU resources and network bandwidth are not the bottleneck of random requests patterns. But according to the relationship of CPU and network bandwidth, before the network bandwidth reach to 1Gbps, CPU will first arrive at its extreme point.

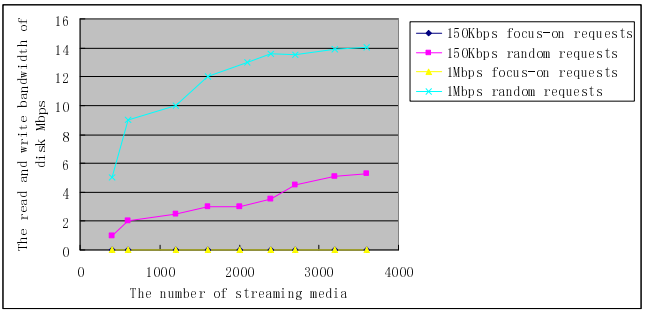


Fig. 4. Analysis of the disk performance

In Fig 4, according to the disk reading and writing performance analysis, for two kinds of bit-rate focus-on requests, the disk reading and writing operations are both close to 0, which indicates that focus-on requests data essentially come from the cache, and disk stress is negligible. In random requests, the disk reading and writing bandwidth increase with the growing number of streaming media nodes. When the disk reading and writing bandwidth of 150Kbps small bit-rate reaches to 5.3Mbps, or disk reading and writing bandwidth of 1Mbps large bit-rate arrives at 14.02Mbps, the disk reading and writing bandwidth stop growing and reach to the performance limits. This is because in the disk access process, the seeking and the waiting time will occupy a considerable proportion. The larger the media number corresponding to small bit-rate video is, the longer the seeking and waiting time will need, the shorter the time for continuous reading and writing disks is. So, the large difference of bit rate also has effect on the disk reading and writing performance.

7.2 Simulation on Demand

We use two kinds of bit-rate video, 150Kbps and 1Mbps and the average response time is shown in Fig. 5.

For 150Kbps bit-rate video, when the concurrent users reached 2200, the average response time will increase dramatically. We believe that the performance limit has been reached.

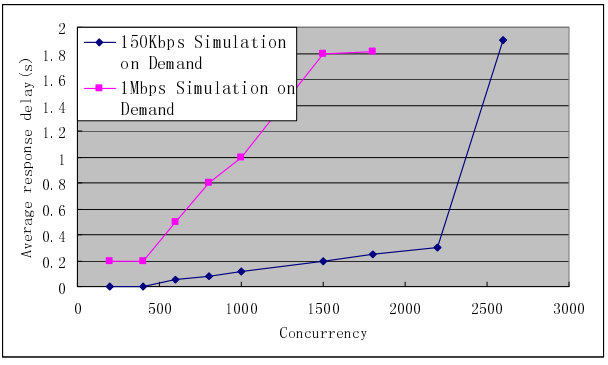


Fig. 5. Simulation test

For 1Mbps bit-rate video, when the concurrent users are more than 1,500, the delay will keep at 1.8 second. We believe that performance this time also reached the limit.

7.3 Performance Bottleneck Analysis

The above analysis shows that in the random request case, the disk reading and writing operation is the performance bottleneck of the system. For smaller bit-rate video, the bottlenecks are more serious; in the focus-on request situation, according to the relationship of CPU utilization and network bandwidth in Fig. 3, CPU will be the system performance bottleneck.

8 Summary

By designing and implementing the client access strategy, FLV video file cutting algorithm and partition file deleting strategy, we offer an online random seeking technology based on cluster stream media server. This technology could meet user's need well. Moreover, to make a better service, this technology could be enhanced in two aspects.

First, prefetching the file partitions should be provided by the media player. This function could be done at server side in order to reduce client's complexity.

Second, λ is settled in present LRFU algorithm and it doesn't change in the practical executing process. We can provide a dynamic adjusting method to make the strategy work better in later work.

In conclusion, in order to improve the overall performance of streaming media servers, first of all we should improve the disk concurrency from hardware and software, followed by improving CPU performance. So we can increase the overall streaming media server cluster's overall performance.

Acknowledgement. This research is supported by the National Natural Science Foundation of China (60773148, 60503039) and the streaming media subsystem of national outstanding courses integration project (No. JPKC-5).

References

[1] Shahabi, C., Zimmermann, R., Fu, K., Shu-Yuen, Yao, D.: Yima: a second-generation continuous media server. IEEE Computer 35(6), 56–64 (2002)
[2] Chai, Y., Wang, K., Li, S.: Study on high-end cluster streaming media server simulation. Journal of Chinese Computer Systems 29(7), 1185–1189 (2008)
[3] http://blog.lighttpd.net/articles/2006/03/09/flv-streaming-with-lighttpd (September 2010)
[4] http://search.cpan.org/~cdolan/FLV-Info-0.24/lib/FLV/Cut.pm (September 2010)
[5] Lee, D., Choi, J., Kim, J.-H., et al.: On the existence of a spectrum of policies that subsumes the Least Recently Used (LRU) and Least Frequently Used (LFU)policies. In: Proceedings of the 1999 ACM SIGMETRICS Conference on Measurement and Modeling of Computer System, pp. 134–143 (May 1999)
[6] Lighttpd (September 2010), http://www.lighttpd.net
[7] FASTCGI (September 2010), http://www.fastcgi.com
[8] Yu, H., Zheng, D., Zhao, B.Y., Zheng, W.: Understanding user behavior in large-scale video-on-demand system. In: Proceedings of the 1st ACM SIGOPS/EuroSys. European Conference on Computer Systems 2006, pp. 333–344 (April 2006)
[9] Dan, A., Sitaram, D., Shahabuddin, P.: Scheduling policies for an ondemand video server with batching. In: Proceeding of the Second ACM International Conference on Multimedia, pp. 15–23 (October 1994)
[10] Wessels, D.: Squid: The Definitive Guide. O'Reilly and Associates (January 2004)
[11] Tate, J., Lucchese, F., Moore, R.: Introduction to Storage Area Networks, 4th edn., IBM International Technical Support Organization (July 2006)

A Self-Adaptive Sliding Window Technique for Mining Data Streams

Yi Yang[1] and Guojun Mao[2]

[1] School of Computer Science, Beijing University of Technology, Beijing, China
[2] College of Information, Central University of Finance and Economics, Beijing, China
cnadrian@126.com, maximmao@hotmail.com

Abstract. The methods of data stream mining have recently garnered a great deal of attention in the field of data mining, and the sliding window technique has been widely used during many researches on it. This paper proposes a new type of self-adaptive sliding window (SASW) model, which has self-adjusting window parameters, and the technique details are presented under the ensemble learning method of single data stream environment. Experimental result shows that the definition of evaluating SASW parameters is appropriate and the gratifying results can be obtained. This idea also can be used in many other algorithms of data stream mining.

Keywords: -data stream mining, SASW technique, ensemble learning.

1 Introduction

The data stream, different from traditional databases, is a real-time, continuous, high-volume and open-ended sequence of data items. In many real-world applications, like sensor networks, traffic analysis, network intrusion detection, data streams have become a powerful information organization form, so building proper streaming mining models for these applications is becoming an important focus.

According to the nature of data stream, the problem of memory-efficiency has been arisen in stream mining models. In fact, a data stream is potentially unbounded but memory resource is limited, thus a mining model must support to find out hidden patterns by the restricted memory from continuous data streams.

The sliding window is one of the popular techniques for data stream mining [1]. It can discover so-far patterns using finite resources through restricting the number of items or the time range to the current window from a data stream. In a general way, the shorter the size of window is defined, the smaller memory is used and the higher memory-efficiency can be gotten. Of course, the longer size of the window can consider more data in a window and so can get a higher mining accuracy.

In addition, when using sliding window model during mining process, the problem of CPU-efficiency should also be noted. When a new window is created, it is necessary to drop some out-of-date data from the old window and add some new arrival data to the new one. However, the number of dropping or adding data for window changes in sliding windows will cause the problem of CPU-efficiency. In

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 689–697.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

general, the longer the dropping or adding steps, the less the times of window transformation for a data stream and the higher CPU-efficiency can be achieved. Comparably, the shorter steps of dropping or adding data can result in mining patterns in a more real-time way. Therefore, setting appropriate window size or moving step values for a sliding window model is a trade-off between mining accuracy and efficiency.

Sliding window models have widely been used in many existing researches of data stream mining [2, 3, 4, 5, 6], however, there are few studies about influence of different size or step settings in sliding windows. In fact, there should be different optimizing parameters of sliding windows for different data streams, that is, optimizing sliding windows is not a universal problem, so such optimizations should be adapted to the changes of a data stream over time. In this paper, we present a new type of self-adaptive sliding window (SASW) model, which can learn the sliding window control parameters that meet the demand given automatically when patterns are mined from data streams.

The rest of paper is organized as follows. Section 2 shows some definitions about the data stream and sliding window model. In section 3, we proposed SASW model and apply it into the ensemble learning of stream mining and give the descriptions on evaluating process of the parameters. Then an example is showed through our experiment in section 4. Finally section 5 summarizes our research.

2 Problem Statement

Given a data stream $S = < s_1, s_2, \cdots, s_t \cdots >$ where s_t is the data item that arrived at the time point *t*. As this paper will discuss the problem of classifying data streams, so a data item can be represented as $s_t = \{attrvec(s_t), class(s_t)\}$, where $attrvec(s_t)$ is the vector of condition attributes and $class(s_t)$ is the class label of s_t.

There are two basic strategies to design a sliding window model in data stream mining, i.e. count-based windows and time-based windows [1]. The former always maintains the same number of data items before and after window changes, while the latter use a fixing time interval to define the size of windows. This paper will focus on count-based sliding window, where the size of a window means the number of items that the window includes.

Let the size of a window *N*, given *k* in (0, 1], we can define a called "*k*-step" sliding window. There are two typical situations are being popularly applied for sliding windows: 1) $k=1/N$, removing only one data item from the previous window and supplying a new arrival item in order to create a new window [4]; 2) $k=1$, data of the old window are all removed and data of the new window are all new arrival items [2, 3, 5, 6]. Obviously, for a "*k*-step" sliding window, the number of removed items (say *step*), is $k*N$. Using the sliding window, we can collect the data in a window and the obtained data can be mined to get valuable classes by a fixing time point. Figure 1 shows a *k*-step count-based sliding window with the size of *N*. Each window collects *N* data items and the next window will be generated to replace the oldest $k*N$ items by new arrival ones.

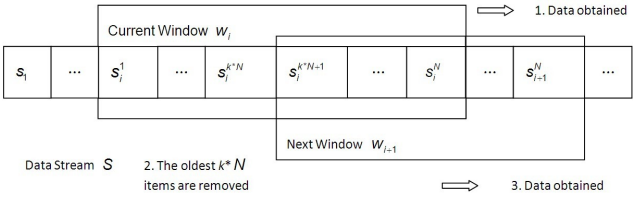


Fig. 1. Processing in a count-based & *k*-step sliding window

According to the existing researches, it can be guessed that when the *N* value is fixed, there should be different optimal *k* values for different applications and computing environments, and so dynamically optimizing *k* value for an ongoing data stream is very important. Moreover, if the value *k* can be learn and adjusted automatically with the changes of data stream, then better mining efficiencies or results may be able to be achieved.

3 A SASW-based Mining Model for Oriented to Ensemble Learning in Data Streams

As we stated, a SASW should adapt to the changes of the mined data stream, and can dynamically adjusted over time. Therefore, a better solution is that integrates adjusting SASW into learning patterns in a mining model to fit dynamic data streams.

In this section, we perform step-by-step analysis to our model, a mining model that integrates SASW technology to ensemble learning for data streams.

3.1 Model Overview

The basic processing framework of our model is described in Figure 2.

As Figure 2 shows, our model has mainly three phases. In phase 1, firstly enough new data is gathered from the data stream. Due to *k***N* items is needed for *k*-step SASW, so after collecting data, a new window can be formed by replacing k**N* data items in the previous window by new items.

In phase 2, a new classifier *NC* can be obtained. *NC* can be learnt by a number of popular classifying algorithms, and we used C4.5 to our experiments in this paper. We use the ensemble learning method in our model; this is because of its better results than single classifier for mining data streams, which have been proven by many researches [7, 8, 9]. Multiple basic classifiers, *M* classifiers at most in Figure 2, constitute an ensemble *EC*. Each classifier has a called weight between 0 and 1 to express its importance in *EC*. After a new window is created and a new classifier obtained, all classifiers in the ensemble will be evaluated and their weight can be updated. The weights of all classifiers can be dynamically changed according to the rates of correct classifying for data in the current window. Furthermore, if the number of classifiers exceeds the maximum defined number of classifiers in the ensemble (say *M* in Figure 2), the classifier with the lowest weight will be removed unless the newest classifier has the lowest weight.

Parameter description

S: the data stream;
W: a SASW (the current window);
N: the size of *W*;
k: control parameter for step of *W* (initial 1.0);
C: classifying method (algorithm, like C4.5);
EC: ensemble (all classifiers in *EC* are with method *C*);
EV: evaluation value of SASW with a fixing *k*;
α and β: the allowed lower and upper bound of *EV*s;
r: the adjustment factor of *k*(between 0 and 1);
ET: evaluation time of SASW in a fixing *k*(by number of windows, initial $ET_{1.0}$ is set by user);
M: the maximum number of classifiers in *EC*.

Model description

Phase 1.Create a new window
1-1. Collect *k*N* new arrival items from *S*;
1-2. Remove the earliest *k*N* data items from *W*;
1-3. Add the new *k*N* items to *W*.

Phase 2. Learn a new classifier and update the ensemble
2-1. Learn a new classifier *NC* in *W*;
2-2. Recalculate weights of all classifiers in *EC* by *W*;
2-3. If the number of classifiers in *EC* is less than *M*, insert *NC* into *EC*; Otherwise, insert *NC* and remove the classifier with the minimum weight in *EC* unless *NC* has the lowest weight.

Phase 3. Evaluate SASW and adjust *k* value
With fixing *k*, for *ET* windows, iteratively do:
3-1. Compute the average *EV* values by the current window (say *EV*).
3-2. If $EV > \beta$ or $EV < \alpha$, then adjust *k*;
3-3. If *k* is changed, recalculate *ET*.

Fig. 2. Framework of the model

In a nutshell, for *k*-step count-based sliding windows, adjusting parameter *k* roughly respects that changing SASW to get better mining qualities, and so adjusting the control parameter *k* value according to the evaluating result to current SASW will be conducted in phase 3. More details about this phase will be given in Part B of this section.

In fact, these phases can simultaneously work. That is, when a new window is being created, pattern learning is doing by the previous window, and the evaluate process is also running and the control parameter *k* of SASW can be adjusted at the same time if necessary.

3.2 Evaluating SASW by Computing EVs

In order to test whether the current SASW still adapted well to the changes of the data stream, a series of measurements must be taken.

Definition 1. Let the current window W_i, and given a classifier c, R_c^i , the correct classifying rate of classifier c in W_i can be computed by the percentage of the correctly classified items to all data items in W_i

Definition 2. Let the current window W_i and given an ensemble $E = \{c_1, c_2, \ldots, c_m\}$, R_E^i , the correct classifying rate of ensemble E in W_i, is defined as the maximal rate of all classifiers in E for W_i:

$$R_E^i = max(R_{cj}^i \mid j=1, 2, \ldots, m) \tag{1}$$

Definition 3. Let the current window W_i, and given a classifying method C and its ensemble EC. If a classifier NC is generated by method C, then $EV(W_i, C, EC)$, the evaluation value of W_i using method C in the way of ensemble learning, is defined as

$$EV_i = R_{NC}^i - R_{EC}^i \tag{2}$$

Definition 4. For a fixing k, given a series of k-step windows $w^k=\{w_1, w_2, \ldots, w_n\}$, and let a classifying method C and its ensemble EC. $EV(k, w^k, C, EC)$, the evaluation value of the k-step sliding window for data w^k using method C in the way of ensemble learning, means the average evaluation value of all windows in w^k:

$$EV(k, w^k, C, EC)=average(EV(W_i, C, EC) \mid i=1, 2, \ldots, n) \tag{3}$$

As far as data stream mining is concerned, the effectiveness of a k-step SASW to a data stream can be reflected by evaluation value of a set of objected k-step windows. In general, the smaller the EV value for a given k, such a k-step SASW can better match the changes of the data stream. The main process of Function $EV(k, w^k, C, EC)$ is given in Figure 3.

Function $EV(k, w^k, C, EC)$
1.set k; 2. FOR each window w_i in w^k DO 3. compute $EV(w_i, C, EC)$; 4.sum up $EV(w_i, C, EC)$ for get their average value; 5.END DO 6. RETURN the average of EVs for all windows in w^k.

Fig. 3. Computing process of evaluation values

3.3 Adjusting Control Parameter k in SASW

In Part A of this section we mentioned that adjusting of the control parameter k value according to *EV*s is an important process of our model. Here, there are two questions have to be answered: when and how parameter k of SASW is adjusted.

As Figure 3 shows, because the number of windows in w^k directly affect the computing efficiency and evaluating effectiveness of Function *EV* (k, w^k, *C*, *EC*), there is no need to check *EV* after every window change. Instead, we use a periodic evaluation time that is called as *ET* in Figure 2. In this paper, *ET* is directly related to the current k value.

Definition 5. For a fixing k, given an initial $ET_{1.0}$, *ET* value related to k, denoted as ET_k, is defined as:

$$ET_k = ET_{1.0}/k. \tag{4}$$

For Example, if the initial k value is 1.0, and the user sets $ET_{1.0}$=100, then when k value is adjusted 0.5, $ET_{0.5}$=200. As Figure 2 stated, if the current k=1.0, evaluation will be done and k can be adjusted after 100 windows disposed; if the current k=0.5, evaluation or adjustment for k can be performed after 200,so *ET* value means the number of windows before the checking of whether the current k value is suitable. When the k value is adjusted, the *ET* value is also recalculated.

Our definition of *ET* based on some basic considerations about stream mining: when the ensemble construction is processing, the value *EV* (w_i, *C*, *EC*) is also calculated continuously, because the data stream is a real-time model, some of the *EV* values may be abnormally compared to other results. For example, after disposing many windows with k=1.0, most computed *EV* values are between 3.5 and 4.0, but a few of them may exceed 5.0 or lower than 2.5, if the adjustment of value k occurs after every window change, such abnormal *EV* values may badly affect the result. So the *EV* value we need to evaluate the k value should be a statistical value which computed by many *EV* (w_i, *C*, *EC*) values to eliminate ill effects. In our paper we use the average value calculated by the values collected after the number of *ET* windows with fixed value k. Besides, the initial $ET_{1.0}$ value should be considered seriously. First, the *ET* value shouldn't be very small based on our analysis above. Second, when k is adjusted, the adjusted *ET* value is getting larger than it before, if it is too large, there may be taking a long time for the processing and hurt the performances of the mining model.

Another important issue we'll describe is how to adjust the k value. As can be seen from Figure 2, whether k value is need to be adjust depends on whether its calculated *EV* value is between the lower and the upper bound(say α and β in Figure 2) defined by user. Generally, if the *EV* value is greater than β, the k value should be decreased, and it should be increased while *EV* is smaller than α.

The setting of α and β value depends upon the considerations about not only the demand of mining effect, but also the potential risk occurs during the mining process: if the *EV* value is very large, it can be seen that the effect of current *EC* is not very well, so a threshold value (say β in Figure 2) of *EV* that meet the user's demand is set. Besides, in Part B of this section we mentioned that: in general, the smaller the *EV* value for a given k, such a k-step SASW can better match the changes of the data stream. But when *EV* is too small, it may cause the overfitting, that is, the *EC* is

overfit the items of new window. To avoid it, the decreasing of *EV* should be controlled, then a lower bound value α is defined to achieve this.

When the current k value is unsuitable, it needs to be adjusted as we mentioned above. An adjustment factor r which between 0 and 1 is defined to decide the length of k to adjust in each time: the increasing process of k is defined as $k=k*(1+r)$, while the decreasing process is $k=k*(1-r)$.

Based on the above analyses, Figure 4 gives the description of the adjusting method of the control parameter k in SASW.

Procedure of the adjusting of parameter k of SASW(ET_k ,k, $ET_{1.0}$,EV, α, β, r)

```
1.  FOR a parameter value k,when  ET_k  arrived, DO:
2.  IF  EV>β BEGIN
3.      Adjust k: k=k*(1-r);
4.      Recalculate ET_k : ET_k = ET_1.0/k;
5.      END
6.  ELSE IF EV<α BEGIN
7.      Adjust k: k=k*(1+r);
8.      Recalculate ET_k : ET_k = ET_1.0/k;
9.  END
10. ELSE   Current k value is suitable.
11. END DO
```

Fig. 4. The adjusting process of parameter value k in SASW

4 Experiment Result

To show the working process of SASW, we conduct an experiment based on the model in Figure 2. Our experiment was implemented using JAVA on a computer with Intel core 2 Q9450 2.66GHz and 3GB-RAM, the WEKA toolkit is also used in the programming.

The dataset used in our experiment is the *spambase* dataset, which has 57 continuous attributed and two type of class label. In order to simulate a data stream, we designed a stream generator with some control parameters such as flow speed etc., which can choose data items from the dataset randomly and continuously.

Some initial parameter values set by us are showed in Table 1.

Analysis of Results: In the first few rounds, because the *EV* value is much larger than the upper bound β, the k value has been decreasing. From round 9, the *EV* value is very close to β, and due to the lower bound value α, which is not very far from β(β-α=0.100 in our example), the k value is increasing and decreasing by turns during the next rounds. Finally in the 19th round, the *EV* value (say 2.389) is the first value between α and β. So its corresponding k value (0.409 as the result shows) is the final adjusting result.

Table 1. Initial Parameter Values in Experiment

Flow speed	100-200 items/100-200 ms
N	1000
Initial k	1.000
α and β	2.300 and 2.400
r	0.1
$ET_{1.0}$	20
C	C4.5 Revision 8
M	8

Table 2 shows a complete adjusting process of parameter k in our experiment, the ET_k and EV value calculated after ET_k windows for each k are showed in it.

Table 2. Results to Our Experiment

Round	k	EV	ET_k
1	1.000	4.098	20
2	0.900	3.891	22
3	0.810	3.702	24
4	0.729	3.511	27
5	0.656	3.304	30
6	0.590	3.097	33
7	0.531	2.902	37
8	0.478	2.695	41
9	0.430	2.489	46
10	0.387	2.280	51
11	0.426	2.477	46
12	0.384	2.265	52
13	0.421	2.463	47
14	0.380	2.261	52
15	0.418	2.440	47
16	0.376	2.241	53
17	0.414	2.419	48
18	0.372	2.236	53
19	**0.409**	**2.389**	48

5 Conclusion

In this paper, we proposed a self-adaptive sliding window (SASW) technique for mining data streams based on our "k-step" window idea, then some technical details

such as the evaluating and adjusting process of the SASW control parameter are showed through an ensemble learning model with our SASW model supported to. The SASW technique has a universal use when mining data streams, e.g. if we use SASW in the local sites when mining distributed data streams, the quality of the mining of local sites can be improved and the maintaining of global schema can be very convenient than ever before.

Acknowledgement. This work has supported by Beijing Municipal Key Laboratory of Multimedia and Intelligent Software Technology, and the National Science Foundation of China under Grants No.60873145.

References

[1] Babcock, B., Babu, S., Datar, M., Motwani, R., Widom, J.: Models and issues in Data Stream Systems. In: Proc. The 21st ACM SIGMOD-SIGACT-SIGART Symp. on Principles of Database Systems (PODS 2002), pp. 1–16. ACM Press (June 2002), doi:10.1145/543613.543615

[2] Hulten, G., Spencer, L., Domingoes, P.: Mining Time-Changing Data Streams. In: Proc. The 7th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD 2001), pp. 97–106. ACM Press (August 2001), doi:10.1145/502512.502529

[3] Babcock, B., Datar, M., Motwani, R., O'Callaghan, L.: Maintaining Variance and K-Medians over Data Stream Windows. In: Proc. The 22nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS 2003), pp. 234–243. ACM Press (June 2003)

[4] Chang, J., Lee, W.: estWin: Online data stream mining of recent frequent itemsets by sliding window method. Journal of Information Science 31(2), 76–90 (2005), doi:10.1177/0165551505050785

[5] Chi, Y., Wang, H., Yu, P.S., Muntz, R.R.: Catch the Moment: Maintaining closed frequent tuplesets over a data stream sliding window. Knowledge and Information Systems 10(3), 265–294 (2006), doi:10.1007/s10115-006-0003-0

[6] Gibbons, B., Tirthapura, S.: Distributed streams algorithms for sliding windows. In: Proc. The 14th Annual ACM Symposium on Parallel Algorithms and Architectures, pp. 63–72. ACM Press (August 2002)

[7] Street, W., Kim, Y.: A streaming ensemble algorithm (SEA) for large-scale classification. In: Proc. The 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2001), pp. 377–382. ACM Press (August 2001), doi:10.1145/502512.502568

[8] Gianluigi, F., Clara, P., Giandomenico, S.: An adaptive distributed ensemble approach to mine concept-drifting data streams. In: Proc. The 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007), pp. 183–187 (October 2007), doi:10.1109/ICTAI.2007.51

[9] Masud, M.M., Gao, J., Khan, L., Han, J., Thuraisingham, B.: A Multi-partition Multi-chunk Ensemble Technique to Classify Concept-Drifting Data Streams. In: Theeramunkong, T., Kijsirikul, B., Cercone, N., Ho, T.-B. (eds.) PAKDD 2009. LNCS, vol. 5476, pp. 363–375. Springer, Heidelberg (2009), doi:10.1107/9-78-3-642-01307-2_34

The CAD Solutions in Highway Landscape Greening

Yin Guangli[1,*] and Qian Shao[2]

[1] Binzhou Vocational College, Sub-College of Information Engineering Binzhou, ShanDong, China
[2] The College of Traffic Management Cadres, Department of Computer Science Institute Beijing, China
yin_3278051@163.com

Abstract. 3D and planar design software, as an indispensable aiding tool for the design industry, is widely used in architectural and landscape designs. This paper centers on the CAD-based study of comparatively important issues existing in freeway landscape designs - landscape "points" (entrances, exits and architectural elements), "lines" (roadside green belts) and "planes" (protecting slopes and rock walls) - in association with actual cases, and brings forward some exercisable design methods and philosophies.

Keywords: Computer aided design, Freeway landscape, Designing method.

1 Status Quo of Freeway Landscape

Freeway is playing an increasingly important role in our daily life. Same increasing is people's expectation of roadside landscape and functions. As a result, study of freeway landscape has become an important program related to multiple subjects.[1] With the furthering of freeway construction in China, roadside landscape problems are gradually exposed: on the "point" level (including entrances & exits, architectural sculptures), the landscape entities are devoid of aesthetic feelings; the cultural nature, spatial scale and building materials are disorderly and unsystematic; and the architectural elements in series are unreasonably spaced; on the "line" level (the linear landscape formed by roadside vegetation), improper vegetation seed brings in survival difficulties of trees or even invasion by foreign species, and spacing between and size of the vegetation are inappropriately schemed; on the "plane" level (the continuous plane landscape formed by protecting slopes and rock walls), both the above-mentioned problems and deformation of the plane image under high driving speed exist.

To solve these problems, investigation and analysis of current freeway landscape are necessary. Besides, it is advisable to experiment on the landscape design projects of freeway sections, and use the experimental results to guide the rational landscape planning of other areas.[2] However, due to the giant workload of freeway construction, function and safety are the first priorities to be considered, and it is unrealistic to take one road section as the experimental materials for landscape design.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 699–706.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Therefore, we apply to computer modeling to simulate the actual freeway landscape, and make analysis and evaluations based on it to work out the best landscape design methodology and philosophy.

2 Computer Aided Problem Analysis

2.1 Experimental Subject

In this study, the author takes two ongoing freeway landscape design cases as the experimental subjects, which are (I) The Cultural Construction for Chifeng City's Freeway in Inner Mongolia Autonomous Region (Project source: Communication Culture Research Institute of Transport Management Institute, Ministry of Transport of PRC); (II) Traffic Safety-based Study of Key Technologies in Freeway Landscape Designs (Project source: China Academy of Transportation Sciences, MOC).

2.2 Research Methodology

2.2.1 Preparation of Experimental Materials (Landscape Entities)

Use SU to build the model for the roadside landscape design schemes, and create 3D animation based on the model to simulate the dynamic effect observed by people in running vehicles on the freeway. For the detailed parts of landscape entities, use PS to visualize the effect drawing of the landscape model exported by SU; then carry out all-practical imitation experiments. In the process of animation and picture creation, several experimental groups are established as per the variables of running speed, size, spacing and appearance of landscape entities. Each group has one variable (for example, different running speed for the same landscape section, or different spacing between landscape points under the same running speed). In the same experimental group, subgroup and number by gradient the experimental results under the same variable, then make an overall appraisal of the landscape design methods based on the different scenarios, and finally work out a function-like method.

2.2.2 SD Method

"Semantic differential" (SD) method is an investigation and analytical method widely used in psychology, market survey and cultural study. This study applies the SD method to qualitatively analyze the experimental materials, and determines the quantification rules of every design segment in association with the SD analytical result under every single variable. In the SD method, questions and scales are two essentials. One scale is divided into seven grades, with two adjectives completely contrary to each other as the two ends. For example, the seven grades may be: (i) "excellent", (ii) "very good", (iii) "good", (iv) "normal", (v) "bad", (vi) "very bad", (vii) "extremely bad". Participants of the experiment shall select one of the seven options based on their emotional evaluation upon the raised issues (here refers to the animations and pictures). Their evaluations will be used for factor analysis. SD method is advantaged by straightforward results, flexible application, and convenience for conceiving and scoring. [3] To get the accurate and authentic feeling

of the simulation scene, several regular drivers and passengers of the exact freeways were invited in the research experiment.

2.2.3 Conclusion

With the experimenters' feelings and evaluations of the SD method-based experiment, we can decide the favorite scenario and favorite landscape design for freeway users, and put forward certain design guidelines on this basis.

3. Problem Solving

3.1 Problem Solving

3.1.1 Entrances and Exits

Entrances of many freeways are designed as the "name card" of the road section by erecting landmark points fitted in the surrounding greenbelts. As the first highlight for the roadside landscape of the section, the design of this "point" deserves special attention.

One subject of this experiment is an entrance to the Hegang-Dalian Freeway of project (II), where there is a triangle greenbelt to be placed with a landmark landscape. In the designing period, several elements with the connotation of Jilin Province's history and culture came into the picture, including North-east tiger, deer, ginseng, ship… The former three are Jilin City's specialties. The "ship" idea originates from the old name of Jilin – "Jilin Wula", which in the Manchu language means cities along rivers. Based on the different design ideas, different landscape entities were designed and effect drawings were created. Then all the design alternatives were analyzed based on the SD method, and finally the "ship" plan was selected. After the outline was fixed, shape and material of the ship were put on the agenda, followed by the illustrations of the features of different materials: metal sculpture is florid but too expensive and corrosion-prone; and stone sculpture is on the contrary.[4] Through SD analysis, the plan as shown in "Figure 1" was confirmed: plums were dry laid to the ship-shape, with metal mast hung by eye-catching slogans.

Fig. 1. Plums were dry laid to the ship-shape

Therefore, we can sum up the requisite features for such entrance landscape types: (i) eye-catching; (ii) representing local culture; (iii) graceful and durable.

3.1.2 Architectural Elements and Sculptures

With the moving of the vehicle, all the landscape points of the freeway show up one by one. The optimum spacing between the points and the size of single points are evaluated as following: by SU modeling, a one-kilometer road section is created, with landscape elements decorating roadside greenbelts. To achieve the intelligible and intuitionist effect, a simplified global sculpture is placed on the central greenbelt, and a cultural totem pole is placed on the outer side. Adjust the observation angle of the model according to driver's visual angle (see Figure2), then export several animations under different running speed as the dynamic scenes observed by drivers and passengers.

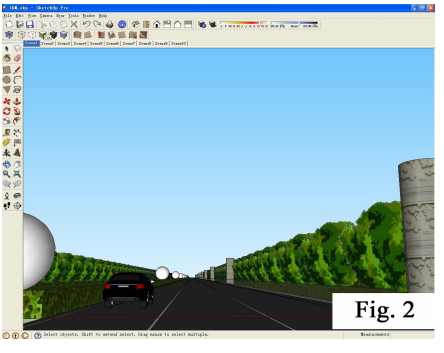


Fig. 2. Driver's visual angle

In the first experimental group, set the average running speed at 90km/h (25m/s), then select the best landscape scene based on the comfort level sensed by people under different spacing and sizes of points. The concrete steps are: in the SU model, set one scene every 100 meters along the travel direction of the one-kilometer section (SU can export animation through scene switches), and set the switching time of every two scenes at four seconds (i.e. to control the advance speed of sight at 25m/s); zero the dead time of every scene (see Figure-3). In this way, the animation effect seen by experimenters is equivalent to running smoothly at 90km/h on the road section. In this group, make constant the speed and landscape size, and choose the spacing between totem poles and global sculpture as the variable (In Figure-2, the spacing is 100m). Then separately adjust the spacing to 50m, 150m and 200m, and create the corresponding animations. Play the animations simultaneously and analyze by the SD method the comfort level of each spacing sensed by experimenters. In this way, the optimum spacing is achieved.

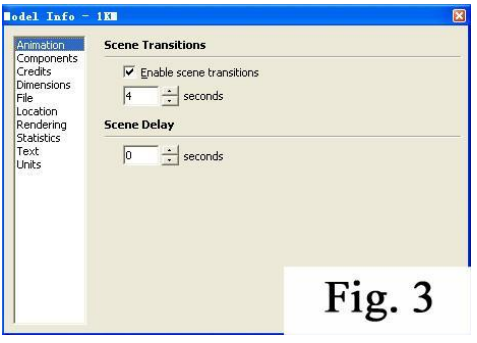


Fig. 3. zero the dead time of every scene

After the above experiment, let's assume the optimum spacing is 100m, and evaluate the comfort level under different running speed (set the speed at 80km/h, 90 km/h, 100 km/h, 110km/h and 120km/h respectively, and select the speed that provides the highest comfort level). Put these two experimental groups into overall consideration, we can get the different optimum spacing of landscape points of road sections with different speed limits. Same, we can get the optimum size, material and appearance of single landscape points. With these two experiments, we can educe the landscape design principle closely related to landscape observers.

Research says that to clearly see the landscape on freeway, the minimum fixation time is five seconds.[5] Therefore, during this period of time, the travel distance of vehicle is the length of landscape, i.e. the spacing of landscape points or the width of landscape planes (see below). Then we get the functional relation as shown in Figure4.

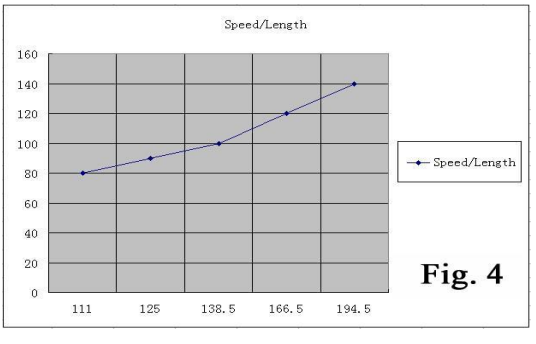


Fig. 4. The spacing of landscape points

3.1.3 Designs of Landscape "Lines"

Generally speaking, there are three considerable factors for the selection of freeway roadside vegetation: firstly, to match tree species with site and focus on indigenous tree species; secondly, to allocate the suitable proportion of fast-growing species and slow-growing species, and of deciduous species and evergreen species; thirdly, to make an overall planning of vegetation colors and the space configuration of trees, shrubs, vines, grass and flowers to enhance the layer sense of the landscape.[6]

The experiment chose a road section of Project (1), where the weather is dry and cold. Suitable roadside trees are poplar, ceiba and dryland willow. Site pictures were PS processed to the effect with different roadside tree species, and then the best scene was chosen. It shall be noted that some road sections used sea-buckthorn as the roadside vegetation, which is characterized by fast growth and no natural enemies. Consequently, it grew out of control and caused influences to traffic safety. Therefore, we must put the ecological principle into consideration when we choose tree species, and use those whose growth can be artificially controlled. Trees' size and planting density shall also be considered to provide the antidinic function. With reference to the above-stated methodology and functional relation, we designed different sizes of trees in the SU model, then exported the animation and pictures, and chose the best tree size based on the SD method. With this best size, we got the best planting density by the same procedure. Different road sections and areas lead to different conclusions. Before conducting detailed designs, we shall make a brief investigation into local features and tailor the most appropriate design scheme.

One important issue is that in creating pictures and animation, we shall attach same importance to both the appearance and the authenticity; otherwise the designs will be "empty talks". See Figure5, the first picture is the actual scene of the experimental case; the second one uses poplar as the roadside tree and plants them based on the real proportion; but in the third picture, emphasis on shape and color outweighs that on the appropriateness of roadside trees. Both the second and the third pictures are design schemes, and the third one is superior in aesthetic measure. However, this beautiful scene is infeasible locally, which makes the second one the reasonable design.

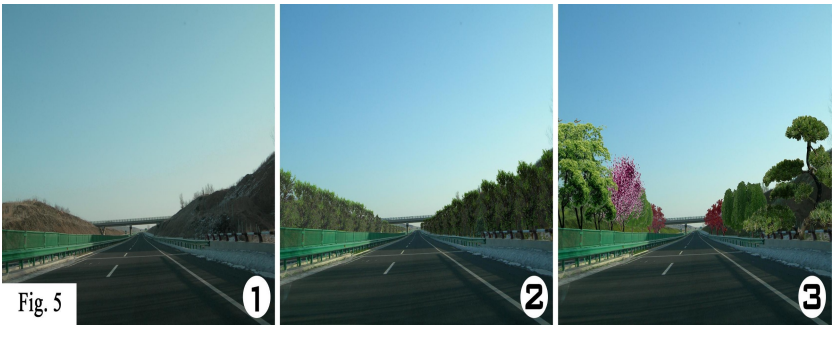


Fig. 5. Both the appearance and the authenticity

3.1.4 Designs of Landscape "Planes": Design of Rock Wall

The experiment uses another freeway section of Chifeng City as the subject. Unique for the section, one side of the road is rock wall (See Figure-6 at left), open for beautiful fresco landscape. Actually, it is unusual for freeway to have rock walls as landscape. But in this case, it is the principal part of design. To paint an already designed picture on the rock wall is not difficult. Difficulty lies in how to determine the planar pattern and measure of the picture based on the visual changes of drivers or passengers, since static normal observation of the picture is different from the dynamic profile view from running vehicles.

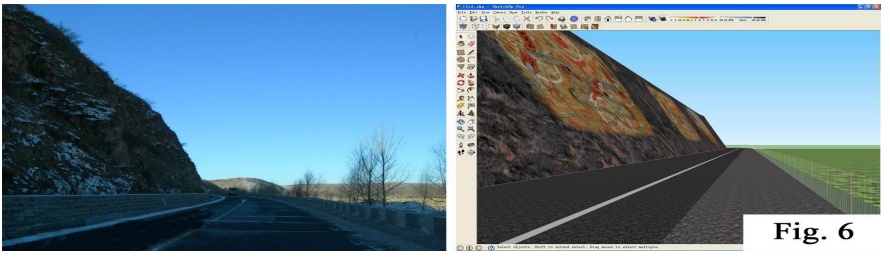


Fig. 6. Create a scene model

Here we shall simulate a dynamic observation of the landscape and refer to the evaluation method of landscape "points" to create a scene model (See Figure-6 at right) with the aid of computer software: paint the planar picture on the rock wall, use picture size or spacing as single variables, divide the designs into several groups and make comprehensive analysis with the SD method. The conclusion is that to achieve the best visual effect of rock wall under fast running circumstances, the picture must be deformation treated based on its perspective principles before being painted. For example, to alleviate deformation, the "big when near and small when far" theory may be applied to have the picture "small when near and big when far" treated; or widen the picture or repeatedly show several pictures at a certain spacing to represent the continuity. The experimental figures closely related to drivers and passengers are:

Table 1. The experimental figures closely related to drivers and passengers

Designed driving speed (km/h)	80	100	120
Visual angle of drivers or passengers (°)	30	20	12
Distance of attention concentration points (m)	300	420	540

With the reference to the functional relation of Figure-4, we use the 1km-section as the experimental subject, and can see that under the normal driving speed on freeways (usually about 100km/h), rock wall pictures simply flash by. Therefore, it is not recommended to paint a complicated picture for such a fleeting observation, because it is simply the waste of construction time and materials. Here we propose simple cultural symbols at the minimum spacing of 100m and the size of 1.5-2 times of normal plan view size (calculated based on the experimental figures in the above table and the Pythagorean proposition, and validated visually by the SD method). In addition, the pictures selected shall also reflect local cultures. In Figure-6, the picture is the lucky emblem originating from the ancient Yuan and Liao dynasties during which Inner Mongolia had a glorious history.

4 Conclusion

With the aid of computer software, we can simulate the landscape of freeway; with the SD method, we bring forward following rational proposals for freeway landscape designs:

Landscape design must not override highway construction standards and transportation rules. Function and safety must be given the first priorities. For example, the landscape design shall not be too flamboyant, otherwise it will distract drivers' attention.

- Landscape materials must be durable and environmental friendly.
- Take fully into account the psychological needs and visual demands of drivers and passengers.
- Landscape designs shall be cultural and vivid, and strive for changes in the unification.

As a forceful validation of the research conclusion, the landscape design schemes of the cases referred in this paper are highly favored by the entrusting parties, and have been gradually executed. The cases of this study are from actual projects. Highlights lie in the vivid simulation of practical scenes with the aid of computer technology, the accurate and rapid analysis process, and the significance of the conclusions drawn. Surely, this research methodology still has room for improvement with the upgrade of software technology (so that more exquisite model may be created, or 3Dmax or Maya will be applied to produce more authentic and flamboyant simulation scenes), and the detailing of gradient grouping by experimental variables.

References

[1] Chen, J.: On the Landscape Design of Freeway. Technology and Market (03), 29 (2010)
[2] Qian, G.: Innovative Design for Expressway. China Communications Press (2008)
[3] Zhuang, W.: Evaluation of Enclosed Public Space Objectives based on the SD Method. Journal of Tsinghua University (Science and Technology) (36), 42–47 (1996)
[4] Li, E., et al.: Ecological Landscape Design for Jilin-Yanji Freeway. Transportation Construction & Management (01) (2008)
[5] Qian, G., Tang, S., et al.: Landscape Design for Freeway. China Communications Press (2009)
[6] Whittaker, R.H.: Gradient analysis vegetation. Biol. Rev. 49 (1967)
[7] Li, T.: Analysis on Issues about Landscape Design of Highway. Science Technology and Industry (08), 86–89 (2009)
[8] Shen, X., et al.: The Simulation System for Freeway Landscape Design. Transportation Science & Technology (03) (2007)
[9] Luo, Z.: Sketchup Tutorial. Posts and Telecom Press (2008)
[10] Tang, J., Huang, Y.: Node Design of Trunk Highway Landscape. Transpoworld (24), 122–123 (2009)
[11] Xi, G.: SketchUp and VRML-based Virtual Tour System Development. Computer Knowledge and Technology (30) (2009)
[12] Feng, L., et al.: Research of Highway Landscape Design Thoughts Based on Moving Features. Journal of China & Foreign Highway (06), 235–239 (2009)
[13] Wolf, K.L.: Freeway roadside management: The urban forest beyond the white line. Journal of Arboriculture (29) (2003); TRB: Proceedings of the 3rd International Symposium on Highway Geometric Design. USA (2005)

Using Genetic Algorithm to Improve Gray Watermark Algorithm

Jianbing Yi and Qiuen Xie

School of Information Engineering, Jiangxi University of Science & Technology, Ganzhou 341000, China
yijianbing8@163.com

Abstract. Robustness and invisibility as two evaluation indicators for watermark algorithm are mutual restraint, this paper employs genetic algorithm to search for the embedded intensity α and the frequency of Arnold scrambling T, with the gray watermark image being embedded in the SVD of image carrier, aiming at achieving a combined optimization of the robustness and invisibility. Experimental results show that the proposed watermarking algorithm not only contains enormous watermark capacity, but quite safe, with invisibility being good and robustness being sound.

Keywords: Gray watermark, Genetic algorithm, Invisibility, Robustness, Chromosome.

1 Introduction

Digital watermark technology as one of the effective means for copyright protection and information security has been a growing research theme. Robustness and invisibility are two fundamental evaluation criteria for digital watermark algorithm[1]. Robustness calls for retaining a good extractability while experiencing general signal processing and common attacks. While invisibility requires the original carrier image of embedded watermark to lose no clear perception distortion. However, these two indicators are mutually restraint. Traditional watermark algorithms often judge whether the selected embedded parameters meet the two indicators by experimental results. Therefore, once, the selected parameters are inappropriate, they should be reselected, and tested again. The process is very cumbersome and finding the best parameter is difficult. This paper presents a genetic algorithm-based embedded watermark technique, which adopted genetic algorithm to search for the embedded intensity and frequency of Arnold scrambling to achieve the overall optimization on watermark robustness and invisibility.

2 Watermark Embedding and Extraction

Ordinary watermark information is normally a binary image or a simple text, expressing insufficient meaning[2]. Therefore, a watermark algorithm is proposed in

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 707–714.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

which gray images are regarded as watermark information. The core of the algorithm is to implement wavelet transform on the blue component of the original image[3], generates multiple sub-band wavelet coefficients, each of which will be divided into different blocks, then to carry out SVD decomposition. The gray watermark is bit-plane decomposed first, and then is embedded in the SVD of each sub-band wavelet coefficient according to the weights of the different coefficients. In Fig. 1, Bit0 to Bit7 illustrates 8-bit plane decomposition of a gray image which is embedded in the each sub-band wavelet coefficient.

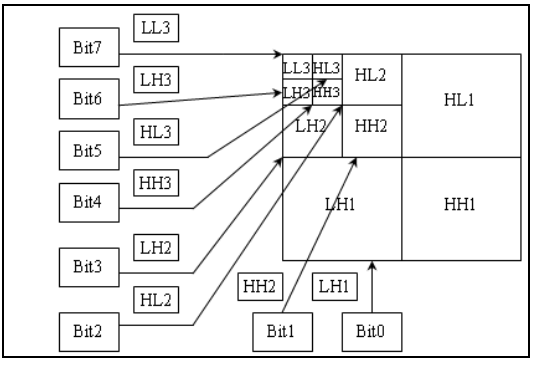


Fig. 1. Bit plane graph is embedded

2.1 Watermark Embedding Algorithm

Because the sensitivity of vision to the blue component of original image is not high, the carrier of watermark embedding selects the blue component of the original image. Watermark embedding process is as follows:

Step 1: Extract the blue component of the original image, and apply three-level wavelet decomposition.
Step 2: Arnold scrambling is conducted on the gray watermark, and then the scrambled watermark is bit plane decomposed.
Step 3: To ensure data security, chaotic encryption is executed.
Step 4: According to the different weights of each bit-plane derived from bit-plane decomposition of gray watermark, embedded the high weights of bit-plane in the critical sub-band wavelet coefficients. For example: the bit7 is embedded in approximation sub-band wavelet coefficients of the third layer. Specific ideas embedded as follows: each sub-band coefficients respectively were divided into 2*2, 4*4 matrix of blocks depending on the circumstances, then the information of each bit of corresponding bit plane is embedded in its SVD of matrix blocks[4].
Step 5: Apply inverse wavelet transform.

2.2 The Watermark Extraction Algorithm

The watermark extraction is the reverse process of watermark embedding. Specific process is as follows:

Step 1: Extract blue components from image containing watermark, apply three-level wavelet decomposition.
Step 2: All sub-band coefficients, according to the embedded situation, are divided into 2*2, 4*4 matrix blocks one by one, SVD decomposition is applied on each block, to extract watermark.
Step 3: Apply chaos decryption for the extracted watermark.
Step 4: each bit plane extracted are synthesized gray watermark.
Step 5: Arnold reverse scrambling on the gray watermark.

3 The Proposed Method

As frequency of Arnold scrambling T determines the degree of watermark scrambling, just as the embedding intensity will directly affect the robustness of watermark and the quality of watermarked image[5]. In order to improve robustness at most, it often anticipates the bit of scrambled watermark to be distributed more chaotic, But the process of applying scrambling technique also witnessed that if the location deviation emerges in the extraction process of digital watermark, Instead, the increasing frequency of Arnold scrambling will yet greatly reduce the visual effect of digital watermark. As a result, during the watermark algorithm, to seek a suitable frequency of Scrambling T and the watermark embedding intensity α has become a key issue to be solved.

Genetic algorithm is a bionic algorithm conceived to search the optimal solution based on principles of biological evolution[6], which simulated the natural process of genetic recombination and evolution, parameters used to solve the problem are compiled to binary code or decimal code-gene, several genes form a chromosome, chromosome is similar to natural selection, matching crossover and mutation operations, After several iterations, it reached the final optimization results[7]. Traditionally, the greater the fitness value, the better solution is. Population-based genetic algorithm, not a single point search-based, could attain many extremes from different points, thus avoiding local optimum[8]

3.1 Chromosome Encoding and Decoding

Genetic algorithms generally use a fixed-length string of binary symbols to represent individuals out of groups, the allele is composed of binary value of (0, 1). Each individual gene of initial population can be generated by uniformly distributed random numbers. This case involves two variables, α and T, of which α is represented by 11 bits, while T is represented by 5 bits; So chromosome structure can be described by the method illustrated in Fig. 2.

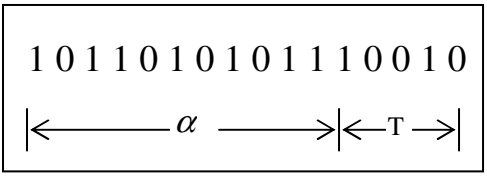


Fig. 2. Chromosome structure

3.2 Fitness Function

During the optimized embedding process of the watermark, peak signal to noise ratio (PSNR) of the image and the similarity of the extracted watermark (NC) are considered as an evaluation criterion of watermark invisibility and robustness. Optimization of the watermark embedding can be viewed as a multi-objective optimization problem, using GA to search frequency of Arnold scrambling T and embedding intensity α in the global context, making the PSNR and NC to be overall optimized. Algorithm based on the degree to relative significance of objectives are weighted, converted the multi-objective optimization problem into single objective optimization problem to work out the solution. The fitness function is represented such as (1).

$$Maxf(\alpha,T) = \omega_1 * PSNR(\alpha,T) + \frac{\omega_2}{n} * \sum_{i=1}^{n} NC_i(\alpha,T) \tag{1}$$

Where ω_1, ω_2 represents the weight of PSNR(α, T), NC(α, T) respectively.

Proportional Selection Operator

Proportional selection factor is to employ the probability based on the proportion to each individual's fitness to determine their chances to contributing children's genes. If the assumed population size is M, the fitness of individual i is f_i, then the probability of individual i is selected as (2).

$$p_i = f_i / \sum_{k=1}^{M} f_k \tag{2}$$

When the probability of individual selection is presented, uniform random number between [0, 1] is generated to decide which individual participates in mating. If the probability of individual selection is high, it can be repeatedly selected, its genes will be expanded in the population; If the probability of individual selection is low, it is to be eliminated.

Genetic algorithm optimization is as follows:

Step 1: Generated 20 initial chromosomes randomly (S);
Step 2: Set population size algebra D = 0, population scale N = 10;
Step 3: Proportional selection method adopted to select N chromosomes from S, Constitute S1;
Step 4: Determine the fitness value of each chromosome;
Step 5: If the predefined conditions meet or repeat iteration to achieve a stable constant , then keep the best chromosome, the algorithm stops and turns to Step (11);
Step 6: N/2 pairs of S1chromosomes is to be two-points crossed by probability 0.5;
Step 7: genetic variation emerges to generate 10 new chromosomes according to the 0.04 probability, stock into S2;
Step 8: Calculated fitness value of each chromosome of S2;
Step 9: S1+S2→S, D+1→D;
Step 10: turn to Step (3);

Step 11: Keep chromosome with the highest fitness value correspondingly to determine T, α according to the encoding rules;
Step 12: Termination.

This method offers the following two conditions as the conditions for the evolution termination: (a) frequency of evolution iterations is greater than the upper bound of D (200); (b) the fitness value of consecutive 10 generations of populations is not changed. Meet anyone of them, the algorithm terminated. Algorithm optimization process is shown in Fig. 3.

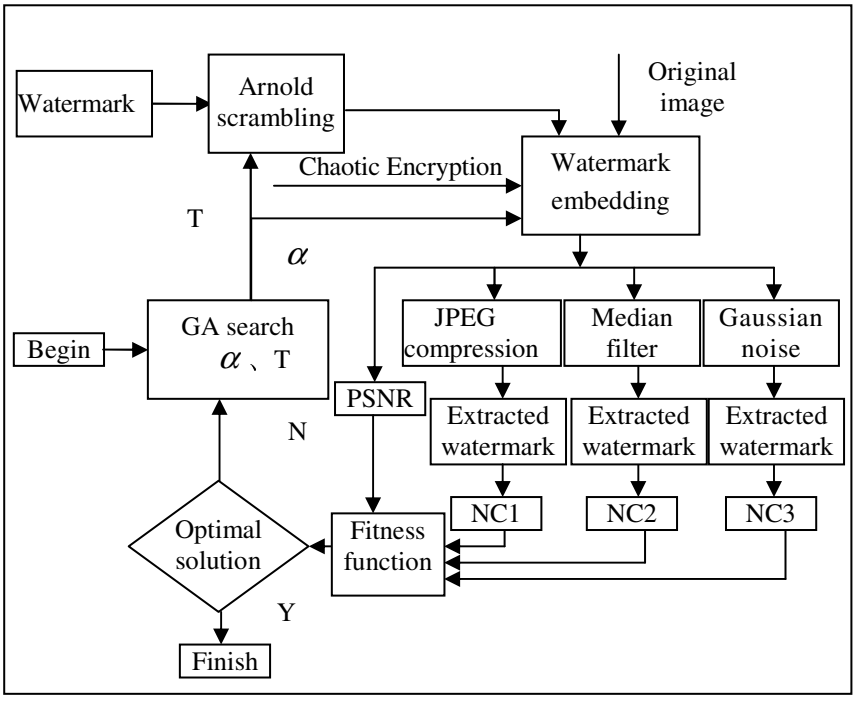


Fig. 3. Flow chart of using genetic algorithm to optimize the watermark

4 Experimental Simulations

The experimental image is a 512×512, three bmp standard test image such as Lena, Baboon and Winter. The experimental watermark is a 32×32 gray watermark image. Simulations are implemented under matlab-7.0 circumstances.

Genetic Algorithm Optimization Experiment
During the experiment, we adopted three kinds of noise attacks, like 90% JPEG compression, 3×3 median filter, and 1‰ Gauss noise to extract the watermark similarity respectively to calculate the fitness function of genetic algorithm. NC1 is the watermark similarity extracted after compression with 90% JPEG, NC2 is the

watermark similarity extracted after 3×3 median filtering, and NC3 is the watermark similarity extracted after attacked by 1‰ Gauss noise. Optimal parameters of three images are searched. The results are listed in table 1.

Table 1. The Results of Optimal Parameters

Image	α	T
Lena	111.8333	24
Baboon	110.6721	17
Winter	110.2264	5

4.1 Robustness and Invisibility Experiment

Watermark algorithm evaluation embraces invisibility and robustness.

Invisibility evaluation is determined by the contrast between the original image and the watermarked image. Visibility indicators are employed to subjectively evaluate the difference between images embedded before and after from visual perceptions, such as Fig. 4, Fig. 5, and Fig. 6. Quantitative objective invisibility evaluation often adopts peak signal to noise ratio to assess. According to human visual characteristics, *PSNR*> 30, the watermark invisibility is much better.

Fig. 4. Lena PSNR=34.2851,NC= 0.9978: (a) original image and watermark, (b) watermarked image and extracted watermark

Fig. 5. Baboon PSNR= 34.5634,NC= 0.9922: (a) original image and watermark, (b) watermarked image and extracted watermark

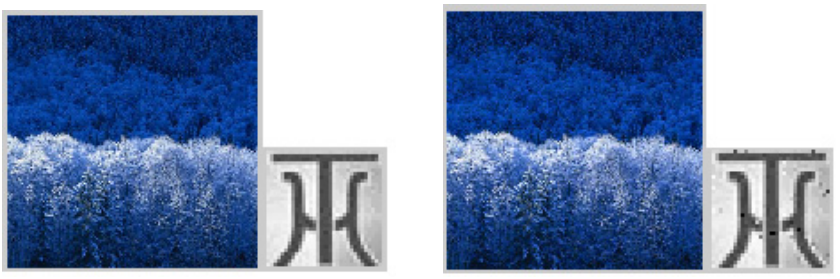


Fig. 6. Winter PSNR= 34.3004,NC= 0.9958: (a) original image and watermark, (b) watermarked image and extracted watermark

Robustness evaluation primarily uses several typical noise attack methods. Table 2 describes the results of watermark similarity and peak signal to noise ratio adopted to calculate GA fitness in noise anti-attack experiments. Table 3 demonstrated the results of extracted watermark similarity after attacked by other common noises.

Table 2. Psnr and NC Values of Three Tested Images

Image	PSNR/dB	NC		
		90%JPEG compression	3×3 Median filtering	1‰ Gauss noise
Lena	34.2851	0.9538	0.9429	0.9608
Baboon	34.5634	0.9269	0.8689	0.9493
Winter	34.3004	0.9301	0.8900	0.9715

Table 3. Other NC Values of Three Tested Images

Image	NC			
	70%JPEG compression	5×5 Median filtering	10‰Salt and pepper noise	Right 1 /4 shear
Lena	0.9321	0.8926	0.9008	0.9762
Baboon	0.8713	0.7721	0.8918	0.9757
Winter	0.8846	0.7721	0.9233	0.9718

4.2 Genetic Watermark Algorithm Analysis

The watermark algorithm uses genetic algorithm to search optimal combination of invisibility and robustness to cope with number of difficulties in which traditional watermark algorithms select embedding intensity and frequency scrambling. It not only saves time but also improves robustness and invisibility.

5 Conclusion

This paper regarded gray images as the watermark information to ensure that the embedded watermark capacity is relatively large; used chaotic encryption to resolve

the embedded information security; introduces genetic algorithm to search the embedded intensity and frequency scrambling of digital watermark algorithm, thus, a comprehensive optimization purposes of watermark robustness and invisibility has been achieved. Experimental results showed that the proposed algorithm can guarantee the robustness of watermark with the damage of original images being minimized. It ensured visual effects after the watermark was embedded, enriched watermark information, and ultimately improved security and robustness.

References

[1] Hsu, C.T., Wu, J.L.: Hidden digital watermarks in images. IEEE Trans., Image Process 8(1), 58–68 (1999)
[2] Pan, W., Li, B., Ma, Y., et al.: Class Structure Refactoring of OO Softwares using Community Detection in Dependency Networks. Frontiers of Computer Science in China Special Issue on Complex Networks and Computing 3(3), 396–404 (2009)
[3] Pan, R., Gao, Y.X.: A new wavelet watermarking technique. In: Proceedings of the 4th World Congress on Intelligent Control and Automation, vol. (3), pp. 2065–2069 (2002)
[4] Liu, R., Tan, T.: An SVD-based watermarking Scheme for protecting rightful ownership. IEEE Trans., Multimedia 4(1), 121–128 (2002)
[5] Kumsawat, P., Attakitmongcol, K., Srikaew, A.: A new approach for optimization in image watermarking by using genetic algorithms. IEEE Transactions on Signal Processing 53(12), 4707–4719 (2005)
[6] Holland, J.H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor (1975)
[7] Lee, D., Kim, T., Lee, S., Paik, J.: Genetic algorithm-based watermarking in discrete wavelet transform domain. Intelligent Computing 4113, 709–716 (2006)
[8] Shih, F.Y., Wu, Y.T.: Robust watermarking and compression for medical images based on genetic algorithms. Information Sciences 175, 200–216 (2005)

A New Secure Windows Terminal Based on Dual-Subsystem Architecture

Li XinMing[1], Li Yi[2], Zhang GongXuan[3], and Wang XiaoMei[4]

[1] EIES Key Laboratory, The Academy of Equipment, Beijing 101416, PR of China
[2] Department of Scientific Research, The Academy of Equipment, Beijing 101416, PR of China
[3] College of Computer Science & Technology, Nanjing University of Science & Technology, Nanjing 210094, PR of China
[4] Company of Postgraduate Management, The Academy of Equipment, Beijing 101416, PR of China
liyi1221@263.net

Abstract. A new secure Windows terminal based on heterogeneous dual-subsystem architecture is proposed to resolve the security problems of Windows terminals. The secure terminal includes a prime subsystem and an assistant subsystem. The prime subsystem, which handles man-machine interactive and mission-related programs, is compliant to Windows. The assistant subsystem, which provides main security mechanisms as security machine, is closely connected with prime subsystem through fast internal interface. In fact, it is an embedded computer with operating system and security mechanisms which control all data flow of terminal passing through network. The design and implementation of the secure terminal which includes hardware architecture, software structure and security mechanisms are given, and the advantages of the architecture are also analyzed.

Keywords: Security, Security Termianal, Architecture, Embeded System, Operating System.

1 Introduction

With widely application of computers and networks, more and more important the security of information terminals connected to network and network itself become. Most of these terminals are using Microsoft's Windows as their operating system (called Windows terminals) now. Unfortunately, the core source code and key technology of Microsoft Windows are proprietary, and vulnerabilities of Windows are exposed evidently. By the way, Windows terminals are the main target of network attacks. In addition, the performance of terminal always dropped rapidly because security processing needs to occupy much CPU time. These problems related to security are some big challenges Windows terminals faced now.

To resolve above problems, many methods such as installing firewall and patching Windows system are taken. Generally, these methods are taken when the system is crashed or the sensitive data are lost after attacking. So the methods are passive ways to defense against network attacks. These passive defense methods cannot meet the

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 715–722.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

security needs of mission–critical and other important data-sensitive information systems. Especially, we have not seen any effective security mechanisms in the market for sealing off special vulnerabilities, such as back doors.

Our research focuses on design and implementation a secure terminal which has special architecture in the point of security and is compatible with Windows at the same time. A secure terminal with security mechanisms in it can detect attacks from network and hidden thoroughfare, and then take adequate defensive actions before damage caused or close possible hidden doors before confidential data divulged. In this way, we can effectively prevent attackers from achieving their intention and greatly reduce possible loss caused by attacking. This kind of secure terminals includes a prime subsystem and an assistant subsystem. These two subsystems connected closely each other. Security mechanisms are mainly implemented in assistant subsystem. More details about the architecture of the terminal and the security defense mechanisms based on it are presented in following sections.

2 System Design

In order to solve security problems of Windows system, our main considerations for designing a secure terminal are as followings: 1) to logically separate Windows system from networks connected; 2) to independently implement the security mechanisms from Windows system; 3) to process and control all the input or output network data of terminal via an embedded operating system irrelevant to Windows, which also runs on independent CPU. The secure terminal we design includes a prime subsystem and an assistant subsystem, which we called dual-subsystem architecture. The prime subsystem handles all the man-machine interaction and all applications run on it. It is totally compatible with Windows. The assistant subsystem is connected to prime subsystem, which processes security mechanisms such as access control, data flow auditing in or out of terminal, etc. Two subsystems run respectively. Each subsystem has its own CPU and operating system. Prime subsystem, off course, runs Windows operating system. Assistant subsystem is really an embedded computer which includes an independent CPU from prime subsystem and an embedded operating system. Since assistant subsystem is in charge of all security processing of terminal in network environment, it decides the terminal security level.

In the implementation of prototype system, assistant subsystem adopts embedded Linux operating system. Windows system and embedded Linux system run independently, they communicate via high speed data transport protocols, and cooperate with each other. For prime subsystem, assistant subsystem is considered as a network card; it is completely transparent for users and applications of prime subsystem. Assistant subsystem is a platform for security processing. Windows system connects to outside networks via assistant subsystem, and all the input and output network data of Windows system are under security control via assistant subsystem firstly.

It just looks like to logically establish a “security wall” between Windows terminal and outside network. Access to Windows system from outside network is under control of assistant subsystem, and any invalid access will be forbidden according to security policies. In addition, all sensitive data of Windows terminal will be audited

under security rules in assistant subsystem, and no confidential data will be disclosed without permission. In other words, the security wall can't be passed by under this architecture. Most of attacks aim at Windows system can be detected and defended by assistant subsystem which runs embedded Linux-like operating system, because these attacks against Windows are generally null and void under Linux environment. This active defense method can effectively prevent data loss, data divulgence or system crashed from attack.

Even some serious loss caused by any intentional vulnerabilities and back doors can be prevented by using this kind of secure terminal. The assistant subsystem can find out the hidden thoroughfare by analyzing abnormal data flow and taking adequate reaction to it. Although vulnerabilities exist in prime Windows system, attackers cannot achieve their ambitions because assistant Linux system will prevent from it actively. In addition, other security mechanisms such as data encryption can be implemented conveniently in assistant subsystem in order to enhance the terminal security. All the security processes run in parallel with other application processes of prime subsystem, which not taking the CPU time of prime subsystem and then improving the performance of terminal.

3 Hardware Architecture

Secure terminal with heterogeneous dual-subsystem means that there are two computer systems named prime subsystem and assistant subsystem respectively in the terminal, each of them runs in different hardware platform and plays independent role. They connect and cooperate with each other. The hardware architecture of this kind of secure terminals is illustrated in Fig. 1.

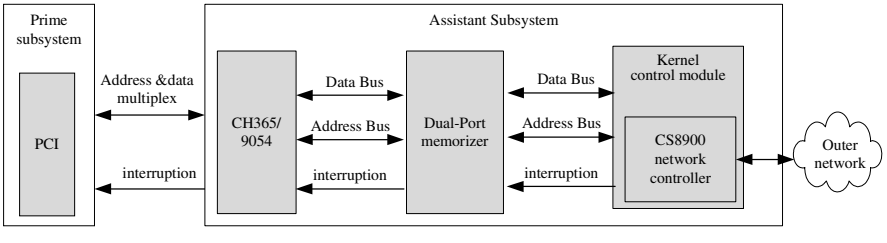


Fig. 1. Hardware architecture of secure Windows terminal with heterogeneous dual-subsystem

The chip in the prime subsystem is Intel CPU with a Windows operating system running on it, and the assistant subsystem has the chip of ARM9, which runs Linux-like embedded operating system. Both of the subsystems connected via PCI bus, and a dual port buffer is used to speed up the communication between subsystems.

Actually, in the prototype terminal, the assistant subsystem is designed as a device of network card-like. The device could be inserted into the PCI slot of prime subsystem in the form of a board card, and the card is an ordinary I/O interface card for Windows users. As an I/O device, prime subsystem can identify and manage it. The two subsystems connect with each other through a high speed bus. Data

transmission can be done via the bus. A special chip of dual-port buffer storage is used to bridge transmission difference.

In our prototype system, the core card adopts the YL2410 of AMR9, the PCI special interface chip adopts CH365, the dual-port memorizer adopts IDT70V05S, and the network controller adopts CS8900A.

4 Software Structure

Based on terminal architecture of heterogeneous dual-subsystem, software module was collocated in Windows prime subsystem and security assistant subsystem. The main functions of software module are as followings: 1) to ensure the normal operating of Windows users and applications. The assistant subsystem is transparent for them; 2) to ensure all the input and output network data of Windows system are in secure controlling; 3) to provide the interface of interaction and deployment for assistant subsystem security controlling. The software framework is illustrated in Fig. 2.

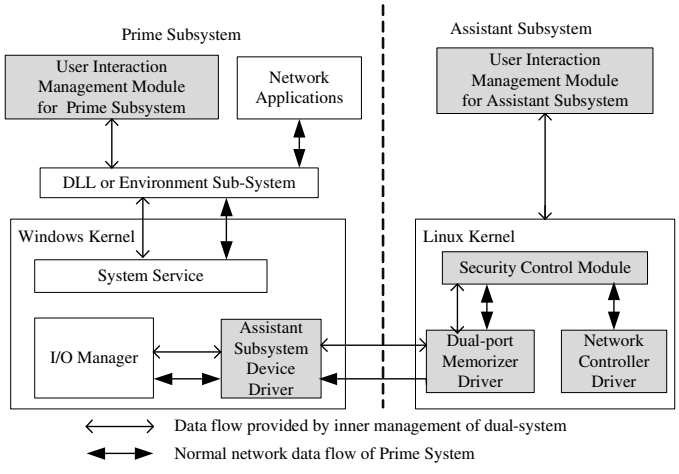


Fig. 2. Software framework of secure Windows terminal system with heterogeneous dual-subsystem

There are three kinds of software modules in a secure terminal: device manager or driver, user interaction management and security control module.

In accordance with the hardware architecture of secure terminal, the chips or controllers in the assistant subsystem, such as dual-port memorizer, internal communication channels and serial I/O interface etc. should be managed coordinately by two subsystems. Three main managers are described as follows: 1) Assistant subsystem device driver, which is located in Windows system and is responsible for transmission all data flow from or to network accessed by Windows prime subsystem. From the view of Windows users and application, assistant subsystem just looks like a normal I/O interface card and is transparent. Assistant subsystem cannot interfere with the operating of prime Windows system; 2) Dual-port memorizer driver, which

is located in assistant Linux system and is responsible for the operations, such as data read and write in dual-memorizer. As a data buffer, the dual-port buffer memorizer is a bridge of the high speed communication between prime subsystem and assistant subsystem; 3) Network controller driver, which is located in assistant Linux system and is responsible for driving the network controller. The network controller in the assistant subsystem is the only interface for a secure terminal to communicate with outer networks connected.

User interaction management module is responsible for interaction of two subsystems. The whole module is divided into two parts: prime subsystem side and assistant subsystem side, and located in prime and assistant subsystem separately. There are two kinds of interaction, one is user enrollment and the other is security event alarming. From the point of security, an internal channel is specially designed for transfer the interaction information of two subsystems. The network user must be enrolled at the first time from graphical user interface provided by prime subsystem side program of user interaction management module, and then the enrollment information is transferred into assistant subsystem via secure internal channel, and at last the validity of network user is checked according with the enrollment information by the assistant subsystem side program of user interaction management module. By contrast, security event alarming information will be reported to prime subsystem from assistant subsystem via secure internal channel.

Security control module runs in assistant subsystem. The core security functions are accomplished in the module. It monitors all network communication processes, analyzes and filters data package by using technologies such as user identification, protocol authentication, keywords matching, and access control etc.

5 Main Security Mechanisms

Since the assistant subsystem is located between the prime subsystem and outer networks, it becomes the sole transfer channel from assistant subsystem to outer networks. Any network data package from or to prime subsystem can be intercepted and captured by assistant subsystem. It is the critical foundation that active and overall defense policies and mechanisms can be implemented in above secure terminal.

Any network data package is checked and filtered by assistant subsystem according to security mechanisms. Secure policies, mechanisms and functions can be added, removed or adjusted by administrators according to the application needs and environments. When data packages are transferred, they are intercepted and captured by assistant subsystem, and their relevant information is analyzed. As a result, packages according with the security polices can be got across, otherwise, they will be thrown away. Security functions provided by assistant subsystem are as follows:

1) Access and control to terminal sensitive data. It can prevent from letting out the confidential data unconsciously or intruding in by hackers or other malicious users from.

2) Identify authenticate and privilege control of user. A security level is designed for every network user in assistant subsystem. The privilege of network user is classified according to security level. Terminal network user can send information matching its privilege or security level, operation beyond authority will be forbidden.

3) Filter and other control for all network data. It is taken by assistant subsystem to all network data crossed according to protocol authentication, keywords match and access control mechanism.

4) Log. All of operations against security rules in assistant subsystem are logged in detail. Log information will be saved in a USB flash disk or the background severs connected through internal connection.

Security processing framework of secure Windows terminal with heterogeneous dual-subsystem is illustrated in Fig. 3.

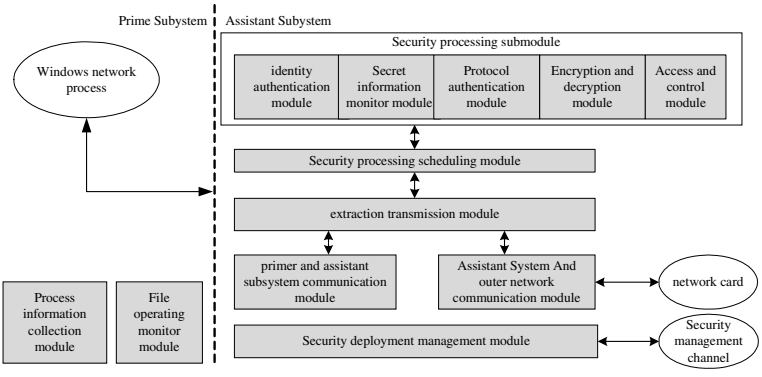


Fig. 3. Security processing framework of secure Windows terminal with heterogeneous dual-subsystem

Following modules are mainly related to security processing:

1) Modules related to process information collection. There are two modules: process information collection module and file operation monitoring module, both in prime subsystem. The former is in charge of the information collection of all network processes, and the later is charged with the information collection of file processing. There are three types of security level: the first is keyword's security level, which is maintained in a table and can be set or modify by administrator; the seconds is file's security level, which is added to file's property and decided by the highest security level of all the keywords in it; the third is process' security level, which is the highest security level of files accessed by processes. The security level of process is changed dynamically in time. Most of the confidential data are kept in files. Process' security level is related to security level of files accessed, and can be determent according to the information collected by above two modules. Accordingly, access control policies can be implemented in line with security level of keywords, files and processes.

2) Security deployment management module. It is in charge of the maintenance of system security rules and deployment information. Security management information managed by the module includes security rules, policies, user password, and security level etc, which is the foundation of all security mechanisms. Therefore, firstly, the security management information is managed only by assistant subsystem; secondly, a security administrator is created in order to carry out the security management;

thirdly, special security management channel is set up. Security administrator account is totally different from Windows' user account. It is a special security management channel used by the administrator to manage user's privilege, keywords, and other security information. So, it is impossible that hackers or other attackers damage any security management information illegally. Via the security management channel, some background server or PDA etc connected can be used by administrator to deploy and manage the system security management information. In addition, the log data can also be kept in a background server via the channel.

3) Modules related to communication modules. There are three modules. Prime and assistant subsystem communication module is in charge of the communication between prime and assistant subsystem. Assistant subsystem and network communication module is in charge of the communication between the assistant subsystem and network. They are both implemented on link layer of Linux network protocols. Extraction transmission module is in charge of extraction of communication data packages. At first, the original data package is extracted at the network layer and given to the network security processing module which will take reaction in responsible for security events. And then, the allowed data package is accepted and submitted to original destination. The assistant subsystem is the only way for communication between Windows mainframe and outer network. All of the network data requested by windows network processes must be checked in these modules before they are input and output successfully through network card.

4) Modules related to security process. These modules include a security process schedule module and a set of security process sub-modules. Security process schedule module submits all the security events to the sub-modules in line with to the security rules for processing. At present, some key sub-modules are implemented in the prototype system such as identity authentication, confidential data monitoring, protocol authentication, access control, encryption and decryption etc.

6 Conclusion

Windows is the most widely used operating system in the market, but security problems of Windows cannot be ignored, especially in mission-critical fields. To solve terminal security problems, a secure terminal based on heterogeneous dual-subsystem architecture is described in this paper. By using this kind of terminal, all the main security mechanisms are processed in assistant subsystem of terminal, data in or out of terminal will be monitored, and possible attacks will be detected and stopped by assistant subsystem according to the security policies. Intentional vulnerabilities or back doors will be detected also by assistant subsystem of terminal. We have designed a model terminal which works as we expected. Now we enhance the system and improve the efficiency so as to industrialize the system in the future.

Acknowledgements. The authors would like to acknowledge Han Xingye for the useful discussions on this topic and her inputs and feedback while writing this paper.

References

[1] Wu, A.-H.: Windows 2000/XP WDM Device driver development. Publishing House of Electronic Industry, Beijing (2006)
[2] Wanf, P.-L., Zhang, G.-X., Wang, N.-L.: Research of Network Driver of dual-processors secure Windows terminal. Computer Science and Explore (2008)
[3] Yu, G.: Research and Implement of I/O and system management technology of dual-processors secure Windows terminal. Master, Institution of Command and Technology of Equipment (2008)
[4] Guo, X.-Y.: Research and Implement of security technology of dual-processors secure Windows terminal. Master, Institution of Command and Technology of Equipment (2008)
[5] Zhao, Y.: Research of User Access Control technology of U-Disk of secure terminal. Master, Institution of Command and Technology of Equipment (June 2009)
[6] Sun, J.-K.: Development technology Describe embedded Linux–Base on ARM. The People's Post and Telecommunication Press, Beijing (2006)

OsgAL-Based Sound Source Management in Simulation Games

Xin Li, Dehui Kong, and Yong Zhang

Beijing Key Laboratory of Multimedia and Intelligent Software Technology,
College of Computer Science and Technology, Beijing University of Technology,
Beijing 100124, China
lixin_903@126.com, {kdh,zhangyong}2010@bjut.edu.cn

Abstract. Sound rendering is a technique which uses auditory display to communicate information to a user and provides an alternative way of visualization. Traditional 3D sound rendering techniques are lower effective when the number and types of sound sources increase. In this paper, we propose a method of 3D sound source management based on osgAL toolkit, by using sound classification and XML document structure for sound properties. We allocate buffers for necessary sound sources and cull inaudible sound sources, which reduce system's consumption of time and space. Experiments on train-driving simulation system show that our method provides a realistic and effective simulation.

Keywords: Sound source management, 3D sound rendering, osgAL.

1 Introduction

1.1 3D Interactive Sound Simulation

3D sound rendering is a technique of generating a synchronized soundtrack for animations. Today, it becomes more and more important in many areas of virtual reality and games. Although techniques of realistic visual rendering of complex dataset are quickly developed with high growth rates of current graphics hardware and advances in rendering acceleration algorithms, simulation of sound or auditory rendering is still a difficult problem to solve. Traditional sound rendering techniques in 3D games have the following disadvantages:

- The real-time rendering is lower effective when the number and types of sound sources increase.
- The simulation has less sense of reality, because it is a simple sound-playing procedure.
- There is no sound localization and special sound effects in simulation.

1.2 Previous Work

Many scholars did research on real-time sound rendering in virtual. Mano et al. give an overview of new algorithms for numerical and geometric propagation [1]; Luo et

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 723–732.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

al. denote to sound rendering technique in virtual reality system [2]; Lokki et al. create realistic 3D audible virtual environment and simulate sound under this environment [3]; Li et al. consider spatial sound model in view of acoustic measurement, and deduce a homophonic stereo synthetic model [4]; Zhang et al. describe essential techniques in sound simulation: spatial sound localization principle, distance model of sound attenuation, Doppler effect formula [5].

Although current techniques produce natural sounding audio and are plausible, these systems use simplified modeling methods. While the number or types of sound sources increase, they become lower effective. In order to produce an accurate rendering of sound, it is necessary to manage large number of sound sources effectively and simulate 3D sound precisely in virtual scenes.

1.3 OsgAL-Based Sound Source Management

Recently, a new open source toolkit "osgAL" is developed based on OpenSceneGraph to deal with 3D stereo [6]. It depends on OpenAL++ and OpenSceneGraph, and contains two libraries: OpenAL++ and osgAL. The former one handles operations regarding sound stream, sound source and listeners, while the latter one implements special effects of 3D sounds by integrating sounds into scene graph.

Nowadays, sound simulation in virtual environment has higher requirement on its accuracy. A solution is needed to organize and manage many kinds of sound sources in virtual scene. In this paper, we propose a multiple sound source management in simulation games, by using sound classification and XML document structure for sound properties. Experiments on train-driving simulation system show that our method is considerably effective. The paper is organized as follows. Section 2 gives a pre-processing on sound sources and data, including sound object classification, sound loading and binding the relations; in Section 3, we give a management solution for sounds by allocating and culling sound sources; the osgAL-based sound source rendering method is given in Section 4; experimental results on train-driving simulation system are given in Section 5; we conclude this paper in Section 6.

2 Sound Modering

In this section, we give a pre-processing on sound sources and sound data.

2.1 Sound Object Classification

To achieve a realistic rendering effect of sounds in simulation games, one has to analyze probable sounds and factors involved appearing in scenes. Sounds in simulation games can be generally divided into two classes: 2D sounds and 3D sounds. The main difference between them is: the volume and tone of 3D sounds directly depend on relative locations of sound source and listeners. Hence, one has to get their spatial properties such as location, velocity, orientation, etc.

- Location: with different relative locations of sound source and listener, different sound effect arises.

- Velocity: the moving velocity of sound object relative to listener, is an essential parameter for enhancing reality sense, which is used to compute Doppler effect.
- Diversing cone: an intrinsic property of sound source, which gives diversing angle of an oriented sound. The diversing cone can be divided into inner cone and outer cone, which is used to judge whether sound source is within the region where listeners are able to hear (see Fig.1 for its principle).
- Orientation: the vector which starts from the origin of virtual environment coordinates, and points to diversing cone's axis.
- Maximum distance: when the distance between sound object and listener is greater than maximum distance, sound can't be heard.
- Minimum distance: when the distance between sound object and listener is smaller than minimum distance, sound won't enhance if the distance becomes shorter.

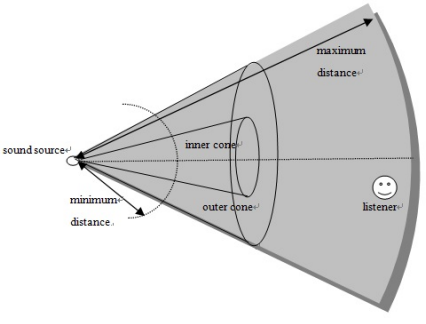


Fig. 1. The diversing cone of sound

Moreover, considering particular requirement of simulation games, sound can be divided into the following types according as its functions.

- Music: 2D sound, such as background music provided in virtual scenes, and sound effects produced by switching interfaces or buttons.
- Sound effect: the natural effect in virtual environment or effect produced by player's operations, such as the sound caused by object impact.
- Voice: sound of dialogue, broadcasting or hint, such as introduction of operations, feedbacks or hints to player's operations.

2.2 Relations between Sound Object and Entity

Because sound is formed by vibration, it is necessary to bind sound data to entity model in virtual environment, in order to enhance stereoscopic sense. For real-time rendering of a 3D scene, scene graph is commonly used for scene management, which organizes all the data in the form of tree structure. Each node of a tree can have arbitrary number of child nodes. A graphic application program can be considered as the root of a tree, each scene contained in program can be considered as a node, and

each node also contains other items. These items can be further classified into geometric object, texture, transform, light source, bounding box, subdivision level, and so on.

Here we consider moving sound source with properties (such as position or orientation) as a leaf node (the brother node of entity object) in scene graph, and bind it to corresponding entity model. The relation is shown in Fig.2, where sound source (SoundNode) and entity model (ModelNode) are simultaneously considered as child nodes of matrix-transform node, so that it contains position and orientation information of entity node and hence can implement real-time rendering of 3D sound while the scene is doing the updating traverse.

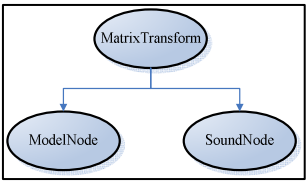


Fig. 2. The binding relations between sound node and entity node

2.3 Sound Data Loading

The sound object used in this paper is an encapsulation of a sound data, which has already been decoded, in standard PCM format, and can be played in hardware directly. The object is loaded into memory in the form of file, and mapped onto virtual environment to produce 3D interactive sound.

As there are many number and kinds of sound sources in virtual scene, loading all the sound files into memory will not only decrease the efficiency of games, but also waste system resources heavily. Therefore, based on XML technique, we use document structure to express sound source information, and load sound files into buffer according as different game scenes and different types of sounds (background music, sound effect, voice), to realize sound rendering. By matching names of scenes, our method avoids loading unrelated sound files. Meanwhile, it is convenient to release sound source of last scene during switching scenes, which saves system's consumption. The XML document structure of sound source information in a given scene is shown in Fig.3, where all the three resource nodes have five parameters: type of sound (category), name of sound (name), storage path of sound file (url), priority of file (priority) and being occupied (active). We shall introduce the use of other parameters of sound sources in the next section.

```
<scene id=“scene of game”>
<AudioFile>
<category>Music</category>
<name>bgm</name>
<url>..// Sound//bgm.wav</url>
<priority>9</priority>
<active>true</active >
```

```
<loop>true</loop>
</AudioFile>
<AudioFile>
<category>SoundEffect</category>
<name>trainRunning</name>
<url>..// Sound//train.wav</url>
<priority>10</priority>
<active >false</active >
<minDistan>5</minDistan>
<maxDistan>50</maxDistan>
<soundCone>30</soundCone>
</AudioFile>
<AudioFile>
<category>Voice</category>
<name>stationReport</name>
<url>..// Sound//stationReport.wav</url>
<priority>10</priority>
<active >true</active >
<timeLen>5</timeLen>
</AudioFile>
```

The XML document structure of sound source information.

3 Sound Source Management

In this section, we propose the sound source loading program. It handles analyzing XML document and initializing sound object. When sound object is loaded into scene, if the number of sound sources is large and the system resource is limited, then reasonable sound source management and culling method are essential to do real-time sound rendering effectively.

3.1 Sound Source Allocating

Sound object not only uses memory during data storage, but also uses buffer of sound system. The buffer can be divided into hardware buffer and software buffer. Usually hardware buffer is supported by sound card with fixed number; while software buffer is supported by driver program whose number depends on memory and performance. Computation of these two buffers is mainly supported by sound card and CPU respectively. To decrease CPU's burden, one has to use hardware buffer in priority when creating sound source, and try to use software buffer as less as possible, which needs to release unnecessary sound source in time. To guarantee the stability of system transplant, for the initialization of sound system we set maximum number of available sound sources: *N*. When the number of sound source playing simultaneously is greater than *N*, a program error will occur, which forces to delete unnecessary sound source in time. So we propose an algorithm to allocate and make use of sound sources reasonably to implement simultaneous playing of multiple sound sources.

At first we set properties: “priority” and “active” to each sound file, which represent the priority of the file and whether source allocated for current file could be occupied by other source with higher priority. Then we define arrays: availableSources and activeSources, which store current available sources and occupied sources respectively. The sources which are not occupied need releasing manually after the playing stops. The sound sources generated during the initialization of sound system are sent into availableSources. When source is needed to be allocated, a source is returned from availableSources; if it is empty, then traverse activeSources and find one which has lower priority than current one and which has lowest priority. If it still doesn't work, then an exception is thrown. Fig.4 shows the flowchart of the sound source allocation algorithm.

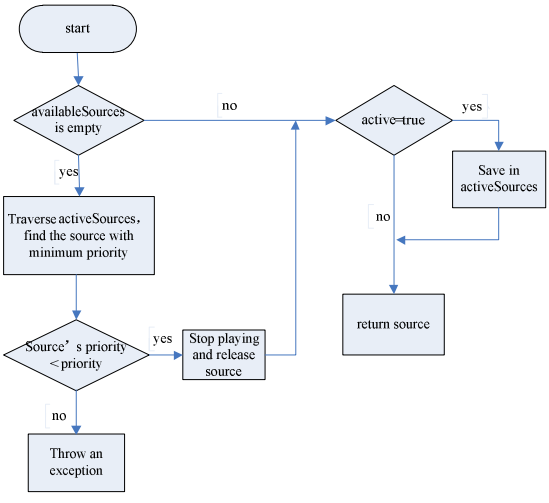


Fig. 4. The implementation flowchart of sound source allocation

3.2 Sound Source Culling

The changing relative location of listeners and sound sources in virtual scene (and other factors) have obvious influence on effects of sound source, such as Doppler effect arising from relative velocity between listener and sound source, echo in closed spaces, etc. Hence listener can’t always hear all the sounds in virtual environment, and that depends on relative location of listener and each sound source. To render 3D interactive sound more accurately and effectively, one has to do real-time culling to sound source in virtual scene. According as its direction and location, 3D sound in virtual scene can be classified into three types:

- Non-oriented sound: such as sound arising from driving trains or crowds in train station.
- Oriented sound: such as sound arising from train in a tunnel.
- Non-positional, non-oriented sound: degenerating into 2D sound, such as sound of rains.

For sound with no orientations, we consider its maximum distance as the threshold of audibility. That is, when the distance between listener and sound source is smaller than this value, then the sound is put in playing list and set associated properties for rendering; otherwise, it is deleted or hung on. For oriented sound, location or direction information (relative to listener) can be used to testify whether listener is both within diversing cone and within maximum audible distance for each sound node. See Fig.1, where grey region represents audible region of listener with respect to the sound source. If the relative location of listener and sound source changes and listener is without audible region, then the sound object can be deleted or hung on. The culling procedure is shown in Fig.5.

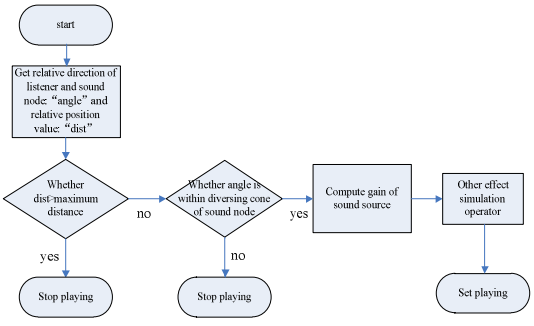


Fig. 5. The flowchart of sound culling

4 Osgal-Based Sound Source Rendering

OsgAL is formed by four basic classes. First of all, SoundState class combines sound sample object in buffer with its properties; then encapsulate the mobile sound source with location and orientation as SoundNode and add into the scene graph, which realizes the binding of the sound source and entity models. During the updating traverse, one can compute its world coordinates to update its velocity, direction and other properties; SoundManager is responsible for initializing sound system, storing and handling status of loaded sound source; SoundRoot updates the SoundManager and transformation of position and velocity for the current listener during cull traversal. The principle is shown in Fig.6.

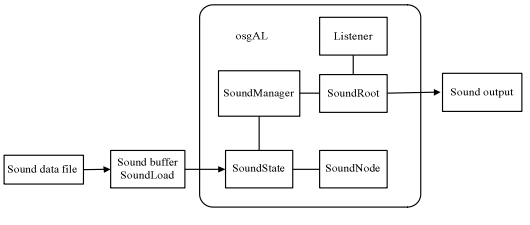


Fig. 6. The principle of osgAL

4.1 Background Music Playing

2D background music is a kind of stationary sound source, which is not needed to be encapsulated as a sound node. Its information is read from XML files of corresponding scene. Then we allocate sound source for it, set volume, playing mode, stopping manner and other properties, register them in sound manager and control its playing by replying keyboard-events.

4.2 3D Sound Rendering

There are many sound ingredients in a game system. Take the simulating train-driving game for instance. The major sound sources are trumpet and roar of the engine during driving. From the viewpoint outside train, there is an obvious occlusion effect when train passes through a tunnel; and when train starts, if the viewpoint does not change with moving train, one can feel a strong attenuation of sound. All these performance can be implemented by setting associated parameters and functions supported by osgAL.

At first we instantiate SoundManager, obtain initial sound environment, set number of sound source, distance model and Doppler factor. Then we create a sound sample object, and load train-driving sound files from hard disk; we allocate sources and set its priority, gain, pitch, relative distance, attenuation factor, size of sound cone, playing mode and other properties. We manage and operate it by putting it into state array of sound management. To show the occlusion effect, we use OccludeCallback to judge whether the sound light starting from sound source is prevented by buildings, and do corresponding volume-decaying operation. Finally we encapsulate it into sound node and combine it with entity node to update its location and velocity.

After finishing setup and encapsulation of sound source, the sound node needs to be added into scene graph immediately. That is, based on structure of OSG, we combine sound source node with the associated train node together, and add it into scene graph as a child node of matrix transform node. The location relations of the whole scene are shown in Fig.7.

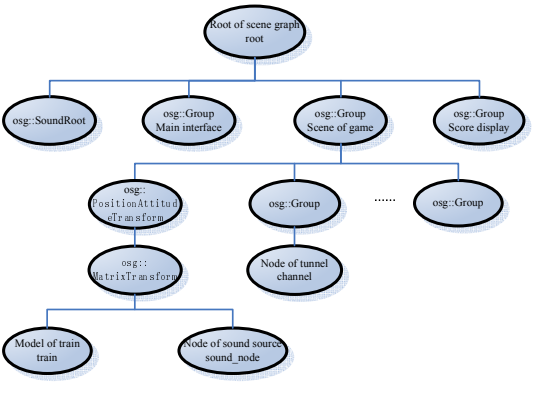


Fig. 7. The scene graph for sound system in train-driving simulation games

In Fig.7, after the sound node is combined with entity node and added into scene graph, one can compute its world coordinates during updating traverse, and update its location, direction, velocity and other properties when status of associated entity changes, which realize the sound simulation.

4.3 Voice Playing

For the playing of voice in system, such as playing of broadcast and advertising material, it is impossible and unnecessary to generate a sound node for each sound, since they have great amount and need to be played according to fixed order. Thus, all broadcasts could share a sound node, and system sets associated sound data file to this node according to different requirements at any time, in order to produce the different broadcastings. That would need additional time factor to control. For example, anytime train is moving, it needs to play the information which station is to arrive and what speed it reaches. As information about the station names and real-time speed are dynamically updated, it is important to control time and arrange the playing order of voice files. In this paper, we implement it by using two approaches according as the lengths of playing time for different sound files. For the sound files with short playing time, we set single thread to check whether current sound is over, and then set the following content to play; for those with long playing time, we save that time for a property value in the node of its XML file, and determine when to play the following content by this value.

5 Test Experiment

We test the sound source management system on train driving simulation games, and do the whole processing including loading sound, managing and clustering sound source, rendering and releasing sound. During the train driving, the sound heard by player can be classified into environment inside driver's cabin and environment outside driver's cabin. Take the case that viewpoint is outside cabin when train starts as an example. After sound clustering, the sounds which exist include: sound from contact with track, sound of winds arising from train movement, sound of broadcasting, sound of hint and background music. The virtual environment is shown in Fig.8.

When player starts train with operation, he can hear the sound of train moving, and he can feel sound attenuation when train drives out of station if his viewpoint remains the same; and when train passes through barrier such as tunnel, player can feel the occlusion effect of sound. When his viewpoint changes, the locations of all the sound sources relative to listener will do real-time update. And rendering effect changes as well.

Experiments show that, by using the sound model and sound source management, sounds can be added into 3D virtual scene effectively, and 3D real-time sound rendering can be implemented.

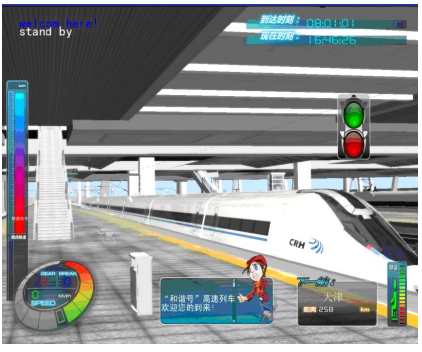


Fig. 8. The Scene that train starts from station

6 Conclusions

This paper proposes a 3D sound rendering method, by classifying sounds and building a sound source management model. We implement loading and releasing sound sources of train-driving simulation system with XML technique, and use scene graph structure and apply osgAL toolkit to realize sound source culling and 3D sound rendering. The result is realistic and satisfactory.

Acknowledgment. This paper is supported by the National Natural Science Foundation of China (Nos. U0935004, 60825203), National Key Technology R&D Program (No. 2007BAH13B01) and Beijing Natural Science Foundation (No. 4102009).

References

[1] Manocha, D., Lin, M.C.: Interactive sound rendering. In: 11th IEEE International Conference on Computer-Aided Design and Computer Graphics, Huangshan, China, pp. 19–26 (August 2009)

[2] Luo, F.Y., Wang, X.R., Peng, X.Y.: Sound rendering technology and its application in virtual environment. Journal of System Simulation 11(5), 364–367 (1999)

[3] Lokki, T., Savioja, L., Vaananen, R., Huopaniemi, J., Takala, T.: Creating interactive virtual auditory environments. IEEE Computer Graphics and Applications 22(4), 49–57 (2002)

[4] Li, Y., Driessen, P.F., Tzanetakis, G., Bellamy, S.: Spatial sound rendering using measured room impulse responses. In: 2006 IEEE International Symposium on Signal Processing and Information Technology, Vancouver, Canada, pp. 432–437 (August 2006)

[5] Zhang, D., Chen, D.F., Jiang, Y.: 3D sound modeling technology and its application in vehicle driving simulator. Journal of Hubei University of Technology 22(3), 22–24 (2007)

[6] `http://www.vrlab.umu.se/research/osgAL`

[7] Li, X., Zhang, J.Q.: 3D sound organization and management in virtual environment DVENET. Journal of System Simulation 13, 169–173 (2001)

An Automatic Extraction Method of Surveillance Visual Context

Haozhe Liang, Shukui Xu, and Guohui Li

National University of Defense Techonology, Department of System Engineering
Deya Road, ChangSha, 410073, China
lianghaozhe@hotmail.com, xskgfkd@163.com, guohli@nudt.edu.cn

Abstract. In this paper, we propose an algorithm to address the problem of automatic extraction of visual context in surveillance video scene. Firstly, we analyze the trajectory information distribution to get the characteristic of the information; then by using Gaussian mixture model we do the cluster analysis of the moving tendency reflected by the trajectory direction. For the similar direction trajectory segments, we do the hierarchical clustering based on the proposed similarity measure to get the key routine of the scene. At the end of this paper, the experiments result shows the efficiency of our algorithm.

Keywords: visual context, trajectory, surveillance, activity understanding, clustering.

1 Introduction

Recent years, visual sensors have been deployed everywhere. Visual surveillance techniques monitor the important scene by visual sensors. But those techniques which monitor only by human eyes cannot be able to construct an effective security visual surveillance system. Intelligence visual surveillance techniques are the hotspot in the video processing domain [1]. They use the sensors to observe the moving object in real-time, including detection, recognition, and tracking to locate and recognize motion in the scene and then understanding individual motion and interaction taken by multi-agent in the visual scene.

With development of the visual motion detection and tracking, how to understanding the content of the scene so called activity understanding is an attractive research. Activity understanding attempts to supply a solution which makes the computers have a consistent perception with human being. The main challenge depends on the transformation from low-level feature to high-level description in the image sequence in a positive way [2][3].

Within the research of activity understanding, automatic solution of transformation is a key point. We define spatial-temporal constraints of the scene with notion 'visual context'. Visual context can promote tracking of moving object and activity understanding techniques because they supply the knowledge of visual scene. Automatic visual context extraction is a basic step to solve the problems of activity understanding.

In a far-field visual scene, trajectory of the moving object is a discrete vector convolving in the state-space. The vector can have attribute values of location, orientation, velocity, and object size. The automation extraction of context based on the trajectory is the prevalent method.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 733–741.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

In this paper we address the problem of automation extraction of visual context in visual surveillance application. Firstly, we analyze trajectory information distribution to get the characteristic of the information. Then by using Gaussian mixture model we do cluster analysis on moving tendency reflected by the trajectory direction. For similar direction trajectory segments, we do hierarchical clustering based on the proposed similarity measure to get key routines of the scene. The rest of this paper is organized as follows. In section 2 we describe the related work. In section 3 specify the trajectory based visual context modeling. In section 4 we present experiments to test and verify our algorithm. In section 5 we conclude the paper and propose our future work.

2 Related Works

Against to some visual scene, which cannot be tracked in an effective way, there are lots of solutions to solve the extraction problem of visual context. [3] proposed a motion template called pixel change history (PHI) which has similar but better attributes than motion history image [4] to describe the moving tendency of individual agents. Based on PHIs clustering and event Hidden Markov Model, the machine can find significant semantic events and topology of events spatial-temporal structure. [15] used a blob detector to find foreground feature vectors with silhouette, size, and direction attribute, which can be clustered by a Gaussian mixture model and been used to find semantic regions in the visual scene. [16] used the optical flow as a low-level feature, and then coding the structure of visual scenes with direction and spatial location which called visual words. Visual topic that molded as the distribution on visual words has obvious semantic meanings about motion concurrency of the scene.

In general far-field camera environment, utilizing tracking information to extract semantic region is a prevalent way. [6] proposed an adaptive Gaussian mixture model against the problem of background subtraction, and then find motion patterns based on tracking information classification. [7] used two similarity measures to clustering trajectories to find key semantic regions like paths, sources and sinks of trajectories.[8] proposed an accumulative trajectory clustering method to learn the path region in visual scene. [9] used the nonparametric Bayesian model to describe distribution of trajectory documents, inference the number of semantic region and spatial-temporal constraints in an unsupervised fashion. [10] utilized the probability theory to mining sources and sinks of trajectories. In the domain of AI and data mining, research of getting patterns of time-serials is a hotspot. [11] proposed to code the trajectory with redundant location which can describe spatial-temporal constrains in the serial. In order to avoid the problem of spatial segmentation size, [12] proposed a disjointed zone coding patterns, and then finding frequent ST patterns.

Our method aims to employ appropriate clustering techniques to solve mining of trajectory information, attributes of different distribution characteristics. We use the Gaussian mixture to model the multi-peak of directions and using key routines to describe the distribution of spatial constraints of trajectories.

2. 1 Trajectory Based Surveillance Context Modeling

2.1.1 Visual Context model Based on the Trajectory Spatial-Temporal Similarity

In this paper, we define the notion of trajectory as: a discrete serial of vectors with multidimensional attributes, which taken by the moving object in the visual surveillance

scene and evolved in the state space: $T=\{P_1,P_2,\ldots,P_n\}$, n is the number of frames which is the duration of moving object in the video. P_i is instantaneous point at time I which we mean like this: $P_i=\{(x_i,y_i,A_i)|1\leq i\leq n\}$, (x_i,y_i) is location attribute, and A_i is attached attribute which is the direction in this paper.

Based on the moving object detection and tracking, we can get the sets of trajectories: $TS=\{T_1, T_2\ldots T_N\}$. In the surveillance scene, the trajectories can reflect motion information of the visual context, for example, the moving tendency and the main moving spatial region. In this paper, we focus on the automation extraction of visual context based on trajectory. At the beginning of the modeling, we assume the moving of similar activities can construct some significant semantic regions, so the trajectories in the certain region may have temporal-spatial similarity:

1) We use clustering of the trajectories direction to modeling the moving tendency.
2) Based on the 1), we cluster the trajectories which are closed with each other spatially.

In order to do clustering analysis effectively, we first observe the data distributions of the trajectory direction and location. In Figure [1] up, red color means the right direction and blue means inversely.

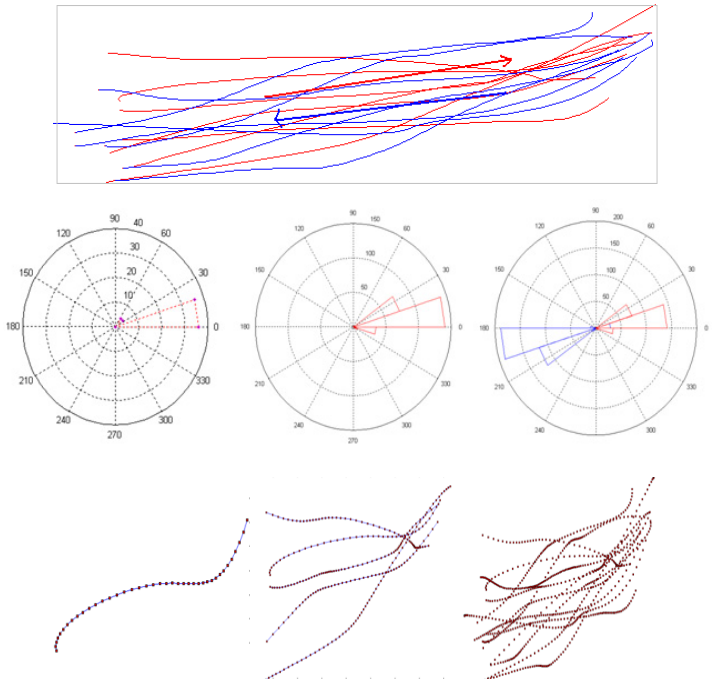


Fig. 1. Data distribution of trajectory

According to the Figure [1] middle and bottom, the distribution of trajectory direction has multi-peaks characteristic because of the rule of surveillance scene. So we can model it by Gaussian mixture model. The spatial location of the trajectory does not have significant regulation. Recently key zones like paths, routines, sources, sinks mining is the mainly method in trajectory spatial analysis.

Based on the analysis above, we propose the algorithm as the cancelled term:

Input: Trajectory Sets: TS = $\{T_1,T_2,\ldots T_N\}$,T_i=$\{P_{i1},P_{i2},\ldots P_{in}\}$,$P_{ij}$= $\{x_{ij},y_{ij},A_{ij}\}$,(x_{ij},y_{ij}) is spatial location, A_{ij} is direction attribute.
Output: Visual Context：R=$\{R_1,R_2,\ldots R_k\}$
Step 1:Using the infinite Gaussian mixture model to cluster the direction of trajectory in an unsupervised way, results in the visual context with similar moving tendency.
Step 2:Using the hierarchical clustering technique to do the trajectory spatial analysis。For every TS_i = { T_{i1}',T_{i2}',…$T_{iN'}$}, members of which have similar moving direction, analyzing the spatial similarity of the trajectory.

2.2 Unsupervised Clustering with Infinite Gaussian Mixture Model

GMM is an effective method to do multi-peaks data analysis, which has probability density function like [6]:

$$p(x)=\sum_{j=1}^{k}p(x|z=j)\pi(z=j)=\sum_{j=1}^{k}\eta(\mu_j,\gamma_j)\pi(z=j) \tag{1}$$

$\eta(\mu_j,\gamma_j)$ is a single Gaussian. $\pi(z=j)$ is the mixture portion. When the structure is known, we can estimate the parameter with the EM. How to determine the 'true' number of mixture is a gloomy problem. Rasmussen [13] proposed Infinite GMM to avoid the 'true' problem which can result in under-fit or over-fit. The IGMM just give the mixture portion a Dirichlet Process prior, which can draw infinite discrete indicator variable c to inference the component which data belong to. The posterior probability of c is:

$$\begin{cases} n_{-i,j}>0, p(c_i=j|c_{-i},\alpha)=\dfrac{n_{-i,j}}{n-1+\alpha} \\ p(c_i\neq c_i i'\neq i|c_{-i},\alpha)=\dfrac{\alpha}{n-1+\alpha} \end{cases} \tag{2}$$

$n_{-i,j}$ is the number of data belong to the component j except data point i. Using the Gibbs Sampling method, we can get the parameters of 'active' components which in fact means the 'true' components.

In figure [2], we show the result of IGMM clustering compared to BIC and AIC which have been applied to the decision of 'true' number of finite GMM with simulation datasets. In figure [2] up, we see that AIC will have over-fit problem and BIC with more or less under-fit while IGMM Gibbs sampling avoid those instances as we show the curve describing the number of GMM components based on AIC, BIC

and the histogram of the distribution of components number based on IGMM sampling. In figure [2] bottom, from left to right, we show the clustering result based on AIC, BIC with four components, BIC with five components, and IGMM.

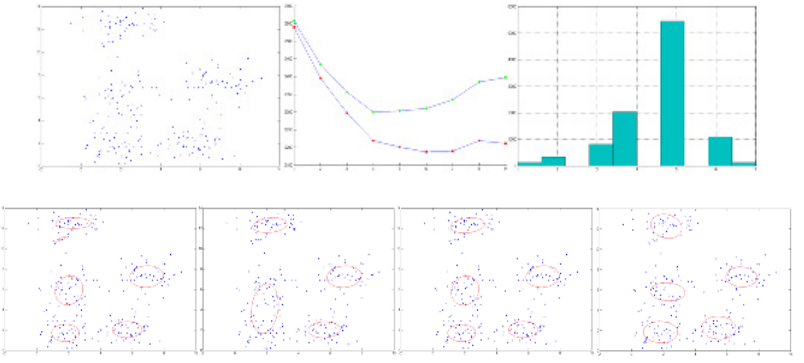


Fig. 2. IGMM BIC AIC clustering and parameter estimation

2.3 Spatial Similarity Hierarchical Clustering

As mentioned above, we find the shape of spatial distribution of trajectory does not follow certain regulation. In this paper, we focus on the learning of the key routine in surveillance scene, which is prevalent in recently years [8][9]. In this paper, similar to [8], the notion of 'routine' means the region which been crossed frequently by the moving object.

Firstly, we use Douglas-Peucker line simplification to reduce the dimensions of trajectory. This process intelligently selects and deletes visually redundant points but preserve sufficient visual details. This result in the trajectory transformation [12]:

$$T:\{p_1,p_2,\ldots p_n\}\rightarrow TDP:\{p_1,p_2,\ldots p_s\},s\leq n_\circ$$

Based on this discription, we use the mostly closed visual point proposed in [7] between trajectories like:

$$\psi(i)=\arg\min_{j\in T_j}\|(x_i^{\,a}-x_j^{\,b},y_i^{\,a}-y_j^{\,b})\| \tag{3}$$

We use the direct distance measure like:

$$direct(A,B)=\max\|(x_i^{\,a}-x_{\psi(j)}^{\,b},y_i^{\,a}-y_{\psi(j)}^{\,b})\|\ \ direct(A,B)\neq direct(B,A) \tag{4}$$

Then we have the symmetrical distance measure:

$$d(A,B)=\max(direct(A,B),direct(B,A)) \tag{5}$$

Just like figure [3], the reduced trajectory points describesobviously curve evolving corner points and the distance measure describes the visual separation between trajectories.

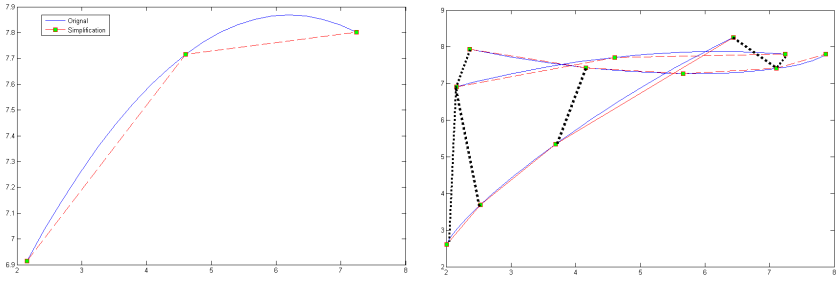


Fig. 3. Douglas-Peucker line simplification and trajectory spatial similarity

2.4 Algorithm Procedures

As mentioned above, we conclude our algorithm about Visual Context model based on the Trajectory Spatial-Temporal Similarity as table [1]:

Table 1. Algorithm procedures

Input: Trajectory Sets: TS = $\{T_1,T_2,\ldots T_N\}$,T_i=$\{P_{i1},P_{i2},\ldots P_{in}\}$,$P_{ij}$= $\{x_{ij},y_{ij},A_{ij}\}$,(x_{ij},y_{ij}) is the spatial location, A_{ij} is the direction attribute。
Output: Visual Context : R=$\{R_1,R_2,\ldots R_k\}$
Step 1:Using the infinite Gaussian mixture model to cluster the direction of trajectory in an unsupervised way, results in the visual context with the similar moving tendency. Represent trajectory direction as two dimensional vector, A=(Δx, Δy), which means the weight of horizontal x and vertical .Define direction sets A = $\{A_{ij}|1\le i\le N,1\le j\le n\}$, use two dimensional IGMM for clustering, estimate parameters of the 'active' components to get the PDF: $p(x)=\sum_{j=1}^{k}p(x\mid z=j)\pi(z=j)=\sum_{j=1}^{k}\eta(\mu_j,\gamma_j)\pi(z=j)$ 。 Get the trajectory class subset : A = $\{A_1,A_2,\ldots A_k\}$, based on this, get the trajectory subset :TS_1, TS_2, …TS_k, TS_i = { T_{i1}',T_{i2}',…T_{iN}'},T_1'$\in T_1$. The result TS_i may segment the original trajectory into sub-trajectory, we call this sub-segmentation T_{ij}'as the segmentation of T_{ij}.
Step 2:Using hierarchical clustering technique to do trajectory spatial analysis。 For every TS_i = { T_{i1}',T_{i2}',…T_{iN}'}, members of which have similar moving direction, analyzing spatial similarity of trajectory. 1) Initialize T_{i1}' as first r_1. 2) From 1 to M: M is the number of routines: Based on the similarity measure d, compute the distance between T_{ij}' and r_m. a) Matching the trajectory against routine: if the distance is more than a threshold, merging this trajectory into routine, r_m, then accumulate the visual corner points of r_m as [r_m, T_{ij}'], because the main direction is known, displacing the corner points, finally using the Douglas-Peucker algorithm to reduce the corner points, updating r_m with these points. b) If there is none matching template routine, initialize a new routine with T_{ij}', updating the number of routines with M = M+1; Finally get R_k= $\{r_1,r_2,\ldots r_M\}$.
Step 3: Out put the complete trajectory sets :R = $\{R_1,R_2,\ldots R_k\}$。

3 Experiments

In this paper, we use the tracking algorithm proposed in [14], including the FCGMM far-field background model and tracking method based on the combination of particle filter and scene feature. We get the set of trajectores during the morning in which the routine taken frequently by the woker as our test datasets. In figure[4] we get the background image with the trajectory drawn on it, obviously spatial points distribute disorderly.

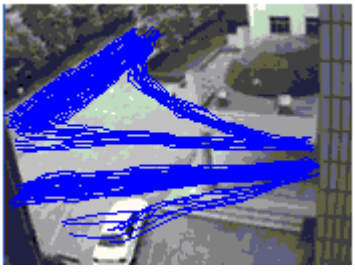


Fig. 4. Background image with trajectories

In figure[5], we do the IGMM clustering on the trajectory direction. We find 4 main direction clusters with the arrow point to the mean point, then we draw regions of the different cluster covering in figure[6]. Folloing each derection is key routines clustering image which come into being based on the direction clustering analysis and described by PD-point lines. Obviously the key routine's dispersal indicates the variance of the trajectory cluster set. As mentioned, our algrithm can segment the visual scene into semantic sub-regions in an effective way.

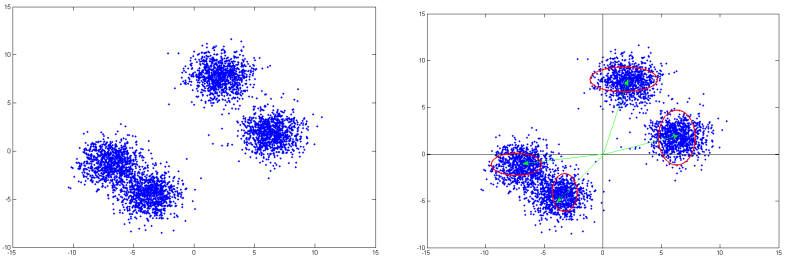


Fig. 5. Direction distribution

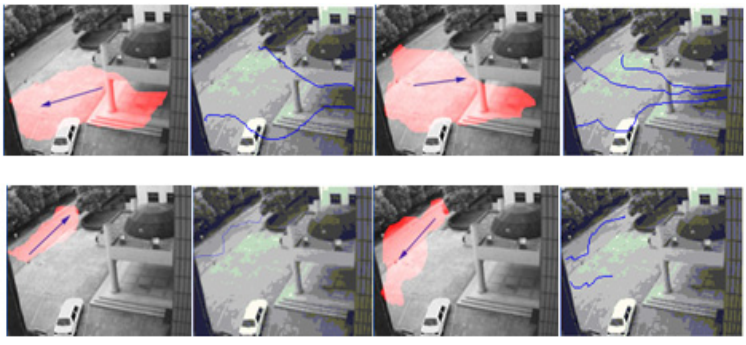


Fig. 6. Direction and routine image

4 Conclusions and Future Work

In this paper we address the problem of automation extraction of the visual scene in the visual surveillance application. Firstly, we analyze the trajectory information distribution to get the characteristic information. Then by using Gaussian mixture model we do cluster analysis of the moving tendency reflected by trajectory direction. For similar direction trajectory segments, we do hierarchical clustering based on the proposed similarity measure to get key routines of the scene. Finally we do experiments based on the surveillance video trajectory datasets.

Based on the results of the algorithm, we get the visual context of the surveillance scene. We can see the context reflecting intuitive and reasonable visual perception in the eyes of human beings, and can help us more in the deeper research of surveillance scene. For example, it can prompt the precise of moving object tracking by supplying rules of walking in the scene. In the future, we will focus on more moving attribute clustering to get greater visual context description.

Acknowledgements. This work is supported by the National Science Foundation (61170158).

References

[1] Hu, Tan, Maybank: A Survey on Visual Surveillance of Object Motion and Behaviors. IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews (2004)

[2] Turaga: Machine Recognition of Human Activities: A Survey. IEEE Transactions on Circuits and Systems for Video Technology (2008)

[3] Lavee, G.: Understanding Video Events: A Survey of Methods for Automatic Interpretation of Semantic Occurrences in Video. IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews (2009)

[4] Tao, X.: Beyond Tracking: Modeling Activity and Understanding Behavior. International Journal of Computer Vision (2006)

[5] Davis, J.W., Bobick, A.F.: The Representation and Recognition of Action Using Temporal Templates. In: IEEE Conference on Computer Vision and Pattern Recognition (1997)
[6] Stauffer, C., Grimson: Learning Patterns of Activity Using Real-Time Tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence (2000)
[7] Wang, X., Tieu, K., Grimson, W.E.L.: Learning Semantic Scene Models by Trajectory Analysis. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3953, pp. 110–123. Springer, Heidelberg (2006)
[8] Makris, D., Ellis, T.: Path Detection in Video Surveillance. Image and Vision Computing (2002)
[9] Wang, X.: Trajectory Analysis and Semantic Region Modeling Using A Nonparametric Bayesian Model. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Patter Recognition, CVPR (2008)
[10] Stauffer: Estimating Tracking Sources an Sinks. In: Proceedings of the Second IEEE Workshop on Event Mining (2003)
[11] Kang, J.: Mining Trajectory Patterns by Incorporating Temporal Properties. In: Proceedings of the 1st International Conference on Emerging Databases (2009)
[12] Kang, J., Yong, H.: Spatial-temporal discretization for sequential pattern mining. In: Proc. of International Conference Ubiquitous Information Management and Communication (2008)
[13] Rasmussen: The Infinite Gaussian Mixture Model. Advances in Neural Information Processing Systems (2000)
[14] Jiao: Research on Moving Object Detection and Tracking Methods for Intelligent Video Surveillance. Ph.D dissertation (2009)
[15] Li, J., Gong, S., Xiang, T.: Scene Segmentation for Behaviour Correlation. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 383–395. Springer, Heidelberg (2008)
[16] Wang, X.: Unsupervised Activity Perception in Crowded and Complicated Scenes Using Hierarchical Bayesian Models. IEEE Transaction on PAMI (2009)

Improved Method of the PSF Parameters Estimation for Image Restoration

Zeng Liang[1], Zhou Yang[2], and Yang Yingyuna

[1] Communication University of China, Beijing, 100024, China
[2] Renmin University of China, Beijing, 100024, China
zengliang1987@126.com

Abstract. The identification of the PSF can be converted to the identification of the parameter of the degradation function. A fine estimation of the blur parameters is a key point for restoration of blurred images. In this paper we propose an approach for estimating PSF parameters. The blur angle estimation based on Hough transform is improved. In addition, some pre-processing steps are applied to the system, such as smoothed filter, spectrum optimization. Experimental results demonstrate that the proposed method estimates the blur angle and the blur length more accurately than the algorithm proposed by R.Lokhande.

Keywords: Image restoration, Point spread function (PSF), Hough transform algorithm.

1 Introduction

Relative motion between the camera and the object being shot during image capture leads to unpleasant image blur. The kernel of motion-blurred image restoration is to establish the parametric model of degradation function and accurately identify the parameters of the PSF.

In the model of degradation function, the PSF has two important blur parameters, which are blur length d and blur angle θ respectively. If the parameters d and θ of the motion blurred images are available, we can get the specific form of PSF, and then the blurred image can be recovered with the classical methods. Various techniques about how to detect two parameters are proposed by many scholars. Cannon proposed an identification using cepstrum of a blurred image [7]. Yitzhaky et al developed the algorithm using spatial properties of a blurred image [4].However, the stability and the recognition accuracy of these previous methods are not satisfying.

In this paper, we improve motion blur parameter estimation, which is based on the Hough transform, proposed by R.Lokhande. Experimental results show that by using the improved method, the estimation accuracy of the blur length and angle can be increased.

2 Analysis of the Motion Blur Model

An image, whose space coordinates and scene brightness level continuously change, is defined as the continuous image. The model of degradation function of the continuous image can be expressed as:

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 743–751.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

$$g(x, y) = f(x,y)*h(x,y)+n(x,y) \tag{1}$$

Where f(x, y) is the original image, h(x, y) plays the role of the PSF, n(x, y) is additive noise and g(x, y) is the degraded image. The operator '*' denotes the 2D convolution.

If there is displacement of image at the time t and the exposure time is T, then g(x, y) can be expressed as:

$$g(x, y) = \int_0^T f[x - x_0(t), y - y_0(t)]dt \tag{2}$$

Formula (2) is the mathematical model of degradation function.

For a discrete image, the discretization of formula (2) becomes:

$$g(x, y) = \sum_{i=0;j=0}^{d-1} f(x-i, y-j)\Delta t \tag{3}$$

Where d is an approximate integer for the number of motive pixels on the image. Δt is the time factor.

Judging from the physical phenomena, motion-blurred images are actually the images which through a series of overlapping delay and then to be cumulated. Therefore if we convert a clear image to uniform motion blurred image, this formula can accomplish it:

$$g(x, y) = \frac{1}{d} \sum_{i=0;j=0}^{d-1} f(x-i, y-j) \tag{4}$$

Using convolution algorithm to simulate the horizontal motion blur, the process can be expressed as:

$$g(x, y)=f(x, y)*h(x, y) \tag{5}$$

Where

$$h(x, y) = \begin{cases} \dfrac{1}{d} & 0 \le |x| \le d\cos\theta;\ y = d\sin\theta \\ 0 & otherwise \end{cases} \tag{6}$$

Here, θ represents the angle between the object motion direction and the horizontal direction. Formula (6) is the expression of PSF which is relative to the motion blur length and motion blur angle. It infers that as long as identify the value of d and θ, the specific expression of h (x, y) can be obtained.

3 Improved Method for Estimating PSF Parameters

3.1 The Basic Algorithm of Detecting the Direction of Movement

In most studies, the identification of the motion blur parameters is based on the Fourier spectrum of image. R.Lokhande supposes a method based on Hough transform to identify the blur angle.

By ignoring the noise and making the Fourier transform on both sides of formula (2), an expression can be obtained:

$$\begin{aligned} G(u,v) &= \int_{-\infty}^{\infty} f(x,y)\exp[-j2\pi(ux+vy)]dxdy \\ &= \int_{-\infty}^{\infty}[\int_{0}^{T} f(x-x_0(t), y-y_0(t))dt)]\exp[-j2\pi(ux+vy)]dxdy \end{aligned} \tag{7}$$

In accordance with the character of the Fourier inverse transform, the equation can be written as:

$$G(u,v) = F(u,v)\int_{0}^{T}\exp\left[-j2\pi(ux_0(t)+vy_0(t))\right]dt \tag{8}$$

Delimit:

$$H(u,v) = \int_{0}^{T}\exp\left[-j2\pi(ux_0(t)+vy_0(t))\right]dt \tag{9}$$

The formula (8) can be expressed as

$$G(u, v) = F(u, v)\, H(u, v) \tag{10}$$

Take Fourier transform for formula (10) and then acquire the logarithmic value of its module, which is called "homomorphic transformation".

$$\log(|G(u,v)|) = \log(|F(u,v)|) + \log(|H(u,v)|) \tag{11}$$

Where g (x, y) is the degraded image, f (x, y) is the original image, h (x, y) plays the role of a linear shift-invariant PSF; G (u, v), F (u, v), H (u, v) are respectively their Fourier transforms.

Make the Fourier transform for the stripe image after homomorphic transformation; in its Fourier spectrum we can see the parallel dark lines clearly. The angle between the dark lines and the x-axis is the blur angle which rotated 90°along the counterclockwise direction. By using the Hough transform to detect location information of these dark lines, the angle of motion blur can be estimated.

The algorithm, supposed by R. Lokhande, uses LOG operator to detect sharp edge of the Fourier spectrum, and then converts it into binary image. There is no clear boundary in the binary image after processing, which is inconvenient to detect the straight line through Hough transform. This paper puts forward an optimization algorithm, making the angle identification more accurate.

3.2 Improved Detection of Movement Direction

In the preprocessing operation, the Fourier spectrum of image is taken more process than the original method. In the Hough transform stage, we can set a threshold to exclude interference, which makes the detection of line orientation more accurate. The specific method is: (1) Get the Fourier transform of the images, and get the modulus of the transformed pixels, in order to make the boundaries of the image spectrum more clearly (Fig.1 (a) (b)). (2) Because the spectrum line after Fourier transform is illegible, it requires many times logarithmic transformations to enlarge the contrast. Experimental results show that 3 times is appropriate (Fig.2 (a) (b)). (3)

Apply the mean filter to process the image of steps (2) to remove miscellaneous point (Fig.3 (a)). (4) Implement the Fourier transform once again, then the spectrum with clear border can be obtained, and the results of Hough transform become more accurate. By setting a threshold for the processed image and then binarization processing it, we can exclude the interference of low-intensity point on the binary image (Fig 3. (b) (c)). (5) Conduct Hough transform to the binary image and get the maximum cumulative array. (6) Set the threshold of the blur angle to exclude interference of the horizontal and vertical lines. (7) One or several motion angles which correspond to the maximum of the cumulative array can be gained through the above steps. Subsequently by using mean operation to process these values, we can get the motion blur angle.

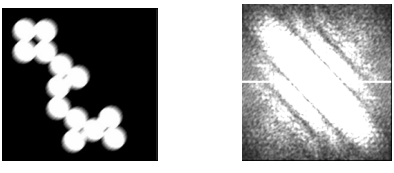


Fig. 1. (a) Original image ;(b) Fourier spectrum

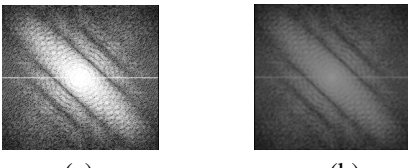


Fig. 2. (a)3 times logarithmic transform;(b) 5 times logarithmic transform

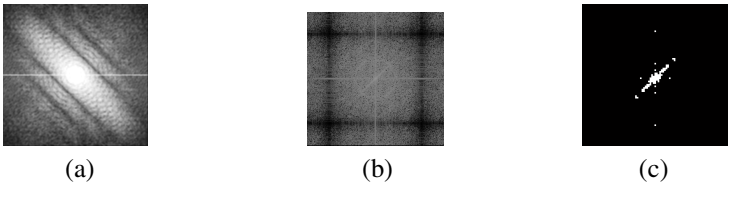


Fig. 3. (a)Fourier spectrum after filtering with the mean filter;(b) Fourier spectrum after 2 times Fourier transform;(c) binary image

Steps (1) to (4) are the pre-processing stages. This paper increases several pre-processing steps, such as logarithmic transformation, the mean filter and 2 times Fourier transform. These operations make the binary images clearer and more visible line which represents the direction of the motion. Consequently more accurate results can be acquired through the Hough transform operation.

In this paper, there are appropriate threshold processing in the Hough transform. We can get the maximum cumulative array M by using Hough transform to the binary image. In the range (M-Δm, M) there are N greater data points (M1, M2 ... Mn), which correspond to N motion estimation angle (θ1, θ2 ... θn). The value of Δm depends on the appropriate minimum number of θ values which can be found within the range of (M-Δm, M). By selecting a threshold value Δm, we can get one or several values of the suitable motion estimation angles θ. Then motion blur angle can be estimated using mean calculation for these θ values.

3.3 Improvement of Estimate Width between the Two Middle Spectrum Lines

Motion blur length d is related to the period of dark line in Fourier spectrum, and the zero spacing of spectrum is 1/d. However, digital images are discrete, so it exists a relationship between s and 1/d:

$$d = L/s \tag{12}$$

Where d is motion blur length (pixels), L is Fourier transform size; s is the width between the two middle spectrum lines.

From (12) we know that as long as in the Fourier spectrum dark line spacing s is generated, the motion blur distance d can be calculated.

In R.Lokhande algorithm experiments we found that using gray projection and difference algorithm to estimate the motion blur length is not very stable. It will lead to considerable error in some cases. The cause of these errors is the discrete image that makes it difficult to determine the minimum using the difference algorithm. This gray cumulative curve in this paper makes it more intuitive and more accurate.

This improved method is: accumulate projection of all the pixels to the horizontal axis, get gray cumulative curve, detect these minimum points on the curve , and then get their pitch. Fig.4 is the gray value waveform graph after accumulating along the stripe direction whose horizontal axis is the pixel accumulate by the column and the vertical axis is pixel gray accumulated value. We can get the width between the two middle spectrum lines by measuring two adjacent minimum point distances in the diagram. Simultaneously motion blur length can be calculated by formula (12).

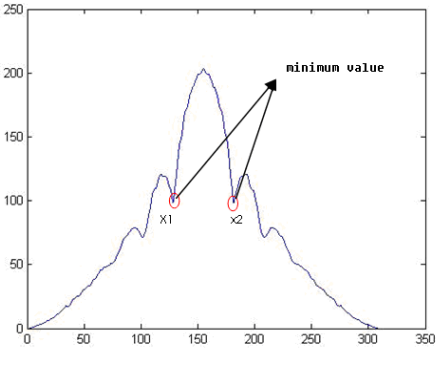


Fig. 4. Gray value waveform graph

3.4 Improvement of the PSF Parameter Estimation

Fig 5 shows the steps of PSF parameter estimation method.

4 Experimental Results

In the environment of Matlab 7.8, we distinguish blur angle and blur length of Fig 6. Experimental results are as follows.

4.1 Compare the Actual Blur Angle with the Test-Out Blur Angle

Fig.6 (a) (b) (c) are separately blurred 20°, 30° ... 80°(for each comparison, using a unified blur length; (a) blur 8 pixels, (b) blur 12 pixels, (c) blur 20 pixels). Then use this method and R. Lokhande algorithms separately to detect the blur angle. Fig.6 (a1) (b1) (c1) reflect the differences between actual blur angle and detected blur angle. Through the above analysis of three pictures we can see that the method which applies 2 times fast Fourier transform and binarization processing to original blurred images and then uses Hough transform to detect the angle is very accurate. The difference is limited within 1° or less. The accuracy of this algorithm in the paper is greatly improved compared with the R. Lokhande algorithm, and the measured angle is closer to the real blur angle.

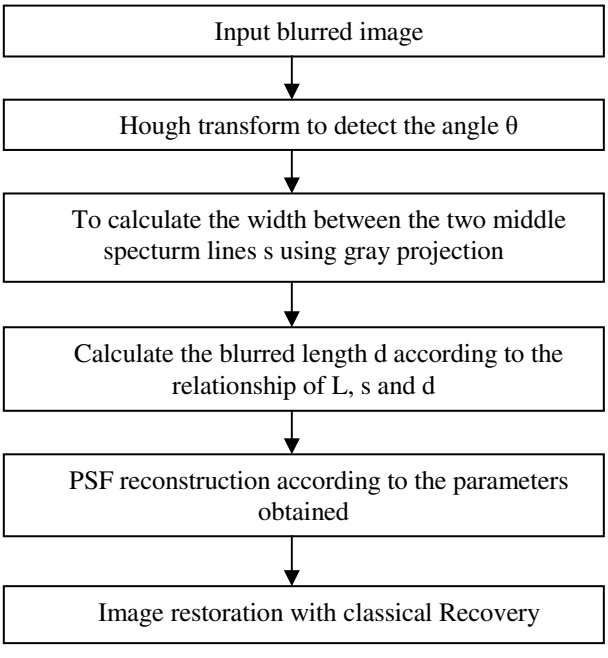
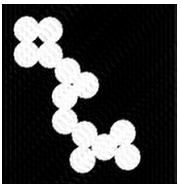
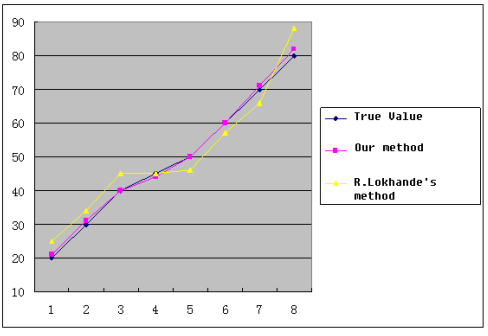


Fig. 5. Simulation flow chart of the algorithm

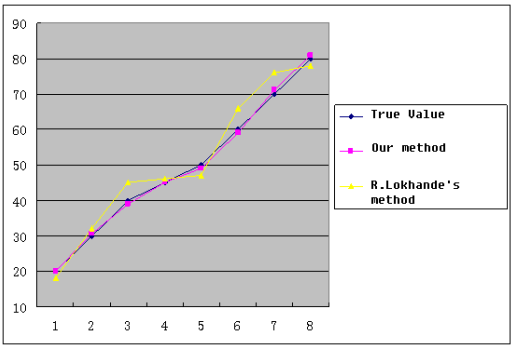


(a)



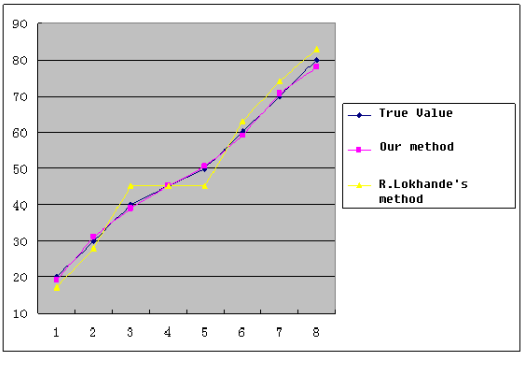
(a1)

(b)



(b1)

(c)



(c1)

Fig. 6. Compare the actual blur angle with the test-out blur angle

4.2 Compare the Actual Blur Length with the Test-Out Blur Length

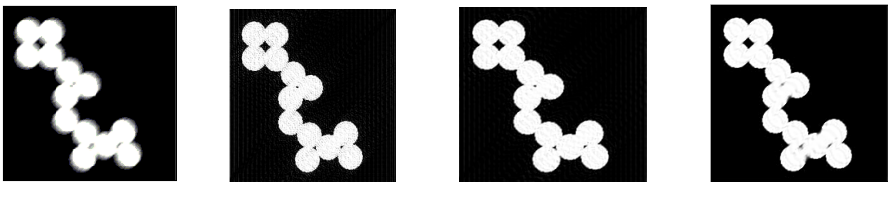
Table 1. Compare the actual blur length with the test-out blur length

actual blur length(pixels)	test-out blur length		
	Figure 6（a）	Figure 6（b）	Figure 6（c）
5	4.9091	4.7500	4.6542
6	6.0563	5.8795	5.9036
7	6.9180	7.0286	6.8333
8	8.0000	7.7500	7.7778
9	8.9583	8.8364	8.7143
10	9.9535	9.9592	9.9592
11	11.0769	10.7391	10.8889
12	12.0000	11.8049	11.7619

Using a unified blur angle (45°), we apply different blur length (from 5 pixels to 12 pixels) to Fig 6 (a) (b) and (c) for test. Then we use the method proposed in this paper to detect the actual blur length. Table I reflects the difference between actual blur length and detected blur length. Improved gray projection algorithm makes the measurement of distance more accurate.

4.3 Recovery Blurred Image with the Classical Method

There are many image restoration methods, such as Regularized Filtering, Wiener filtering and Lucy-Richardson algorithm. These algorithms need to know PSF to reconstruct the image. By using the proposed method to measure the PSF parameters, greater image restoration result can be obtained.



(a) (b) (c) (d)

Fig. 7. (a) Blurred image;(b) Deblurring Images using Wiener Filter;(c) Deblurring Images using Regularized Filter;(d) Deblurring Images using the Lucy-Richardson Algorithm

Through multi-group comparisons of experimental results that, Wiener filtering method is the best to deblur and obtain a clear outline, but there is much background noise; Lucy-Richardson algorithm is good to recover and has no noise, but the deblurring effect is the worst in the three ways; the effect of Regularized filtering is in the middle.

The above three results show that, the proposed method can accurately determine parameters of PSF. As long as we get the accurate parameters, we can use different deblurring methods to restore the image and get very satisfactory results.

5 Conclusions

This paper studies the estimation method of motion blurred image PSF parameters. We first set up motion blurred image degradation model, and then study how to properly establish the PSF. Further, we determine PSF parameters based on Hough transform and gray projection. Finally we analyze the experimental result to verify the accuracy of this method.

Acknowledgements. This work was supported in part by the "211 Project" of Communication University of China.

References

[1] Li, Q., Yoshida, Y.: Parameter estimation and restoration for motion blurred images. IEICE Transaction Fundamentals E80-A(8), 1430–1437 (1997)
[2] Moghaddam, M.E., Jamzad, M.: Motion blur identification in noisy images using fuzzy sets. In: IEEE International Symposium on Signal Processing and Information Technology, vol. (5), pp. 862–866 (2005)
[3] Lokhande, R., Arya, K.V., Gupta, P.: Identification of Parameters and Restoration of Motion Blurred Images. In: Elissa, K. (ed.) Title of Paper if Known, 1-59593-108, pp. 301–304. ACM, Dijon (2006) (unpublished)
[4] Yitzhaky, Y., Kopeika, S.S.: Identification of blur parameters from motion blurred images. Graphical Models and Image Processing 59(5), 310–320 (1997)
[5] Qi, H., Guo-Ji, Z., Jing, T.X.: Improvement of motion blurred direction identification based on Hough transform. Computer Applications (January 2008)
[6] Zhou, Y., Zhou, X., Dai, S., Peng, S.: An Improved Motion-Blurred Image Restoration Algorithm. Computer Engineering & Science (2009)
[7] Cannon, M.: Blind deconvolution of spatially invariant image blurs with phase. Trans., Acoustics., Speech and Signal Processing 1, 58–63 (1976)

Image Enhancement Based on NSCT

Ma Changxia, Bi Ye, Jiang Hongan, Chen Yunping, and Zhang Zhanqiang

Department of Computer Science Huaihai Institute of Technology, Lianyungang 222001 China
machangxia2002cn@yahoo.com.cn

Abstract. An algorithm for image enhancement based on nonsubsampled contourlet transform (NSCT) is proposed. NSCT is multiresolutional, localized, multidirectional and anisotropic so it can more effectively capture high dimensional singularity. Firstly, the coefficients at different scales and in different directions are obtained by image decomposition using the NSCT, then with these coefficients thresholds are adaptively set and the generalized nonlinear gain function is used to enhance the features with low contrast while protecting the strong contrast features from over enhancing in the NSCT domain. The experiment results show that the algorithm achieve a good effect.

Keywords: Image Enhancement, nonsubsampled contourlet transform, Multiresolutional.

1 Introduction

Histogram equalization [1] is one of the most well-known methods for image contrast enhancement. This technique makes use of global statistical information of the image and is not capable of diagnosing the local variations. These changes usually constitute the high frequency part of the image. So one way for image contrast enhancement is to enhance or amplify the amplitudes of these high frequency features.

Wavelet based contrast enhancement methods are proposed in the literature [2]. Unlike Fourier-transform, WT coefficients are partially localized in both spatial and frequency domains and form a multi-scale representation of the image [3]. Therefore the transformed coefficients can be successfully selected as the features of images for classification. However, it is well known that the commonly used separable extensions of one-dimensional transforms, such as WT, are limited in capturing the geometry of image edges. It cannot effectively represent textures and fine details in images for lacking of directionality.

The Contourlet transform (CT) is a flexible multi-resolution, local, and directional image expansion using contour segments. It can provide a multiscale and directional decomposition for images. Due to subsampling, the CT and WT are lack of shift-variant and result in pseudo-Gibbs distortion and blur edge of pavement cracks. The NSCT is built upon iterated nonsubsampled filter banks to obtain a shift-invariant directional multiresolution image representation. Thus, the NSCT is fully shift-invariant, multi-scale, and multi-direction image decomposition.

In this paper a new approach for image enhancement is described using the nonsubsampled contourlet transform. The rest of the paper is organized as follows. In Section 2, image enhancement based on nonsubsampled contourlet transforms. Section 3 is the experimental results. Finally, conclusion is drawn in Section 4.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 753–759.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Image Enhancement Based on NSCT

2.1 Contourlet Transform

The contourlet transform (CT), firstly proposed by M.N.Do and M.Vetterli [4], is a new image representation scheme which owns a powerful ability to efficiently capture the smooth contours of image. CT is implemented via a two dimensional filter bank that decomposes an image into several directional subbands at multiple scales. This is accomplished by combining the Laplacian pyramid (LP) [5] with a directional filter bank (DFB) [6] at each scale. A Laplacian pyramid (LP) is first used to capture point discontinuities, then followed by a directional filter bank (DFB) to link point discontinuity into linear structure. The overall result is an image expansion using basic elements like contour segments. Fig. 1 shows a flow graph of the contourlet transform.

However, because of the downsampling and upsampling, the CT is lack of shift-invariance and results in ringing artifacts.

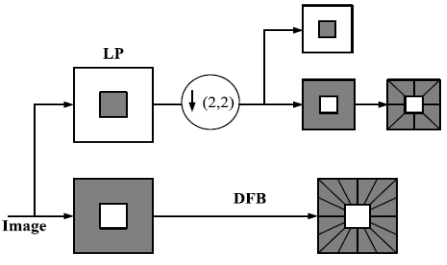


Fig. 1. Block graph of CT

2.2 NSCT

The nonsubsampled contourlet transform (NSCT) developed by Cunha, Zhou, and M.N.Do avoids the frequency aliasing of contourlet and enhances directional selectivity and shift-invariance [7]. The construction is based on a nonsubsampled pyramid structure that ensures the multi-scale property and a nonsubsampled DFB structure (nonsubsampled directional filter banks) that gives directionality. It is different from the CT that the multi-resolution decomposition step of NSCT is realized by shift-invariant filter banks satisfying Bozout identical equation (Perfect reconstruction, PR) and not LP. Because of no downsample in pyramid decomposition, the low-pass subband has no frequency aliasing; even the band width of low-pass filter is larger than $\pi/2$. Hence, the NSCT have better frequency characteristics than the CT. The two-level NSCT decomposition is shown in Fig.2.

The original image can be decomposed into a low-frequency approximate and high-frequency part by the algorithm that ensures multi-scale characteristic of NSCT. At the core of the NSCT structure is the 2-D two-channel NSFB. Shown in in Fig. 3 are the NSFBs.

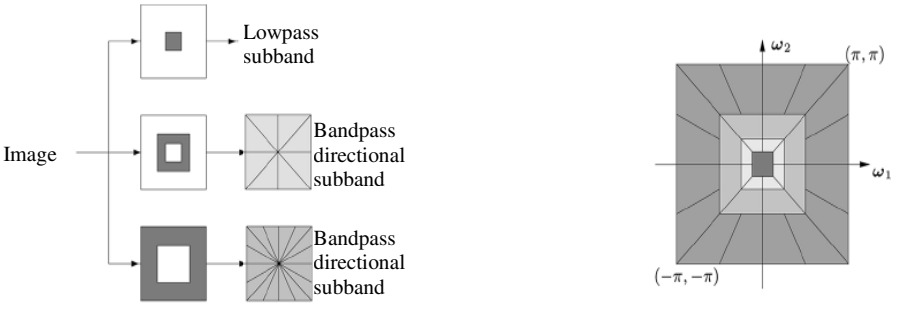


Fig. 2. NSCT (a)NSCT structure (b)Idealized frequency partioning

In this paper, the NSFB is built from a lowpass analysis filter H0(z) and H1(z) = 1−H0(z). The corresponding synthesis filter G0(z) = G1(z) = 1. The perfect reconstruction (PR) condition is given as

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 1$$

The Bezout relation puts no constraint on the frequency response of the filters involved. Therefore, to obtain good solutions, one has to impose additional conditions on the filters. As a result, we get a flexible multi-scale, multi-direction, and shift-invariant image decomposition. At the core of the NSCT is the nonseparable two-channel nonsubsampled filter bank (NSFB).

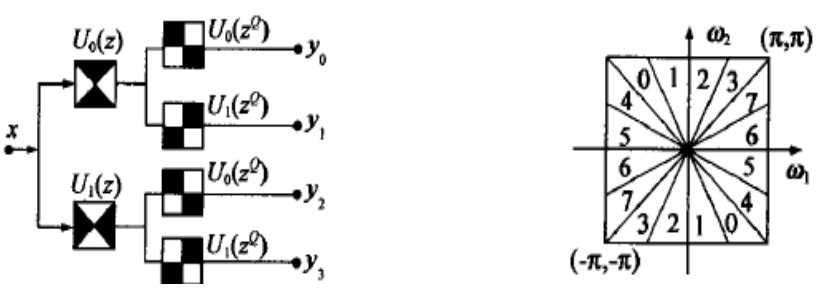


Fig. 3. Four-channel NSDFB (a)Filtering structure (b) Frequency decomposition

2.3 Image Enhancement in NSCT Domain

In practical problems, images contain thin edge, but also some noise. Therefore, the key and difficult of images enhancement are to enhance the contrast between weak information and the background, while suppressing noise and maintaining signals as far as possible without loss. The NSCT not only can accurately grasp characteristics of images but also can prevent edge blur caused by pseudo-Gibbs. This method includes two points: signal classification through selected threshold, and enhancement signals through the generalized nonlinear gain function.

- Threshold Setting

Threshold setting is crucial. Good threshold selection approach and enhancement function can enhance the weak edge and suppress the noise, rather than the clear edge is distorted and the noise component is enlarged in the enhancing process. Images to be enhanced include four categories information: image energy, clear edge information, weak details information, and noise. Corresponding to the four categories information, we analyzed the coefficients obtained by the NSCT transform:

Image energy is concentrated in low frequency;

The coefficients corresponding to obvious cracks information are larger in all high frequency subband;

The coefficients corresponding to faint details are larger in some high frequency subband, but in other high frequency subband they are smaller;

The coefficients corresponding to noise are smaller in all high frequency subband.

We mainly deal with the latter three types of information. So we first set the threshold T1 to separate the latter three types and image energy. Then, set the T2 to treat the latter three types. T1, T2 is related with the coefficients in the NSCT domain, rather than a fixed set of all input images. Threshold set by the approach can be adaptive with the change of pixels environment, and reflecting the gray trend of the current region.

First set the threshold value T1. As the gradient can reflect the image details and texture information, the gradient can be used as a measure.

The NSCT coefficients matrix is denoted as $\boldsymbol{f}_k^l(x, y)$, and then its grey gradient can be expressed as:

$$Grad[\boldsymbol{f}_k^l(x, y)] = [\frac{df}{dx}, \frac{df}{dy}] \overset{def}{=} \boldsymbol{\Delta f}$$

Its vector module is $|\boldsymbol{\Delta f}|$, $|\boldsymbol{\Delta f}| = \sqrt{(\frac{df}{dx})^2 + (\frac{df}{dy})^2}$ and then we can get the vector module matrix denoted as $\boldsymbol{Gf}_k^l(x, y)$.The value T1 is expressed as follow:

$$T_1 = c * Max(\boldsymbol{Gf}_k^l(x, y)) \tag{1}$$

Where $f_k^l(x, y)$ is the l-scale and k-subband NSCT coefficient matrix, Max() the maximum value of subband coefficients.

The threshold value T2 is needed to classify the latter three types of information. If the threshold is too small, the filtered noise signal is still there. But if the threshold value is too large, the important image features will be filtered out, thus causing image distortion. We use the median estimation method proposed in reference [8]:

$$\sigma = \frac{Median(|\boldsymbol{f}_k^l(x, y)|)}{r}$$

Where r is generally 0.6745, Median () is the median value of l-scale and k-subband coefficients.

The threshold value T2 is shown as (2):

$$T_2 = \delta * \sigma \tag{2}$$

Where δ is a parameter between 1 and 5.

- Generalized Nonlinear Gain Function

The key of enhancement is weaker parts and details. For strong parts, we only need to maintain in order to avoid over enhancement. For small coefficients, we can take them as noises, and make them zero. So we adopt generalized nonlinear gain function to increase high-frequency coefficients in NSCT domain. Usually the design of generalized nonlinear gain function shall meet the following requirements: Low contrast composition should be increased greater; steep edge should be retained. The function definition [9] is shown as (3).

Where 0<p<1 is enhancement amplitude of faint edge. P is smaller, and the rate of increase is greater. Mean and Max is respectively the mean and maximum coefficient in l-scale and k-subband.

$$f(\boldsymbol{f}_k^l(x,y)) = \begin{cases} \boldsymbol{f}_k^l(x,y), & Mean(\boldsymbol{f}_k^l(x,y)) \geq T_2 \\ Max(\dfrac{T_2}{|\boldsymbol{f}_k^l(x,y)|\wedge p}, 1) * \boldsymbol{f}_k^l(x,y), & Mean(\boldsymbol{f}_k^l(x,y)) < T_2 \ \ and \ \ Max(\boldsymbol{f}_k^l(x,y)) \geq T_2 \\ 0, & Mean(\boldsymbol{f}_k^l(x,y)) < T_2 \ \ and \ \ Max(\boldsymbol{f}_k^l(x,y)) < T_2 \end{cases} \tag{3}$$

3 Experimental Results

Our experimental computer's configure is Intel Pentium IV 2.93GHZ, 1G memory, 80G hard drive, Matlab7.0.4 is programming tool. The size of the image being processed is 512 * 512 pixels. Specific experimental steps are as follows:

- The input image is transformed by NSCT transformation and we can get the transform coefficients in different scales and in different directions.
- According to (1) and (2) to determine the threshold T1, T2, in which, c is 0.1, δ is 4, and r is 0.6745.
- According to (3), nonlinear gain function is used to enhance the transform coefficients, which p is 0.3.
- Enhanced coefficients are done by inverse transform to obtain the enhanced image.

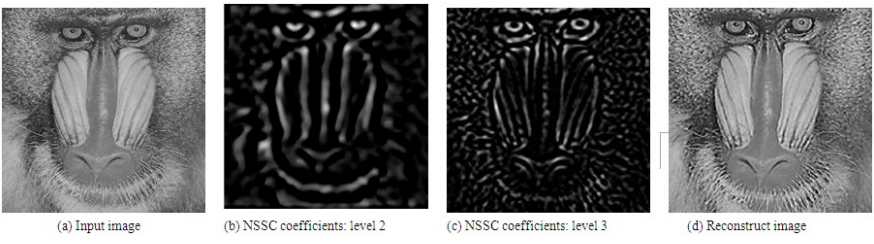


Fig. 4. Enhanced baboon image (NSSC:Nonsubsampled Contourlet Coefficients)

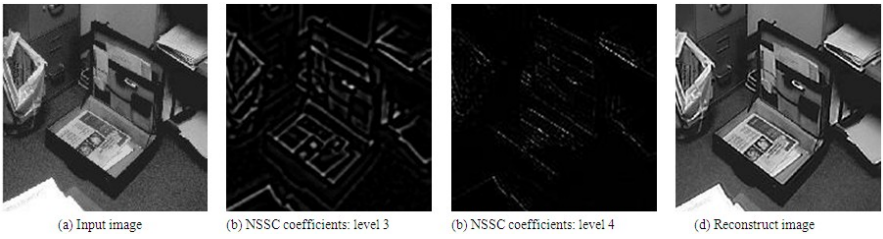


Fig. 5. Enhanced briefcase image (NSSC:Nonsubsampled Contourlet Coefficients)

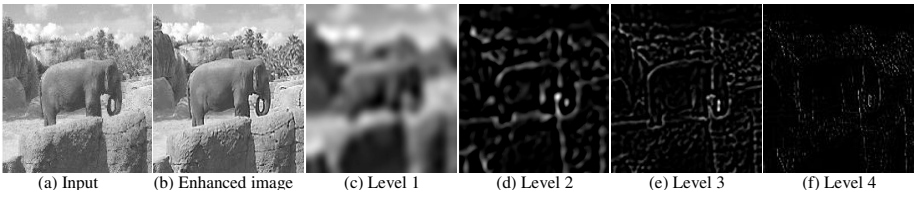


Fig. 6. Enhanced elephant image (NSSC:Nonsubsampled Contourlet Coefficients)

4 Conclusions

Using the characteristics of multi-scale, multi-direction and translation invariance, we proposed a new approach for image enhancement based on NSCT. Compared with other algorithms the method proposed in this paper can not only effectively enhance weak signals but also suppress noise while preserving the rich details of the image, and can well inhibit the edge blurring caused by the pseudo-Gibbs distortion. Experimental results show the effectiveness of the algorithm, but how to improve the timeliness and efficiency of the algorithm is issue to further study.

References

[1] Humme, R.: Image Enhancement by Histogram Transformation. Computer Graphics and Image Processing 6, 184–195 (1997)
[2] Zong, X., Laine, A.F., Geiser, E.A., Wilson, D.C.: Denoising and Contrast Enhancement via Wavelet Shrinkage and Nonlinear Adaptive Gain. In: Proc. of SPIE, Orlando, vol. 2762, pp. 566–574 (April 1996)
[3] Rioul, O., Vetterli, M.: Wavelet and signal processing. IEEE Signal Processing Magazine 8, 14–38 (1991)
[4] Do, M.N., Vetterli, M.: The contourlet transform: an efficient directional multiresolution image representation. IEEE Transactions on Image Processing 14(12), 2091–2106 (2005)
[5] Burt, P.J., Adelson, E.H.: The Laplacian pyramid as a compact image code. IEEE Trans. Commun. 31(4), 532–540 (1983)
[6] Bamberger, R.H., Smith, M.J.T.: A filter bank for the directional decomposition of images: Theory and design. IEEE Trans. Signal Proc. 40(4), 882–893 (1992)

[7] Cunha, A.L., Zhou, J., Do, M.N.: The Nonsubsampled Contourlet Transform, Theory, Design and Applications. IEEE Trans. on Image Process. 15(10), 3089–3101 (2006)
[8] Chang, S.G., Yu, B., Vetterli, M.: Adaptive wavelet threshoding for image denoising and compression. IEEE Transactions on Image Processing 9(9), 1532–1546 (2000)
[9] Liu, G.J., Huang, J.H., Tang, X.L., et al.: A novel fuzzy wavelet approach to contrast enhancement. In: Proceedings of 3rd International Conference on Machine Learning and Cybernetics, Shanghai, pp. 4325–4330 (2004)

Research on Design of Network Teaching System

Wenpeng Lu[1], Ruojuan Xue[2], and Haixia Li[3]

[1] School of Science, Shandong Polytechnic University, Jinan, Shandong, China, 250353
[2] LiShan College, Shandong Normal University, Jinan, Shandong, China, 250014
[3] School of Information Engineering, Shandong Youth University of Political Science, JiNan, Shandong, China, 250014
lwp@spu.edu.cn

Abstract. In order to satisfy the requirement of network teaching in computer classroom, the paper designs and implements an network teaching system for computer classroom. The software includes four functional modules, which are teaching demonstration of teachers, screen monitoring of students, remote instruction of teachers and transfer of teaching material. The software is suitable to network teaching circumstance in computer classroom, which is based on Client/Server model and is composed with teacher-side and student-side. The software is helpful for the development of educational informationization, which would promote the efficiency of teaching and learning in computer classroom.

Keywords: network teaching system, teaching demonstration, screen monitoring, remote instruction.

1 Introduction

With the development of educational informationization, all kinds of schools has built up a large number of computer classrooms, which are fit for practical course on computers. In practical course, teachers need to distribute many teaching materials to students and need to demonstrate detailed operations for them[1]. In traditional solutions, in order to facilitate the distribution of teaching materials, a data server is required to be set up, which is used to upload and download teaching materials; in order to facilitate the demonstration, a projector is required. The traditional solution would increase the hardware cost of computer classroom. Besides, the installation and maintenance of server and projector is a trouble and tiresome work.

In order to avoid the trouble in traditional solution of computer classroom, an network teaching system for teaching in computer classroom is designed in the paper, which is a simple and cost-effective solution. The solution can satisfy many kinds of requirements of teaching without any hardware cost. The network teaching system includes four kinds of functions, which are teaching demonstration of teachers, screen monitoring of students, remote instruction of teachers and transfer of teaching material. With the network teaching system, teachers can easily control the teaching procedure on computer classroom. This is helpful to improve teaching and learning effectiveness.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 761–765.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

In the paper, the detailed design and implementation of network teaching system for teaching in computer classroom are described. The rest of the paper is organized as follow. Section 2 introduces the framework of network teaching system. The detailed implementation of each module is described in Section 3. As last, we give the conclusion and future work.

2 Framework of Network Teaching System

According to function requirements of network teaching system for teaching in computer classroom, the software is composed with four models: module of teaching demonstration of teachers, module of screen monitoring of students, module of remote instruction of teachers and module of transfer of teaching material. The software is based on Client/Server model and involves teacher-side and student-side.

Module of Teaching Demonstration of Teachers. In order to demonstrate detailed operations for students, the module is required to realize demonstration of teachers, which transfers the screen of teachers to student side in real time.

Module of Screen Monitoring of Students. In order to monitor the operations and behaviors of student, the models is required to realize screen monitoring of students. With the model, teacher can monitor the student of any student. When some student is selected, the model would transfer his screen to teacher side.

Module of Remote Instruction of Teachers. In order to satisfy the requirement of remote individual instruction, the model is required to realize remote control function to student side. When a student ask teachers for help, teacher can control remotely student's computer and help him solve the problem.

Module of Transfer of Teaching Material. In order to facilitate the transfer of teaching material, the model is required to realize the function of data transfer, which allows the two-side transfer between teacher side and student side.

3 Implementation of Network Teaching System

3.1 Implementation of Module of Teaching Demonstration of Teachers

The module is responsible to realize the task of teaching demonstration of teachers. The teacher can broadcast his screen to all or part of students so as to demonstrate detailed operations. The module can capture the screen in teacher side, convert it to a special formation and compress it, then continuously transfer it to student side.

Sequence diagram of the model is shown in Fig.1

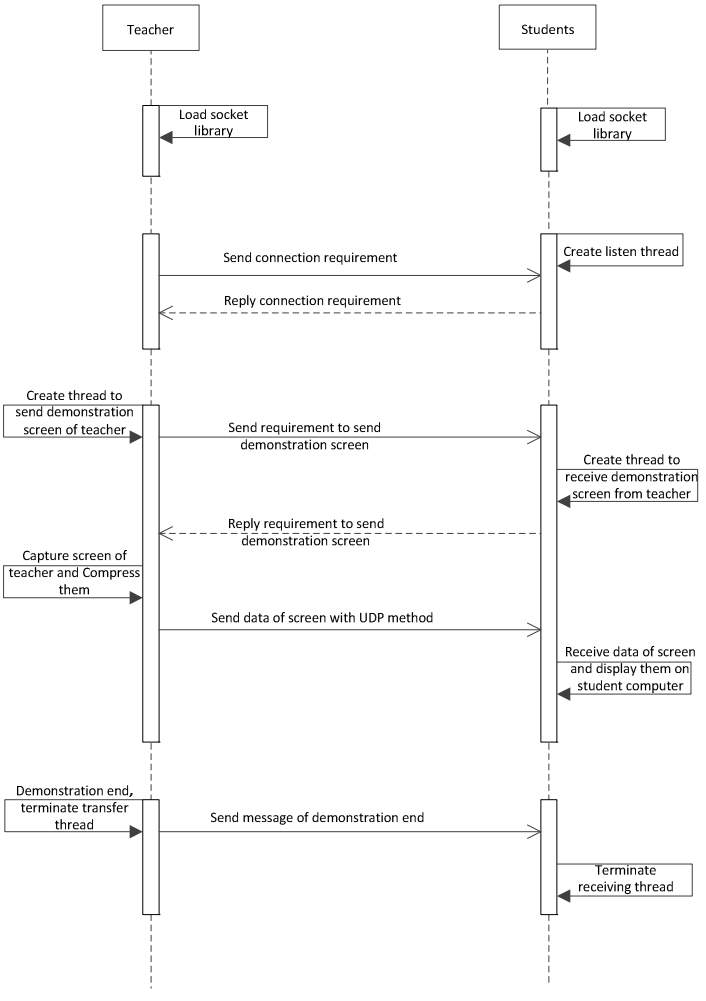


Fig. 1. Sequence Diagram of Module of Teaching Demonstration

The detailed program flow for teacher side is as follow[2, 3]:

1) Load socket library;
2) Send connection requirement to student side;
3) Create thread to send demonstration screen;
4) Send requirement to send demonstration screen to student side;
5) Capture screen of teacher computer and compress them;
6) Send data of screen to student side with UDP method;
7) Send message of demonstration end to student side;
8) Terminate transfer thread.

The detailed program flow for student side is as follow:

1) Load socket library;
2) Create listen thread;
3) Reply connection requirement to teacher side;
4) Create thread to receive demonstration screen from teacher;
5) Reply requirement to send demonstration screen;
6) Receive data of screen and display them on student computer;
7) Terminate receiving thread.

3.2 Implementation of Module of Screen Monitoring of Students

The module is responsible to realize the task of monitoring operations of students. Similar with module of teaching demonstration of teachers, this module capture the screen of student side, compress them and transfer them to teacher side. The sequence diagram and program flow is similar with Section 3.1. Therefore, we don't give unnecessary details for this model.

3.3 Implementation of Module of Remote Instruction of Teachers

The module is responsible to realize the task of remote instruction of teachers. The teacher can remotely control the action of keyboard and mouse of student computer with his keyboard and mouse on teacher side. Compared with module of teaching demonstration of teachers, this module need to transfer control message from teacher side to student side[4].

The detailed program flow for teacher side is as follow:

1) Load socket library;
2) Send connection requirement to student side;
3) Send requirement to instruct student remotely;
4) Create thread for remote instruction on teacher side;
5) Receive screen of student and display it on teacher side;
6) Send control message of teacher (keyboard, mouse) to student side;
7) Send message of instruction end to student side;
8) Terminate instruction thread on teacher side.

The detailed program flow for student side is as follow[5]:

1) Load socket library;
2) Create listen thread;
3) Reply connection requirement to teacher side;
4) Create thread for remote instruction on student side;
5) Reply requirement to instruct student;
6) Capture and compress screen of student;
7) Send screen of student to teacher side;
8) Receive messages from teacher side and control student computer with them;
9) Terminate instruction thread on student side.

3.4 Implementation of Module of Transfer of Teaching Material

The module is responsible to realize the task of transferring teaching material between teacher side and student side. Compared with the other modules, the function of this module is simply, which is easy to implement. Therefore, we don't give unnecessary details for this model.

4 Conclusion and Future Work

In order to satisfy the requirement of teaching in computer classroom, the paper designs and implements an network teaching system, which can help teachers demonstrate operations, monitor students, instruct students remotely and transfer teaching material conveniently.

Though the software has been designed and implemented, it only has basic functions for teaching in computer classroom. In the next work, we would try to make it more reliable and implement more functions for it.

Acknowledgement. This work is supported by two projects of Shandong Province Higher Educational Science and Technology Program (J12LN09 and J10LG20), China, and by Natural Science Foundation of Shandong Province (ZR2011FQ038), China.

References

1. Zhou, H.E.D.: A Study of the Construction of Multi-media Network Classrooms and the Coaching Method for Computer Experiment Classes. Journal of Jimei University (Natural Science), 167–172 (2001)
2. Luo, L., Zhan, Z. (eds.): Windows Network Programming. The People's Posts and Telecommunications Press, Beijing (2011)
3. Dai, Y. (ed.): Network Communication Programming Using Visual C++. Mechanic industry Press, Beijing (2011)
4. Wang, M., Qi, Z.: Design and Implementation of Multimedia Network Classroom Monitoring System. Computer Engineering 36, 291–293 (2010)
5. Liang, W. (ed.): Classic case of Network programming in Visual C++. Tsinghua University Press, Beijing (2010)

A Semantic Framework for Integration of Prison Internet

Peiliang Sun[1], Wei Wei[2], and Ming Xia[3]

[1] Zhejiang Police Vocational Academy, Hangzhou, 310018
[2] College of Information Science and Engineering, Henan University of Technology, Zhengzhou, 450003
[3] College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, 310023

Abstract. The Internet of Things (IOT) provides a good chance for prison control by integrating all sensors deployed in the prison into a unified monitoring system. However, because the data format and data semantic of the monitoring systems are not unified, the integration of systems in different prisons is difficult. In this paper, we designed a semantic framework to promote the data integration and interoperation of different prison monitoring systems. The framework provides a unified prison data source above concrete data sources, while every published data source is separately mapped to main semantic ontology model. At the same time, the framework provides different query and reasoning service interfaces for different high level applications and registry service. As a result, multiple heterogeneous sensor and stored data sources are integrated on-the-fly by using the semantic ontology data stored in knowledge base. With our framework, data interoperability, information search and retrieval and automatic inference can be realized.

Keywords: Semantics, Framework, Integration, Prison Internet.

1 Introduction

Prison security is a national challenge. To ensure a secure prison environment, different Internet of Things (IoT) infrastructures have been built to monitor the status of prisoners and the activity between prison and outside. With IoT [1,2,3], we can build condition monitoring systems in prisons and integrate them into a global monitoring system based on the wireless sensor network and cloud computing technology. Nevertheless, these systems are frequently designed, operated and maintained by different organizations, and the information format and the semantic information used are not unified, so it's hard to interoperate among these systems.

The prison environment is a restricted area, in which all prisoners, people and cars in and out are required to be monitored. In order to meet the above monitoring requirement, different sensors are deployed and connected to form a sensor network to gather various information, e.g., the location, status and identifier of all prisoners. The integration of this complex information has become a greatly challenge in building prison monitoring systems. In this paper, we designed a semantic framework to promote the interoperation ability of the different prison monitoring system. The goal

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 767–771.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

is to provide semantic interoperation middleware to the national prison monitoring network that can realize data interoperability, information search and retrieval, automatic inference etc.

2 Related Work

In the area of prison monitoring, solutions based on sensor networks have been used for several years for monitoring or controlling purposes. H. Xu et al. at Jiangnan University developed a small wireless sensor network to monitor the status of prisoners' handcuffs, in which sensors were integrated into handcuffs and were used to collect position and status information of prisoners [4]. In [5], an electronic watch which provides the utility to identify and locate prisoners in pre-established gateway was developed. In [6], S. Li provided a detailed illustration of a RFID application in public security domain of China. In [7], a real-time prisoner locating system based on RFID technology was created.

For interoperation problem, the proposition of Semantic Sensor Web [8] is to integrate the Semantic Web with sensors networks [9,10], which annotated sensor data with semantic metadata to increase interoperability as well as provide contextual information essential to situational knowledge.

Based on such mechanisms, semantic framework is proposed for different applications, such as federated ocean observation, environment decision support [11,12], etc.

3 Design of the Framework

The framework exploits ontology to support the semantic integration and operational collaborations between different monitoring systems. The proposed framework does not require a specific technical architecture such as Service Oriented Architecture (SOA). The collaborations are realized through unified semantic ontology. This provides the flexibility of implementation when SOA or web services are not available.

A service-based architecture exploited in our framework is shown in Fig. 1. From the aspects of end users, there are Data Acquisition Service (DAS) (event-based and query-based) and Representational State Transfer (REST) Service, Semantic Registry Service (SRS), Semantic Integration Service (SIS), and Storage Management Service (SMS).

Both DAS and REST Service interact with the backend data sources through a Semantic Integration Service. Here DAS is used for querying small amounts of data through standard web service interface which allows users to set flexible query conditions and makes various semantics-related queries and semantic data exchange, while REST is used for more resource oriented query. Any user can then execute queries over the resulting data source to explore aspects of the data, or be informed of interesting events as queries over the integrated source.

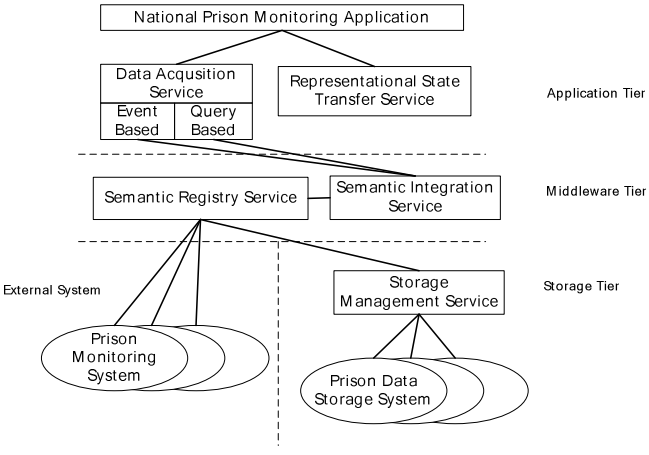


Fig. 1. The service architecture of the framework

Here SIS Service is used to combine data from multiple sources, which gives an ontological view over a set of data sources through mapping documents which relate to the conceptions in data sources to a unified data model. The mappings are expressed in terms of selections and transformations over the data sources, and can be created either manually or with the help of mapping tools.

A data provider could create the virtual source by providing a mapping document to the SIS through the Integration interface which states how to relate the source streams of the interface of the individual monitoring system. Or they can just provide a data storage system to the SMS Service which registers the storage system to the SIS service. Then SIS can serve any query on-the-fly, i.e., make real time transformation of semantic queries to traditional queries over raw sensor data or over interface exposed by individual prison monitoring system.

The SRS service accepts RDF (Resource Description Framework) documents describing the interfaces supported by the service, and the spatiotemporal and thematic description of the datasets. These semantic descriptions are expressed with respect to the network of prison ontology. The service supports the discovery of datasets by answering queries from DAS service, where queries have additional support for spatiotemporal constraints which makes it possible to query prisons based on the spatiotemporal and thematic coverage of the data, expressed in terms of concepts from the unified ontology.

The SMS service integrates data stored in repositories such as relational databases and files. This includes the retrieval of values from a prison monitoring source which are stored in a database, access to the historic values from the data stream or data associated with the deployment of the sensor network, e.g., the location of the sensors or associated threshold values.

4 The Prototype Implementation

To evaluate the framework design, we have implemented a prototype of the semantic framework for the Prison Internet on CubicWeb [13], which is a framework for building Semantic Web applications. It provides a programmatic environment for RDF, RDFS, OWL and RQL similar to SPARQL Language. The reasoning server is the open source Pellet [14] which provides standard reasoning services for OWL ontology. The query controller use Pellet to represent and reason about information using OWL. An ontology created from OntoSensor [15] is used in the prototype as an experimental ontology. In our implementation, the SRS service will transfer the RQL queries into SNEEQL, which queries relational database through web service. For example, from a monitoring system we get query result from database as an xml packet in following format.

```
<PrisonMonitorPacket>
<RawPeronalInfo>
        <Name>Pid</Name><ConvertedValue>7341852-53</ConvertedValue>
</ RawPeronalInfo >
<RawPersonalInfo>
        <Name>PersonelName</Name><ConvertedValue>Minggang
Hu</ConvertedValue>
</RawPeronalInfo>
<RawPeronalInfo>
        <Name>Sex</Name><ConvertedValue> Man</ConvertedValue>
</ RawPeronalInfo >
<RawPeronalInfo>
        <Name>Age</Name><ConvertedValue> 37</ConvertedValue>
</ RawPeronalInfo >
<RawPeronalInfo>
         <Name>Location</Name><ConvertedValue>cell-0305-413</ConvertedValue>
</ RawPeronalInfo >
</ PrisonMonitorPacket >
```

After semantic transformation, the data is as follows:

```
 <IOnto: PersonalName    rdf:ID="Chinese_Name_1">
      <IOnto:hasStringVal rdf:datatype=" #string">Minggang Hu
      </IOnto:hasStringVal>
</IOnto: PersonalName>
<IOnto:PersonalAge    rdf:ID="PersonalAge_AD_1">
      <IOnto:hasUnit rdf:resource=" #Years"/>
      <IOnto:hasIntegerVal rdf:datatype=" #integer">37 </IOnto:hasIntegerVal>
</IOnto: PersonalAge>
<IOnto:PersonalLocation    rdf:ID="PersonalLocation_Cell_1">
      <IOnto:hasUnit rdf:resource=" #No."/>
      <IOnto:hasStringVal rdf:datatype=" #string"> cell-0305-413</IOnto:hasIntegerVal>
</IOnto: PersonalLocation>
```

From the output we can see the all sensor readings follow formal semantics and can ensure semantic interoperability.

5 Conclusions and Future Work

The semantic integration of the national prison monitoring system can greatly improve the utilization of the potential of data resources on existed prison monitoring information systems. The semantic interoperability is a basic requirement for automatic search, retrieval and processing of prison internet sensor data. This paper is a step further towards to unify the domain objects, sensor readings, time/space semantics, and lifecycles of prison monitoring. The major contributions of our work are to automate the syntax and semantic annotation of sensor data and provide the semantic unification, which provides a solid base for integration of heterogonous prison monitoring systems to form a global system.

In the near future, we are planning to perform more comprehensive performance analysis by integrating more types of prison monitoring systems in China, which will be a further step in the direction towards enabling semantic web to access and process the national prison monitoring system.

References

1. Katasonov, A., Kaykova, O., Khriyenko, O., Nikitin, S., Terziyan, V.: Smart Semantic Middleware For The Internet Of Things. In: 5th International Conference on Informatics in Control Automation and Robotics, pp. 169–178. INSTICC, Madeira (2008)
2. Yan, L., Zhang, Y., Yang, L.T.: The Internet of Things: From RFID to the Next-Generation Pervasive Networked Systems. Auerbach Publications, FL (2008)
3. Brock, D., Schuster, E.: On the Semantic Web of Things. In: Semantic Days 2006, Stavanger, Norway (2006)
4. Xu, H., Fan, X.X., Wang, P., Peng, L.: Design of Things Network Node Used in Prison. Journal of Computer Research and Development 12(4), 317–320 (2010)
5. Tian, M.: The Application of Intelligent Sensor Monitoring System in Prison Management. Internet of Things Technologies 3, 28–29 (2011)
6. Li, G.S.: The Application of RFID Technology in Public Security Domain. China Security and Protection 4, 41–44 (2011)
7. Li, W.F.: The Real Time Personel Location System in Shushan Prison. Internet of Things Tehnologies 4, 30–31 (2011)
8. Sheth, A., Henson, C., Sahoo, S.: Semantic Sensor Web. IEEE Internet Computing 12(4), 78–83 (2008)
9. Thirunarayan, K., Pschorr, J.K.: Semantic Information and Sensor Networks. In: 24th Annual ACM Symposium on Applied Computing, pp. 1273–1274. ACM Press, New York (2009)
10. Brunner, J.S., Goudou, J.F., Gatellier, P., Beck, J., Laporte, C.E.: SEMbySEM: a Framework for Sensors Management. In: 1st International Workshop on the Semantic Sensor Web, Herkalion, Greece, pp. 19–34 (2009)
11. Cameron, M.A., Wu, J.X., Taylor, K., et al.: Semantic Solutions for Integration of Federated Ocean Observations. Proceeding of Semantic Sensor Networks, 64–80 (2009)
12. GrayA, J.G., et al.: A Semantic Sensor Web for Environmental Decision Support Applications. Sensors 11, 8855–8887 (2011)
13. Cubicweb, `http://www.cubicweb.org`
14. Pellet, `http://clarkparsia.com/pellet/`
15. Russomanno, D.J., Kothari, C., Thomas, O.: Building a Sensor Ontology: A Practical Approach Leveraging ISO and OGC Models. In: The 2005 International Conference on Artificial Intelligence, Las Vegas, USA, pp. 637–643 (2005)

Architecture of the Cloud Computing Platform for Enterprises

Shi Xingjun and Xuezhen Zhu

Zhejiang Industry & Trade Vocational College, Wenzhou, Zhejiang, China

Abstract. The paper analyzed the current actualities of the IT application for enterprises, and pointed the main deficient issues existed. Based on the target analysis, it then gave a four-layer architecture of the cloud computing platform for enterprises: physical layer, platform layer, supporting layer and application layer. Meanwhile, it gave the data management and operation mode of the cloud computing platform, as well as its network configuration. Finally, the paper put forward that, once the cloud computing platform structured, it could benefit enterprises to reduce the input cost on IT tremendously, and improve the availability and expandability of IT resources immensely.

Keywords: Cloud Computing Platform, IT Resource, Data Center.

1 Background

IT (Information Technology) is now indispensable for most modern enterprises. It is extensively used in every aspect of their production and management, such as OA (Office Automatic), ERP (Enterprise Resource Planning), SCM (Supply Chain Management), CRM (Customer Relationship Management), CAM (Computer Aided Manufacturing), E-Business, etc. However, with the rapid popularity and promotion of IT application process and the mushroom expansion of data size, enterprises having to increase input in IT continuously, more and more problems hereby arise in data storage and management, shown as follows:

a. The current data centers in enterprises pay most attention on the IT equipment, but not the operational management. This obsolete mode of data management makes it difficult to satisfy the demand of the operational management.

b. Weak carrying capacity, isomeric information islands and incompatibility between different information systems are featured prominently in the rigid architecture of current data center.

c. The information Resources cannot be configured rationally or dispatched in time, which caused the low utilization and the serious waste of them.

d. Enterprises have to consume a tremendous amount of time and money for the construction, upgradation, configuration and management about IT.

2 The Objective

Hereby, it is imperative to rebuild the IT data center and improve the data management mode. A brand-new platform--- Cloud Computing, directing the trend of demand

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 773–777.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

& future and being of great extensibility, is urgently needed. With it, dynamic computing and storage resources allocation management can be easily implemented, and automatization of service could be achieved as well. Simultaneously, the cost of IT equipment, operation expense, manpower and energy consumption could be drastically reduced.

3 Design of the Cloud Computing Platform

A veracious evaluation of enterprises' IT service demand and data size is needed before the design of the cloud computing platform. Take the case of a typical minor enterprise for instance. Its cloud computing platform should be structured in 4 layers: physical layer, platform layer, supporting layer and application layer, as illustrated in Figure 1.

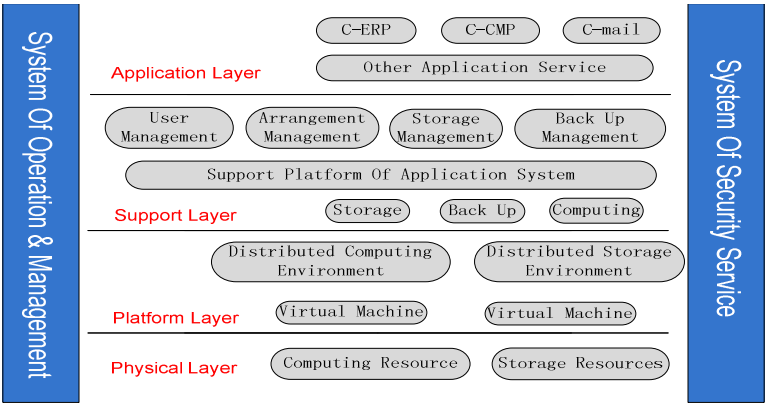


Fig. 1. Overall architecture of the cloud computing platform

3.1 The Resource Pool

The resource pool of cloud computing is distributed on the physical layer, including computing and storage resources, making up of a group of prefabricated server nodes and one-bit memories which can be automatically supplied and recovered. It is of great necessity to arrange disk arrays and servers according to the demand assessment of enterprises' functions and performances, such as the data scale, load balancing, dynamic configuration, etc. For a typical minor enterprise, customarily, 3 or 4 medium servers and an IBM or DELL disk array with a capacity of one-hundred T are enough as its resource pool. Meanwhile, for optimum efficiency,it is better to adopt servers with the same mode.

3.2 The Virtual Platform

On the platform layer, computing and storage resources are virtualized. We can integrate XenServer as the virtualized platform, as well as KVM, VMware, etc. Based on original hardware and software, we install virtualized applications on the bare servers,

and then virtualize several computing and storage environments, forming just as several cloud machines.

With the virtualized platform and its mighty cloud management capability, we could implement from application to virtualized arrangement, and the standardized installation of OS to the application arrangement.

3.3 The Support Layer

The support layer renders the basic service upon the administrator and users, including functions as follows:

- User management

There are 2 major roles on the platform: The administrator and users.

Users can submit applications when they need to acquire, increase or reduce their computing and storage resources, or when ending their projects or change the time. Moreover, they could operate most services on their own, such as restart the virtualized servers or reset the password.

- Disposing management

Disposing management provides fullly automatic services to users' applications. It deals with the applications from users automatically at a certain time,Thus what the administrator has to do is just to monitor the execution status and maintain the regular operation of the OS and applications.

- Storage management

The cloud computing platform manages to supervise both the internal and external storage resources. To meet users' requirements of large storage capacity and high transmission rate, SAN (storage area network) is a fairly good choice.

- Back-up management

The huge data sets emphasize the importance of back-up. The backup agent software is arranged into every business server. Once the backup resources and programs are defined, the data will be backed up from servers to designated memories on time. The backup procedure schema is shown in Figure 2.

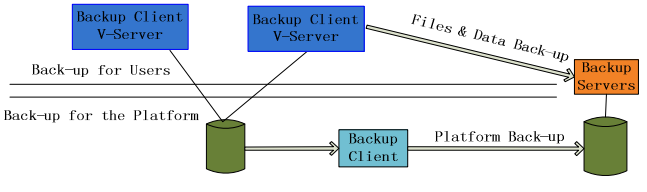


Fig. 2. The backup procedure schema

- Operation monitoring

The function of opeartion monitoring mainly carries on the real-time monitoring to the operation status of the virtual servers applied. It includes the basic computing status, the resources assigned, the utilization ratio, etc. The administrator is in charge of the monitoring, while users can keep abreast of the operation status of their resources.

- Security management

The cloud computing platform is capable of ensuring the isolation and security of each item. Users have easy access to the platform by login with authorized username and password, or the virtual machines of items with VPN authentication and VLAN authorization. It is the isolation of virtual machines and VLAN that ensure the security of the data, as illustrated in Figure 3.

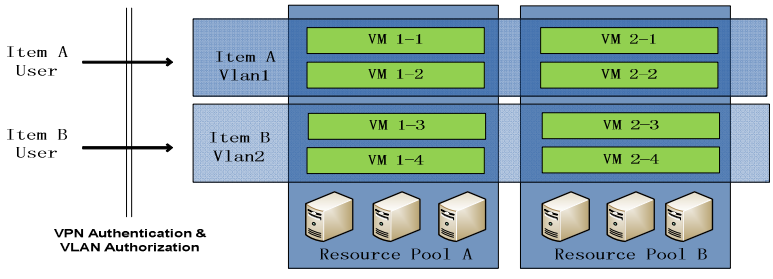


Fig. 3. Arrangement for the security management

3.4 The Application Layer

Enterprises can arrange Cloud-CMP, Cloud-ERP, Cloud-Mail and other cloud applications to demand on the application layer.

4 Design of the Network Configuration

Considering that the cloud computing platform is a complex system with numerous servers and immense data size, the design of the network configuration should comply with the principle of high security, high reliability, good extensibility and good compatibility.

The network configuration is divided into 3 layers: the entrance, the exit, and the core. The exit and the core adopt 2 core switches, and 2 firewalls are applied for a full redundancy connection. Furthermore, considering the extensibility and the compatibility of the network technology of future generation, we'd better choose ten-gigabit smooth transition equipment between the core layer and the entrance.

Taking a full cosideration of redundancy, arranging a gigabit link between the 2 core switches is essential in designing the backbone network.

When designing the exit, policy routing support should be considered, as well as the rational QoS control. In addition, we should ensure the security of the network system while executing the auto-switch between the 2 exits.

5 Conclusions

To see the tangible effect of the cloud computing platform, we take an ordinary company with a scale of one-thousand staffs as an example, as illustrated in Table 1.

Table 1. A compare between cloud computing & traditional mode

Items	Mode 1: traditional	Mode 2: cloud computing	Remarks
IT resources utilization ratio	20%	70%	The IT resources being demand-assigned & dynamic- allocated in mode 2
Human resourse time input	90min	10min	Taking the cost of per capita time input in IT software environment as an example
Capital investment	RMB5500	RMB3000	Taking the cost of per capita IT hardware for instance
Hardware life cycle	3-5 years	5-8 years	Update needed when the hardware is obsolete

In sum, the architecture of the universal cloud computing platform for enterprises can greatly improve the IT resources utilization, and drastically reduce the cost of IT equipment, its operation expense, human resource and energy consumption. However, enterprises differ in type, scale and business scope. Therefore, when designing the architecture, we should take into full consideration enterprises' actual demand and hereby take the superiority and availability of cloud computing platform to extremes.

References

[1] CSA. Security Guidance for Critical Areas of Focus Cloud Computing V2.1 (2009)
[2] Chen, K., Zheng, W.: Cloud Computing: System Instances and Current Research. Journal of Software 20(5), 1337–1348 (2009)
[3] North, K.: Cloud computing platform as service. Information week (2009)
[4] Wikipedia. Cloud computing (EB/OL) (March 10, 2009), http://en.Wikipedia.org/wiki/Cloud_computing
[5] Cloud computing versus grid computing (EB/OL) (March 03, 2009), http://www.ibm.com/developerworks/web/library/wa-cloudgrid
[6] Amazon elastic compute cloud(AmazonEC2), http://aws.amazon.com/ec2 (2008)
[7] Gannon, D.: Head in cloud. Nature (2007)
[8] Weiss, A.: Computing in the Clouds. Networker 11(4), 15–25 (2007)

The Application of SSL Protocol in Computer Network Communication

Tian Huan

Lanzhou Vocational Technical College, Gansu Lanzhou 730070, China

Abstract. Since nineties of last century , the world have feaced the Internet storm.From then on,network application have began to go into every aspect of our lives gradually. But the traditional Internet Protocol (TCP / IP) does not provide information transmission security mechanism.From the technical level, how to solve the problems of network information thefting, tampering, hacking attack have become a pressing matter of the moment of our science and technology workers. This paper is precisely based on this perspective. Through this views,we analyse and discuss the computer network communication protocol deeply. From the theoretical level, the SSL protocol is applied to the TCP protocol. From doing so,the problem of information transmission security that cannont be solved by traditional TCP/IP protocol will be solved successfully. At the last, an example of online bank specific will be discussed, the reaserch of this paper will be uesd in this example.

Keywords: TCP/IP, SSL;Signature, Information encryption.

1 Introduction

Since nineties of last century , the world have feaced the Internet storm.From then on,network application have began to go into every aspect of our lives gradually . With the remarkable achievements made at the same time, we also have to pay attention to an intensified phenomenon: of hacking in internet. What is the reason leading to the phenomenon of deterioration? This paper will discuss which factors lead to the occurrence of this phenomenon? At the same time, the corresponding technical means will be applied to correct this problem. At last, the latest research results applied to a sample unit specific description.

2 Basic Theory(TCP and SSL)

At present in the world computer network communication, we use the standard of TCP / IP[1] (Transmission Control Protocol / Internet protocol) protocol which was put forward by American Researchs. The protocol is designed for computer communication development application protocol, it through the hierarchical classification method, the complicated logic relationship into five distinct layers relationship. The five layers are: Application Layer, Transport Layer, Network Layer, Data Link Layer, Physical Layer. See chart below(Table 1):

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 779–783.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Table 1. The hierarchical structure of TCP/IP protocol

Application Layer
Transport Layer
Network Layer
Data Link Layer
Physical Layer

The TCP / IP protocol in the establishment of communication, which solve network congestion, processing network conflict has obvious advantages. But the agreement in the development stage, how to prevent the computer communication from hacking was not fully taken into account (also may not consider). So the agreement, not to the network security, information security for any given (from Table2.1 can be seen, the agreement of the five layer system structure, without any one layer according to the safety specific processing mechanisms). Through the computer network to transmit the information, are clear (unencrypted treated) transmission, and transmitted in the network without the corresponding information transmission tamper resistant, anti attack security validation. So through certain technical means, can be implemented easily steal each other's transmission information (all the final transmission data in the Table2.1 Application Layer), or for any party to transfer the information to distort, attack. How to solve the above problem, become the computer development, must solve the imminent problems.

In order to solve the above problems, especially the introduction of the NetScape SSL protocol[2] (Secure Socket Layer). The SSL protocol in TCP / IP protocol model of network layer and application layer (Network Layer) (Application Layer), which uses TCP to provide a reliable end-to-end security services, it ensures that the client / server communication between the not to be attacked, tapping, and always on the server for authentication, can also choose on the client side authentication. The SSL protocol also adopts the hierarchical method,, it also adopts a hierarchical classification method, the complicated logic relationship into four distinct layers relationship. This four layer respectively(Table 2):

Table 2. The hierarchical structure of SSL protocol

Application Layer
SSL Handshake Protocol Layer
SSL Record Protocol Layer
TCP layer

3 The Application of SSL Protocol in Computer Network Communication

Based on the traditional TCP / IP network communication information[3] without any encryption and authentication mechanism, we will apply the the SSL protocol which was introduced in section 2 to computer network transmission. Through respectively in client and server sides, SSL protocol is added to the TCP / IP layer and Application

Layer. The realization of the user identity confirmation mechanism (Digital signature), as well as to the transmission of information encryption system (Information encryption), specific see below(Fig 1):

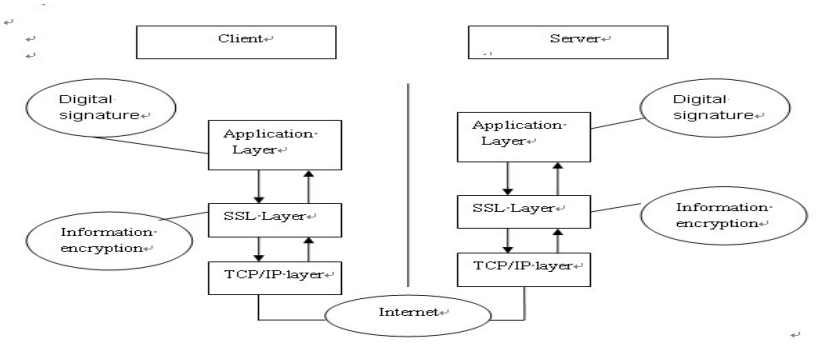


Fig. 1. The diagram that security communication with Digital signature

4 Example of Application

In this paper, the theoretical research results applied to some bank (due to network information security, so here inconvenience made the company's specific name) online bank transfer (Professional Edition) process of business process reengineering. Herein we dispose of two steps on the online bank transfer (Professional Edition) process are described, they are respectively:Online bank account application, Online bank accounts transfer.

First of all, we will explain of Online bank account application. The entire process of it will be show as below(Fig 2):

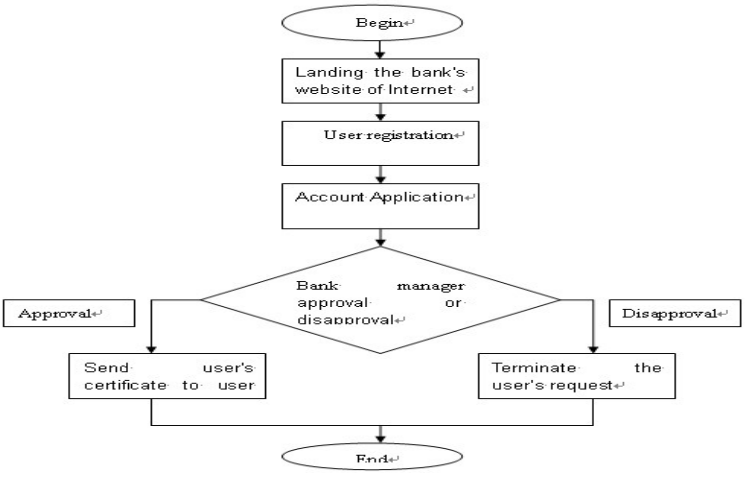


Fig. 2. Online bank account application

Finally, we will explain of Online bank accounts transfer. The entire process of it will be show as below (Fig 3):

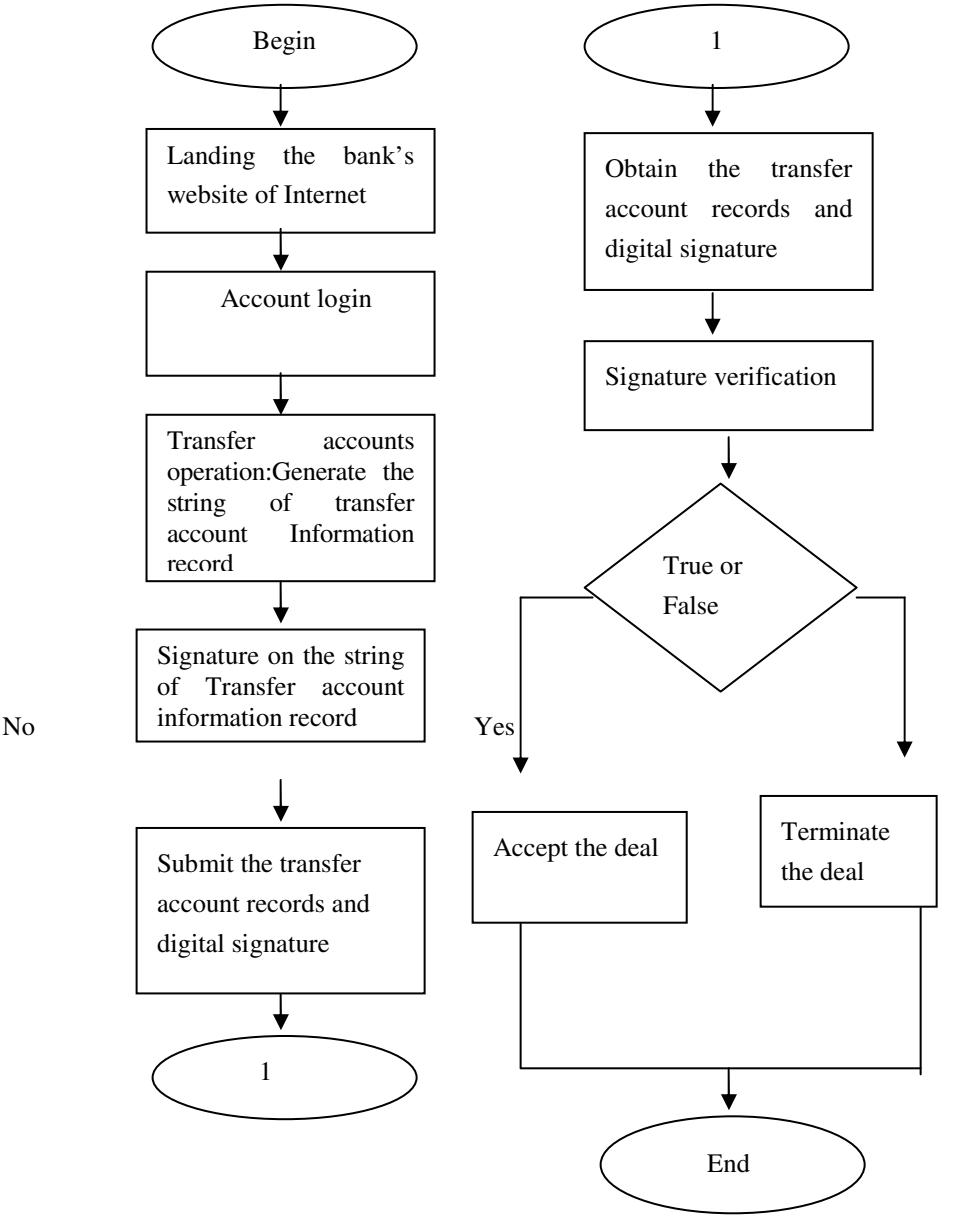


Fig. 3. Online bank accounts transfer flow chart

The user will import his own electronic certificate into the Internet computer (also can use this certificate placed in his portable U disk). Subsequently, the user lands located on the Internet of the bank, with oneself in the line once bank card information (user card + password inquiry) landing. During landfall, the bank not only to detect the user number and the password is correct, the simultaneous detection of the users have set the electronic certificate is valid. The information above all confirmed after the success, the user successfully landed in the bank system on the net. Otherwise, the landing failed, and gives the corresponding description information. Users of the site into the online bank transfer page, fill in the transfer of information, are completed, submit the transfer application. The system will now users transfer information to generate Generate the string Of transfer account Information record, and with the bank transfer system for electronic certificate public key encryption processing this information. Subsequently, the encrypted information to enclose the electronic certificate private key based on the user 's digital signature. After completion of the process, with a digital signature encryption information is transmitted to the bank server (Figure 3, 1). Bank transfer server receives the information, the decomposition of digital signature and encryption information. The user's public key to decrypt the digital signature confirmation, with its own private key to decrypt encrypted transfers information processing. The above information is successfully parsed, bank transfer system corresponding to the requested operation. The operation is successful, will request the success of information, using its own private key encryption, and attach a corresponding to the user's public key corresponding to the digital signature process. Users of the reverse of the operation, to be a success message. Thus, an online banking transactions completed successfully. If there any abnormalities or unrecognized certificate request, all operation failed.

5 Conclusion

In this paper, we discuss the computer network communication protocol for a more in-depth. From the theoretical level, through the introduction of SSL protocol.The problem of information transmission security which can not be solved by traditional TCP/IP protocol,was solved by using SSL protocol into TCP/IP protocol. Then through online bank specific examples, our theory research results of the application in practice. In order to achieve the theory comes from practice, and finally guide the purpose of practice. We hope that through this study, the method of solving computer network communication security will give other reaserchs some useful reference.

References

[1] Seth, S., Ajaykumar Venkatesulu, M.: TCP/IP Architecture. Design and Implementation in Linux 12, 40–220 (2008)

[2] Freier, A., Karlton, P., Kocher, P.: The SSL Protocol Version 3.0, draft-frier-sslversion3-02.txt (EB/OL) (June 18, 2006), http://wp.netscape.com/eng/ssl3/draft302.txt

[3] Karanjit, S., Chris, H.: Internet. In: Firewalls and Network Security, New Riders, USA, pp. 43–47 (1995)

Research on Wireless Network Security Model

Linmao and Wangyong

Department of Information Engineering, Xuzhou College of Industrial Technology, China

Abstract. Wireless network is an open architecture which is different form wired network. The open architecture is very convenient to communicate, but it also brings a lot of risk on network security problem. Therefore, this paper designed a new wireless network security model based on self-adjustment mechanism. It consists of three modules: detection module, safety model management module and backup module. After stimulate experiment in Swarm compared with traditional model PPDR, it shows the safety model designed in this paper can intercept dangerous data effectively.

Keywords: Software Engineering Education, Experience Based Approach.

1 Introduction

There is a trend that more and more people use their mobile devices on the way home or during waiting bus. The traditional wired network can't meet high demands on mobility, it's easy damaged, hard to locate and expand its network[1]. Therefore, smart mobile devices, such as smart phones, iPhone, android phone, tablets, iPad and so on, can meet the high demand on mobility network and become more and more popular in recent years. It's easy to bring and access to internet by wireless network, easy to be installed and low cost [2]. There are more and more corporations realize the importance of wireless network and add them to internet infrastructure. Now days, applications on wireless network are mainly based on the following area: traffic and travel services, public services, yard management, mobile office services, e commercial, smart community, and personal usage[3]. Enterprise and personal applications will take more and more important role in wireless network, and with the high development of wireless network and smart mobile devices, it will bring a deep revolution for people's life. Contract to the rapid development of wireless network applications, the network security is the bottleneck of wireless network. The wireless network has no physical connection node at all which is different with wired network. The data transferred in completely open environment, anybody can get the communication content by some specific tools in the coverage of wireless network[4]. The attacker can pretend to be a valid identity and then access to wireless network to attack and theft information. Wireless network is more dangerous than wired network for its open policy. Therefore, it's important and necessary to research on the security model on wireless network [5].

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 785–790.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Wireless Network Synopsis

2.1 Wireless Network Overview

Wireless network are referring to all types of computer networks which is not connected by any kind of cables. People can use this method to connect network devices in home, enterprise, hotel and telecommunications networks. This method can avoid bring cables into a building which is costly or a cable connection among various equipment locations [6]. Wireless networks are commonly administered and implemented by using radio waves. This method takes place at physical layer level at the OSI seven layers network model.There are many wireless network standards in use, such as 802.11, 802.11b, 802.11g, Bluetooth, HomeRF, WiMax, 3G (WCDMA, TD-SCDMA, CMDA 2000), 4G(LTE, TD-LTE) and so on[7].

2.2 Current Wireless Network Security Model

Conventional wireless network security model is based on static and open loop control system. This method can't effective response to dynamic network security thread and the low robust of wireless network system. According to high development of wireless network technology, the closed loop control security system is becoming more and more popular for its high dynamic response; its working flow diagram is showed in figure 1.

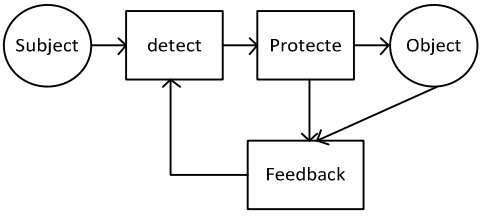


Fig. 1. Working flow of closed loop control system

There are many closed loop control system as PDR (protect, detection, and response) model, PPDR (polley, protect, detection, and response) model and APPDRR (analysis, polley, protect, detection, and response) mode.

3 Research on Wireless Network Security Model

3.1 Security Model Architecture

Since the wireless network is dynamic, the security model is also need to be dynamic. The model needs to not only keep the wireless network security in basic level but also can adjust the model itself according to the network system working status and keep the wireless network working even under attack. It's necessary for wireless network take distributed security model because of the disadvantages of centralize security

model. Current distributed security model can't dynamic adjust itself according to network working status. Therefore, this paper designed security model added self-adjust module to realize self-adjust functions. This module is designed to dynamic adjust security model between mobile devices and wireless network connection nodes, and keep the globalizations and security model update timely. The self-adjusting module is the main component in this paper designed security model, its architecture is showed figure 2.

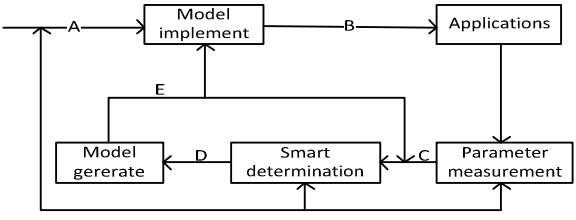


Fig. 2. The Self-adjusting module architecture

The self-adjusting module contains the following functions:

1. Capture the network package. We can't cascade out detection in wireless network which is different from wired network. And there are many package copies in wireless network; therefore, this feature is very important in wireless network security model.
2. Security detection. Self-adjusting module can matching the network package feature with dangerous packages. If the package is very similar with dangerous, it will handle the package into the next step virus detection.
3. Analysis the input and output to generate security model. Self-adjusting module analysis the input and output to dig out further data feature, and then generate a new security model which is more matching current network status.

3.2 Security Model Management

Security model management is very important in the whole security model system. According to the high development of wireless network load and scale, it's highly necessary for us to design a mechanism to manage these security rules. Therefore, this paper designed security model management module to manage security models. It's manage framework is showed in figure 3.

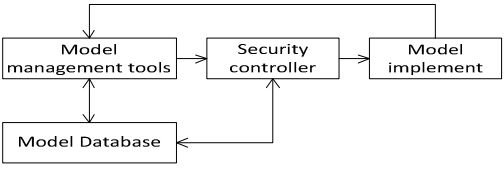


Fig. 3. Security model management framework

Security model management module consists of the following parts:

1. Security model management tools. It's the network interface for wireless network administrators, these tools can add and deleted user information, add, update, sync, and delete overall security rules, search the wireless network status and history logs.
2. Safety controller. It's a safety workstation or proxy server, which is designed to dispatch switch and pass down the security model.

As showed in figure 4, the working flow of security model management module is a layering flow. The network security rules in management tools needs to be combined with generated security rules in security model implement module and then execute the combined security policy, because overall security rules probably becoming the bottleneck of dynamic wireless network and the different demands of the different sub network defense rating.

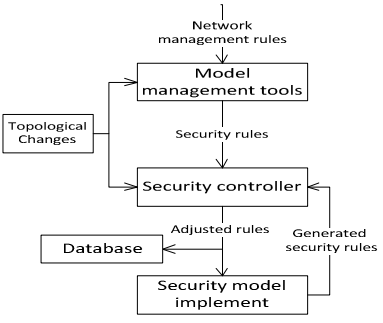


Fig. 4. Security model management working flow

3.3 Working Flow of Security Model

In order to improve the efficacy of wireless network security model, we need to set parameters of management tools first, and then interpret the network security rules into security models and dispatch them into model controller. The model controller combined these new security models with local security rules and processed them, then generated a final security model and sent it to model implement module by communication component in order to protect the wireless network safe.

If self-adjusting module find a mobile devices attack the wireless network in security model implement component, or attacks form outside of wireless network, the whole security system will record the attack behavior, target host, attack time, network protocol to security controller, the security controller will record them, notify dangerous attack to users and launch related safety response, finally send the updated security rules to security model implement component. If the attack is coming from internal network, the security model implement component can also notify AP to isolate the attackers, and update this network topology changes to security model controller.

3.4 Stimulate Experiment

We test the wireless network security model in Swarm stimulate software. Swarm is a stimulate software for agent based models. We test the ability of self-adjusting module of this paper designed wireless network security model and compare with conventional PPDR model (polley, protect, detection, and response). The swarm stimulate model framework is showed in figure 5.

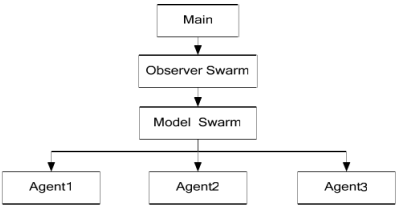


Fig. 5. Swarm stimulate frameworl

Main is the whole entrance of stimulate program and then we create a class named ObserverSwarm and an instance, after call the class, it will create a class named ModelSwarm and an instance, and then after call ModelSwarm class we will create graphics user interface and control panel, finally we can combine the diagram and data. ModelSwarm take charge in model execute, it will create couples of agents and control agents timescales. In this model, we setting the network flow from 1 to 1000, and choose network samples in every 10 nodes. Therefore, we get 100 network samples which are combined with attackers and normal network nodes. The stimulate result in first test is showed in table 1.

Table 1. The stimulate result in first test

Models	Network Samples	Samples Status Numbers(Attackers)		Test ResultsNumbers(Attackers)	
PPDR model	1000	0~500(100)	501~999(100)	0~500(100)	501~999(100)
This mode	1000	0~500(100)	501~999(100)	0~500(100)	501~999(100)

And then we change the network sample status, 200 attackers in 0~600 nodes, and 80 attackers in 601~999 nodes. The result is showed in table 2.

Table 2. The stimulate result in second test

Models	Network Samples	Samples Status Numbers(Attackers)		Test Results Numbers(Attackers)	
PPDRmodel	1000	0~600(200)	601~999(80)	0~600(120)	601~999(80)
This mode	1000	0~600(200)	601~999(80)	0~600(196)	601~999(98)

4 Conclusion

According to the high development of wireless network, it's more and more important to provide a safety environment for mobile devices. Therefore, this paper designed a

new wireless network security model with self-adjusting module and closed loop control system. After stimulate in Swarrn with PPDR model, it shows this paper designed wireless network security model can response to attackers effectively in dynamic network. The result gave a reference for researching on wireless network security model.

Acknowledgement. The work was supported by the Project of Jiangsu Provincial Department of Education of Social Science Foundation (2011SJB880048).

References

[1] Newman, M.E.J.: The structure and function of complex networks. SIAM Review (2003)
[2] Myers, C.R.: Software systems as complex networks: Structure, function, and evolvability of software collaboration graphs. Physical Review E Statistical, Nonlinear and Soft Matter Physics (2003)
[3] Botha, M., Von Soims, R.: The Utilization of Artificial Intelligence in a Hybrid Intrusion Detection System. In: Proceeding of SAICSIT (2002)
[4] Spitzner, L.: Honeypot - Definitions and Value of Honeypots (2003), http://www.tracking-hackers.com/papers/honeypots.html
[5] Yu, R., Wang, G., Mei, S.: Designing energy efficientrouting scheme with delay constraint for wireless sensor networks. In: Consumer Communications and Networking Conference (2006)
[6] Bouhafs, F., Merabti, M., Mokhtar, H.: A Semantic Clustering Routing Protocol for Wireless Sensor Networks. In: IEEE Consumer Communications and Networking Conference (2006)
[7] Eschenauer, L., Gligor, V.D.: A key-management scheme for distributed sensor networks. In: Proceedings of the 9th ACM Conference on Computer and Communication Security (2002)

Application of Network Platform in EFL Writing Feedback

Kang Hui[1] and Shi Ziming[2]

[1] Software School, Dalian University of Foreign Languages, Dalian Liaoning 116000, P.R.C.
cutekaitlyn@163.com
[2] School of Japanese Studies, Dalian University, Dalian Liaoning 116000, P.R.C.
szm_zz@163.com

Abstract. Computer technology can assist in giving and receiving EFL writing feedbacks. This paper aims at building an effective network platform with the directing idea of combining teachers feedbacks and students feedbacks together in a scientific way to achieve the goal of enhancing both students involvement and their confidence in the process of EFL writing learning; giving the right and appropriate directing feedbacks from the teachers; and simplifying and speeding up the process of the giving and receiving combined feedbacks.

Keyword: EFL writing, feedback, teacher and peer, network platform.

1 Introduction and Literature Review

1.1 Present Situation of EFL Writing Feedback

Providing feedback to students is an important part in the EFL Writing class in order to enable students to become aware of the recursive nature of the composing process [1]. Since the role of feed back in EFL writing classes and its effect on student writing was first recognized, various areas have been identified as key research areas. Basically speaking there are four kinds of feedback in the EFL Writing class, namely, teacher feedback, peer feedback, self evaluation and the combination of either two or three of the former three kinds of feedbacks.

1.1.1 Teacher Feedback

In the field of teacher feedback, researchers have focused on its content, style and effectiveness and have tried to determine what types of teacher comments are beneficial. Cumming argues that giving feedbacks in all aspects like word choices, grammar, rhetorical organization etc. would overwhelm any earnest student, and the available research is not clear as to whether error corrections lead directly to writing improvement [2]. Zamel implied that feedback should focus on what the writing means as a whole unit of discourse rather than on surface grammatical correction [3]. Truscott suggested that grammar correction should be banned due to the sparse evidence for its benefit and even suggested it could be harmful [4]. However, this

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 791–796.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

criticism has been contradicted by research which shows that although the reader can cope with minor grammatical errors if errors are more serious and very frequent there will be difficulty in knowing what the L2 writer means or what the writer is trying to say. In this case it appears that grammar correction is necessary. To both keep students' confidence and improve their grammar, the feedback mode of this paper chooses to show students main categories of grammar mistakes in class by citing examples and avoid the teachers' feedback in grammar and word choice.

The style of teacher feedback is also studied. It was found that specific questions in teacher feedback asking for information about students' personal experience or comments on syntax use lead to an effective impact [5]. Researchers have also suggested that helping "student ownership" in pragmatic thinking and meta-cognitive knowledge of the writing process is crucial [6].

So when giving feedbacks, teachers should focus on the general aspects like structure and style in an encouraging and concerning way.

1.1.2 Peer Feedback and Self Evaluation

Peer feedback provides students with opportunities to work cooperatively to examine each other's drafts and to give oral and/or written comments. It refers to the comments, suggestions and corrections that students provide on each other's writing for the purpose of revision. Responses to L2 writing from peers in classrooms help develop local communities of writers interested in and focused on students' writing [2]. Although through peer feedback, students involvement is high, the problems still exists: if the peers are not on the same English level, the credibility of feedbacks might get affected; peer feedbacks are not valued as highly as the teachers' feedbacks.

Self-assessment likewise helps to foster individual goals, self-monitoring of writing processes, and reflection about accomplishments.

Peer and self-assessment usually require orientation, coaching, or training for students who may not be accustomed to interacting with their L2 writing or to achieve the aims intended in a specific course [7] [8].

1.1.3 Combination Feedbacks

Some researchers have noted that only teachers feedbacks or peer feedbacks is not enough, as teachers feedbacks alone can not involve students actively and peers feedbacks are not accepted and taken by some students. As research [9] shows that 5% of the total revisions are influenced by peers and 35% of the total revisions are influence by teachers. Some research states that peer feedback should be combined with teacher feedback as in [10]. Yet most of the researches did not propose the specific way of combining.

The feedback mode proposed in this paper is to carry out two-time-anonymous-feedback with the aid of computer networks. The first feedback is carried out among peers in their groups anonymously to achieve the goal of objectiveness. The second feed back is carried out by the teachers in the aspect of structure, style and organization. The final mark of each student is composed of the student's feedback mark (first feedback mark+/-teacher's feedback) and the teacher's feedback mark. The detailed and specific scoring system will be discussed in Part 2, Construction.

1.2 Present Computer Application in EFL Writing Teaching

In modern teaching, a lot of computer appliances are employed to enhance the teaching itself. Presently there are three kinds of computer devices that are used in the field of writing teaching: teaching media (online writing labs); peer evaluating blogs and intelligence evaluating and scoring systems.

1.2.1 Online Writing Labs

Usually as an extension of the university writing center, OWLs offer service to those students who cannot come to the center for face-to-face consultation, allowing them to submit their writing online for feedback, asking questions online, or access the online writing resources. The best know online writing lab is perhaps the Purdue University Online Writing Lab (http://owl.english.purdue.edu/).

1.2.2 Blog

In addition to being an individual place for writing practice, blogs can also serve as a collaborative writing tool because bloggers can comment and give feedback to other bloggers in this online community.

1.2.3 Intelligence Evaluation

Intelligence evaluation needs comparatively higher technology. It bases on the machine translation technology, employing database to evaluate the writing and score it. Recent years, some researchers have been focusing on the construction of the online evaluation system by employing artificial intelligence.

1.3 Goal of the Platform

Based on the above researches and studies, it is concluded that one effective feedback mode including both the teachers' feedback and students' feedback with the assistance of the computer networks should be built.

The goal of the feedback platform is to involve the students in the active process of learning writing through feedbacks giving and writing revision (meta-cognition in writing); to build up students confidence in writing; to fully play the teachers' guidance role; to make teachers work effective and to reduce time and simplify the process of giving the combined feedbacks.

2 Construction

Based on the introduction and literature review above, the present network platform is constructed. The platform is directed by certain theories in feedback giving and receiving and is realized by using ASP.net and a database.

2.1 Feedback Thoughts in the Platform

As is discussed above, the ideal feedback should be the combination of peer feedback and teacher feedback with their respective focuses.

Peer review and feedback can be conducted first for the teacher to get the situation of both students' writing and feedback giving. When doing the peer feedback, certain groups should be divided, as certain researchers have found that students at the same or similar level doing the peer feedback usually achieves good effect. Before the peer feedback is actually conducted, the proper trainings and constructions should also be given in order to make the feedback appropriate. For example, the components of the score for each writing and the credits giving ways should be introduced and explained in details. We should also educate students to give feedbacks according to the principle of "appreciating first, and assisting second" to build students' confidence and the feeling of achievement [2]. At the same time the peer feedback might also be anonymous. Although this is not supported by any academic research (the writer is to research on the effect of anonymity on students writing and feedback giving performances) so far, there is hypothesis that anonymity shall avoid students from being biased or stereotyped because of the writers' daily academic performances during their feedback giving process.

Teachers' feedback should be focused on the structure and style, being generous and instructive. As Cumming, Zamel and Truscott have argued and proposed that too much focus on the grammar and spelling of the article shall reduce students' confidence and interest in practicing and improving their writing [2] [3] [4], and some other researchers also pointed out the importance of grammar and details correction, the platform proposed in this paper proposes the idea that the first peer feedback should focus on everything from grammar, spelling to the general structure, while teacher feedback focuses only on the general style, structure and coherent of the whole article. By doing this we can achieve the goal of practicing the detailed aspects of writing (presented by class teaching in catalogues and pointed out by peer students; mistakes pointed out by peers are less discouraging than those pointed out by teachers [11]), learning the general aspects of the writing (directed by the teachers), not being discouraged by the teacher feedback on grammar and spelling mistakes and saving the teacher's energy and time.

The final score is composed of each student's peer feedback giving score and the teachers' final score. Peer feedback has been carried out for a long time in the field of EFL writing teacher, it is stated that in students should be careful, responsible and encouraging in their feedback giving, yet so far there is no credit restraint on the peer feedback providers. The relevant education, training and coaching is crucial and necessary, yet the credit restraint may enhance the peer performance further in the feedback giving (the writer is to prove this by research). The specific implementing way is that students in the feedback system usually get two scores—one is the writing score (given by the teacher with reference of the score give by the student feedback giver) and the feedback score (given also by the teacher with reference of their performance in feedback giving). Each student's feedback giving score is decided by their attitude, efficiency and the difference between the score given to other student's writing by him/her and by the teacher. Of course a certain margin would be allowed in the difference, like 10 credits. For example, if student A gives student B 64, yet the teacher's score for student B is 85, the difference is 21. 21 minus 10 is 11, so "-11" is the feedback score student A gets, if his writing score given by the teacher is 80, the student's final score of the writing would be "80-11", "69". By scoring on both students writing

and feedback giving may greatly enhance student's level in feedback giving, thus in turn promotes their writing.

2.2 Technology Tools Employed in the Platform

As is discussed above, the process of two-time-anonymous feedback giving is objective, scientific, effective yet very complex. The network platform should be a very efficient assisting tool in the feedback giving and receiving process by putting all the related work after class. The platform is designed and constructed by ASP.net with an SQL database attached to it. The platform has three interfaces, namely, the administer interface, teacher interface and student interface. The administer interface will carry out all the back-stage management, including the teachers, students and exercises management; the teacher interface carry out the relevant work by teachers, like checking the peer feedback giving condition, pointing or arranging the anonymous peer feedback (anonymous among students yet clear to the teacher), and crediting the students' writing and feedback; the student interface helps to carry out student's tasks such as the writing composing and handing in, the feedback giving and receiving, they will also have an e-archive to record their practice and gaining in the writing course.

3 Conclusion

By constructing such a two-time-anonymous feedback network platform, the EFL writing teaching will be more effective and efficient.

It is scientific in feedback giving by combining teacher feedback and peer feedback appropriately, and by arranging the peer feedback in a certain anonymous group so that the efficiency and credibility of the feedback should be enhanced. Students involvement in the whole process of writing learning shall be enhanced greatly by giving them credits on feedback and their confidence would be built by giving encouraging and assisting feedbacks both by the peer students and the teacher. By employing the network platform the whole complex process will be greatly simplified and convenient. All related work can be done after class and the whole time involved and the teacher's energy can reduced to make the teacher's performance even effective.

References

[1] Paulus, T.: The effect of peer and teacher feedback on student writing. Journal of Second Lanuage Writing 8(3), 265–289 (1999)
[2] Cumming, A.: Curriculum and assessment: Trends in L2 writing research and pedagogy. In: Cheng, X. (ed.) Teaching and Researching English Writing as a Profession. Foreign Language Teaching and Research Press, Beijing (2010)
[3] Zamel, V.: Responding to student writing. TESOL Quarterly 19, 79–102 (1985)
[4] Truscott, J.: The case against grammar correction in L2 writing classes. Language Learning 46, 327–369 (1996)
[5] Ferris, D.: The influence of teacher commentary on student revision. TESOL Quarterly 31(2) (1997)

[6] Ferris, D.: Teaching writing for academic purpose. In: Flowerdew, J., Peacock, M. (eds.) Research Perspective on English for Academic Purpose. Cambridge University Press, Cambridge (2010)
[7] Cumming, A.: Goals for Academic Writing: ESL Students and Their Instructors. John Benjamins, Amsterdam (2006)
[8] Liu, J., Hansen, J.: Peer Response in Second Language Writing Classrooms. University of Michigan Press, Ann Arbor (2002)
[9] Connor, U., Asenavage, K.: Peer responses groups in ESL writing classed. Journal of Second Language Writing 3(3), 257–726 (1994)
[10] Fan, L.: The important role of peer feedback in EFL college writing. In: Cheng, X. (ed.) Teaching and Researching English Writing as a Profession. Foreign Language Teaching and Research Press, Beijing (2010)
[11] Ge, L.: Research on the effects of teacher feedback and peer feedback. Shangdong Foreign Languages Teaching 3, 54–57 (2011)

The Bayesian Network and Trust Model Based Movie Recommendation System

Du Wei and Chen Junliang

State Key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. Recommendation systems are services which recommend users new items such as news articles, books, music, and movie they would like. With the rapid development of information technology especially worldwide web, the information on the internet is exploring, how to screen out the information needed from multitude data in internet becomes a primary problem for solution. To solve this problem, a Bayesian network and Trust model based movie recommendation system is proposed, the Bayesian network is imported for user preference modeling and trust model is used to filter the recommending history data and enable the system to tolerant the noisy data. The simulation experiment use the movielens dataset as a source and validate the validity of the algorithm described in this paper, and a conclusion is reached that compared with collaborative filtering algorithm, the algorithm proposed holds advantage in the field of efficiency and noisy tolerant capability.

Keywords: Bayesian network, trust model, recommendation system.

1 Introduction

With the development of internet technology, the information on the internet is growing exponentially. People usually have to make decision on goods purchasing or service choosing without any priori information on them. To solve this problem, information filtering methods are proposed to remove unnecessary information from the original data. And also recommendation system is highly required to help customer make rational decision. Recommendation system is highly required to help customer make rational decision. Recommendation system is widely utilized for contents such as news, articles, movies, books and music. After the emergence of e-commerce, recommendation system development, predicting preferred manufactures and offering useful information, is fashionable to promote consumption. Recommendation Systems [1] based on AI technology have been explored especially around 1990s, when the WWW and internet services grew explosively all around the world. As the WWW evolved and the number of items with which were provided for users augmented considerably, it was getting more difficult for users to make choices of 'good' items. On the other hand, the evolution of the WWW has led various

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 797–803.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

researches on information retrieval techniques including NLP (natural language processing). Both ask for solutions which facilitates choices and technical conditions for implantation have made recommendation systems researched and technical conditions for implementation have made recommendation systems researched and developed actively. In additional, Network system complexity increased, large-scale users, environment tends complex, credible problem becomes the problem can not be avoided. So research on the recommendation model with the ability of avoiding cheating is of great significance.

2 Related Works

Regarding the problem of personalized trustworthy service providing in movie recommendation, lots of scholars in this field have made some achievements. Some of them with representatives can be listed as follows. Ahn, Shinhyun [2] paid attention to WoM's role as cultural metadata. 'Recommendation systems' are services which recommend users new items such as news articles, books, music, and movies they would like, developed a simple and low-costmovie recommendation system harnessing vast cultural metadata, about movies, existing on the Web, evaluated the system, and analyzed its strength. Kirmemis, Oznur [3] proposed a content based movie recommendation algorithm to make recommendations for the target user through building content based user models from collaborative-based user models and characteristics of the movie domain. Park, Seung-Taek [4] proposed a new ranking method, which combines recommender systems with nformation search tools for better search and browsing. Our method uses a collaborative filtering algorithm to generate personal item authorities for each user and combines them with item proximities for better ranking. Singh, Vivek Kumar [5] combined the content-based approach with sentiment analysis task to improve the recommendation results. The final recommendation list contains only those items that are both similar in content to the items liked by the user in the past, and also labeled as positive on sentiment classification. The researches mentioned above mainly focused on how to improve the efficiency of recommendation, yet the trustworthy attributes of the model proposed is neglected. Regarding this problem, the trust model is imported to the traditional recommendation algorithm based on Bayesian network with the filtered dataset.

3 Structure of Movie Recommendation System

The Global Design of the Recommendation System. The system architecture of our system is shown in Figure 1. The content database is constructed of four database including user database contains the user information included in the workflow of recommendation, movie database contains the movie information for handing, particularly about actors, directors and genre, rating database contains

rating information for specific movie from certain user's rating as input and calculates the trusted the rating value for further processing. Bayesian propagation engine will take trusted rating value from trusted filtering module and predict the probability the rating value the movie, which will be shown to the customer for four databases mentioned with user information, movie information and rating information received.

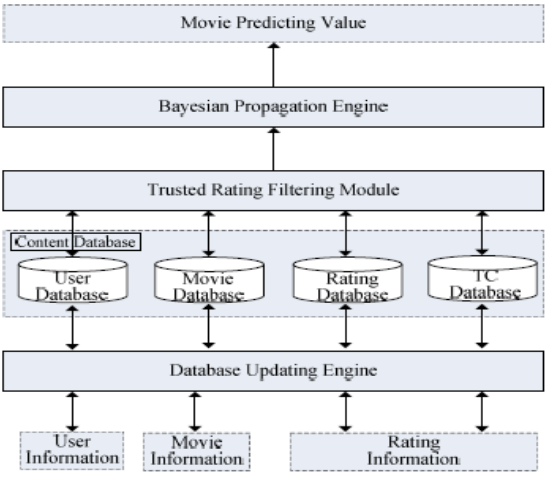


Fig. 1. Global design of recommendation system

Trust Model Description. Current research has pointed out that people tend to rely more on recommendations from people they trust like friends than on recommendations based on anonymous ratings. This factor is even more important when we are performing a group recommendations usually follows an argumentation process, where each user defends his preferences and rebuts other's opinions. Here, the trust between users is the major factor when users must change their mind to reach a common decision. A promising approach is to collect the trust knowledge from existing social networks. We believe that prediction accuracy of a user in the past is an important factor for measuring the trustworthiness of him and two different kinds of trust metrics, including Global Trust Metric and Trust Propagation Metric should be taken into consideration, which can be shown like Figure 2.

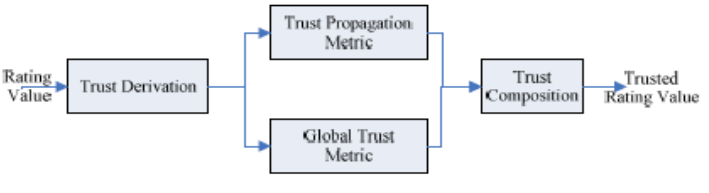


Fig. 2. Trust model workflow

Definition of Bayesian Network. A Bayesian network can be characterized into three categories according to the structure of the network and the conditional probability tables (CPT). The first is specified with the network structure and the conditional probability manually. For the second, both the network structure and the probabilities can be estimated from the data automatically. When both learning data and domain knowledge are available, the mentioned two categories of Bayesian network are the most practical and can be applied to some real-world problems and therefore are generally used. The third means that the variables which represent the user's impression on a movie will define the overall rating.

4 Working Steps Description

Evolution of Trust Model. With the rapid recommendation, trust value ought to be updated periodically. What's more, we should collect the trust metric and the local metric for the next step procedure.

To simulate the social network, friendship links should be generated randomly. Let's define a trust function for analyzing these links to calculate the trust among users which are different from the distance inside the social network. Herein we use equation (1) to evaluate the trust value.

$$trust(a,b) = \begin{cases} 1, & if\ distance(a,b) == 1 \\ 0.5, & if\ distance(a,b) == 2 \\ 0 & others \end{cases} \tag{1}$$

The predicted rating of item i for the user u_a by another user u_b is given by the expression (2), where a simple Resnick's prediction formula is utilized.

$$p_{a,i}^{b} = \overline{r}_a + (r_{b,i} - \overline{r}_b) \tag{2}$$

Where $\overline{r}_a$ and $\overline{r}_b$ denote the mean ratings of user u_a and user u_b respectively. $r_{b,i}$ refers to the rating of item i given by u_b, which also is the trust score of u_a in relation to u_b and then the trust score is derived by averaging the prediction error of co-rated items between the two users.

$$t_{a \to b} = \frac{1}{n(I_a \cap I_b)} \sum_{i \in I_a \cap I_b} \left(1 - \frac{p_{a,i}^{b} - r_{a,i}}{m}\right) \tag{3}$$

Where I_a and I_b are the set of rated items of user u_a and user u_b respectively, and m is the size of the rating range.

A user's global trust in relation to another user will combine the local trust with recommendations received from other users. The global trust score of user u_a can be

defined as the average of the local trust scores given by neighbors who are directly connected to u_a in the trust web.

$$gt_a = \frac{1}{n(NB(u_a))} \sum_{j \in NB(u_a)} t_{j \to a} \tag{4}$$

Where $NB(u_a)$ refers to the neighborhood of u_a. The sparse rating matrix usually leads to two users having no co-rated items, which will result in no direct trust relationships between them. To solve this problem, we can use trust propagation to infer the indirect relationships. In the trust web, a trust path between a source user u_s and a target user u_t exists in line with trust propagation. Assume that an intermediate user u_m is connected u_s and a target user u_t in the trust path. With the weighted average of the two direct relationships of $u_s \to u_m$ and $u_m \to u_t$, we can compute the trust score of u_t through u_m.

$$t_{\substack{s \to t \\ m}} = t_{s \to m} \oplus t_{m \to t} = \frac{n(I_s \cap I_m) t_{s \to m} + n(I_m \cap I_t) t_{m \to t}}{n(I_s \cap I_m) + n(I_m \cap I_t)} \tag{5}$$

The expression can be used when two users have more co-rated items.

When multiple paths between two users exist in the trust web, each path separately infers its own score.

Recommendation Process for the Movie. The recommendation process contains the following two steps: trust filtering and content based filtering. Firstly, a request is sent to find candidate user segments with information relating to a target movie and then for each target user segment, a request is sent to find appropriate solicitation points. Finally, personalized promotion information to target user segment will be sent by a operator.

5 Simulation Experiment

Design of the Experiment. The uniform resource locator is hosted on a small server with a AMD Sempron 3100+ CPU and 1GB of RAM. University of Minnesota has collected MovieLens data set by using the GroupLens Research Group. The well-known MovieLens recommendation system established in 1997 can accept the user of the film's score, which is recommended to user's personalized list of movies. Currently, there are more than 70,000 user recommendation records and more than 5000 movie. According to Movie Lens, data set of user score is an integer from 1 to 5. Moreover, the time necessary attributes also are included in the experimental data. The three data tables from

movie lens dataset will be needed for the experiment, which are user data table, data tables and the film score data table.

Result Analysis. The content of movie lens dataset are formatted and exported to database, and advertisement are illustrated with semantic annotation process, and then stored in the database consequently. In order to improve the quality of the dataset, users recommend at 10 movies will be needed. In the experiment, for Bayesian network training 90% of the dataset can be used and the rest 10% is reserved for testing.Comparing our algorithm with collaborative algorithm integrated in the movie lens dataset, we obtain the result as described in Figure 3. According to Figure 3, we know that as the number of users increases, the user interest model can continue to accumulate learning, the user interest model will capture the theme of science fiction growing areas when enjoying the same movies, and the accuracy of their recommendations continue to improve.

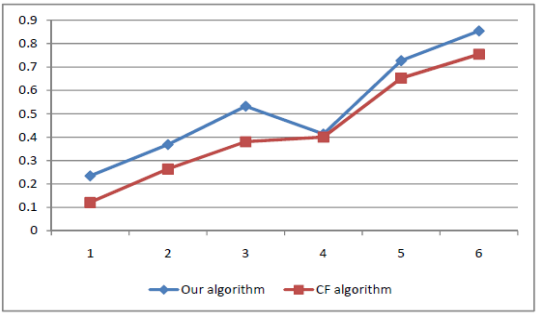


Fig. 3. Comparasion of our algorithm with CF algorithm

6 Conclusion and Future Works

In this paper, we have proposed a Bayesian network and a trust model based on movie recommendation system. The Bayesian network is used for preference modeling and trust model is imported for filtering the recommending history data which the system used has immunity to the noisy data. The simulation experiment has validated the validity of the algorithm proposed in this paper by using the movie lens dataset as a resource. And then we get the conclusion that the algorithm proposed in this paper holds the advantages in the field of efficiency and noise tolerant capability compared with the collaborative filtering algorithm. However, in the future, there are more work to be done on the performance improvement of the algorithm.

Acknowledgment. The work presented in this paper is supported by “973” program of National Basic Research Program of China (Grant No. 2011CB302704); National Natural Science Foundation of China (Grant No. 61001118, 61132001).

References

[1] Eric Hoster, R., Yoon, V.Y., Guimaraes, T.: Recommendation agent impact on consumer online shopping: The Movie Magic case study. Expert Systems with Applications 39(3), 2989–2999
[2] Ahn, S., Shi, C.-K.: Exploring movie recommendation system using cultural metadata. In: Tiergartenstrasse 17. Springer, Heidelberg (2009)
[3] Kirmemis, O., Birturk, A.: A content-based user model generation and optimization approach for movie recommendation. American Association for Artifical Intelligence, Chicago (2008)
[4] Park, S.-T., Pennock, D.M., DeCoste, D.: Applying collaborative filtering techniques to movie search for better ranking and browsing. American Associationfor Artificial Intelligence, Boston (2006)
[5] Singh, V.K., Mukherjee, M., Mehta, G.K.: Combining a content filtering heuristic and sentiment analysis for movie recommendations. Springer, Bangalore

The Network Security Management System Based on the Fuzzy Dynamic Evaluation

Deng Fan, Pei Tao, and Ji Hongcheng

Training Department, Engineering University of CAPF, Xi'an 710086, China

Abstract. Since the 21 century, the office mode of enterprise or business institution has rapidly developed toward digitization, informatization and network. At the same time the office efficiency is enhanced, the information security hidden danger has been a serious problem. This paper analyzes frame the structure and safety requirements of information network of the small and medium enterprise or business institution, designs and realizes network security management system (FDA-NSMS) based on fuzzy dynamic evaluation. FDA-NSMS collects data utilizing man-machine way, Delphi (Delphi) method is introduced for fuzzy dynamic assessment of the security degrees of network systems, and considering security measures of the management system, safety domain and the technology protection, information network are uniformly planned and realizing the multi-layer safety guarantee of the system.

Keywords: Information security, Fuzzy dynamic assessment, Network security.

1 Introduction

Digitization, informatization and network is the new operation mode of enterprise or business institution. It promotes office efficiency, perfects the management system and improves the decision-making scientificity, which has a lot of positive significance for the development of enterprises or business institution. However, due to the network system needs to connect external network such as Internet connection, which produces a degree of information security problems, therefore, how to maintain the information network security has become an important topic. For this reason, many institutions purchased some network security products, such as firewalls, invasion insist on system, vulnerability scanning system, network anti-virus software and so on, to protect the information network security from different sides, but also has shortcomings, such as limitations and dispersion, has a lack of unified management scheduling between each other, can't give network system bring all-directions protection [1].

The network security management system (network security management system based on fuzzy dynamic assessment) built in this paper ,which is based on fuzzy

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 805–811.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

dynamic evaluation, makes fuzzy evaluation for the information network system safety by collecting the static and dynamic two aspects of the data, provides the decision-making instruction for the entire system, at the same time considering the management system, safety domain and various technical protection means (including firewall, intrusion detection, virus detection, and data backup, etc), uniformly plans for the whole system, having ensured that the information security and normal operation of the system[2,4,5].

2 The Key Technology-Fuzzy Dynamic Evaluation Model

Because of the uncertainty and complexity of information security elements, and there are many difficulties such as the historical data is limited and data is hard to collect. It is difficult to use quantitative method to evaluate information security degrees and the traditional evaluation method is used more qualitatively. SunWeiGong and other people put forward a fuzzy comprehensive evaluation model which refers to WSR methodology and the hall three dimensions structure, selects evaluation factors with classification, uses the way of computer data acquisition, questionnaire survey method, field survey method to acquire data, scores each evaluation factors with the Delphi method, finally using the fuzzy comprehensive evaluation methods for the quantitative score, a good way to solve the quantitative evaluation of the information security degrees of information network system [3].

Because the fuzzy comprehensive assessment model depends on the expert scoring too much, it does not satisfy the real-time demand of network security management system. This paper will consider the questionnaire and on-the-spot investigation data as static data, its elements scores being from expert scoring; consider the real-time gathered data as dynamic data, its elements scores being from historical statistics experience, a fuzzy dynamic evaluation model is established which realizes the real-time evaluation of network security management system.The process of establishing the fuzzy dynamic evaluation model is following:

1) Establish the evaluation factors set X and divide the elements into two subsets: dynamic element XD and static element XS satisfy the condition $X = \{XD, XS\}$, $XD \bigcap XS = \phi$,the above two subsets respectively repeat steps 2)-7).
2) Divide the elements in subset (represented by XX) into m subsets: $XX_1, XX_2, ..., XX_m$ satisfy the condition $XX = \{XX_1, XX_2, ..., XX_m\}$, $XX_i \bigcap XX_j = \phi (i \neq j)$, and establish a evaluation subset

$XX_i = \{Y_{i1}, Y_{i2}, ..., Y_{in}\}, i = 1,2,...,n$ for each subset $XX_i, i = 1,2,...,m$, where n denotes the number of evaluation elements of XX_i.

3) Establish a comment set for all assessment elements, $V = \{V_1, V_2, ..., V_s\}$, where s is the number of comments.
4) Establish a fuzzy matrix, score the assessment elements which meet the condition $\sum V = 1$, evaluate all the single assessment element and get the matrix $R_i = (r_{ij,k})$,where $i = 1,2,...,m; j = 1,2,...,n; k = 1,2,...,s$ and $r_{ij,k}$ is the membership of single assessment element for comments [3].
5) Establish a weight set and show the weight of each evaluation factor in $XX_i, i = 1,2,...,m$:

$$A_i = \{a_{i1}, a_{i2}, ..., a_{in}\}, \sum_{j=1}^{n} a_{ij} = 1$$

6) Make a fuzzy comprehensive evaluation and calculate the comprehensive evaluation vector of $XX_i, i = 1,2,...,m$: $B_i = A_i \bullet R_i = (b_{i1}, b_{i2}, ..., b_{is})$.
7) Multi-grade fuzzy comprehensive evaluation. Consider $XX_i, i = 1,2,...,m$ as an alone assessment element and use B_i as the single element evaluation vector of Y_i which constitute the fuzzy evaluation matrix from XX to V :

$$B = \begin{vmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{vmatrix} = \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{12} & b_{12} & \cdots & b_{2s} \\ & & \vdots & \\ b_{m1} & b_{m2} & \cdots & b_{ms} \end{vmatrix}$$

According to the importance of XX_i in XX , the weight can be obtained: $A = (a_1, a_2, ..., a_m), \sum_{i=1}^{m} a_i = 1$ and the final comment vector is obtained: $TX = A \bullet B = (tx_1, tx_2, ..., tx_s)$.

8) Evaluate the dynamic and static elements jointly. According to the importance of XD and XS , the weight can be obtained: $C = (c_1, c_2), \sum_{i=1}^{2} c_i = 1$, the dynamic and static comprehensive assessment vector of

X is $T = c_1 \bullet TD + c_2 \bullet TS$, assign a value for the comment set $V = \{V_1, V_2, ..., V_s\}$,and the security degree is $G = T \bullet V = \sum_{i=1}^{s} t_i v_i$.

3 The Overall Design of FDA-NSMS

According to the above requirements analysis, the FDA-NSMS comprehensively considers the factors of management system, personnel distribution, security domain management, technology protection, emergency treatment and safety assessment. FDA-NSMS quotes separation of powers, the user group is divided into three roles: administrators, operators and inspectors. Administrator make operation specification, rewards or punishment system and the responsibilities of other two roles, and responsible for personnel allocation, overall control on the operation of the FDA-NSMS; The operator is responsible for all kind of specific function of FDA-NSMS, and its work relevant situation is recorded in log; Supervisor is responsible to supervise the implementation of the FDA-NSMS concrete function, and connecting with the log and various system, and put forward to the administrator reward or punish suggestions. In aspect of the personnel distribution, the FDA-NSMS request administrator position taken by institution vice president; the operator establishes the technical working group according to the institution technical force, and the group leader is business backbone and supervisors should be the backbone of the discipline inspection department who are familiar with the business. The two aspects of safety domain management and technical protection have stronger work technique, which in addition to need to deal with related business, but also need equip related equipment with good performance, such as computer room environment, computer terminals, server, the network media, RAID storage arrays and alarm and etc hardware, operating system, office software and a firewall, and intrusion detection software and antivirus software and etc software equipment. FDA-NSMS sets up hardware level firewalls between the information network system and Ethernet, decorate the intrusion detection system, and establish the virus database, equip online anti-virus software, in order to ensure data security, also sets up RAID storage arrays, regularly make backup for the network system database. FDA-NSMS also sets up a separate emergency treatment server, which makes some intelligent analysis on the error messages of information network system, launched the emergency response plans, and alarm through alerter and GSM etc way.

In order to ensure the real-time and efficiency of safety evaluation function, the FDA-NSMS sets up safety assessment server, according to expert opinion, historical data and real-time data, makes the quantitative evaluation on the safety degree of the whole information network system, and feedback the assessment results and processing opinion.

Table 1. The evaluation elements composition and weight

Static/ dynamic evaluation elements	Primary evaluation elements	Secondary evaluation elements	The elements composition
Static evaluation elements	Management system elements (0.3)	Organic safeguards (0.2)	The attention of leadership
			The position of Information security administrator in the institution
			Whether have special information security organization
		Budget planning (0.2)	Long-term and short-term strategy plan
			Budget and accounting for the proportion of the investment information
		Institution building (0.4)	Whether Information security system is sound
			The implement situation of information security system
		People management (0.2)	The specialized level of information security personnel
			The Information security personnel mobility
	Technology protection elements (0.4)	System security (0.3)	Server security measures (antivirus software and a firewall, and intrusion detection system and so on)
			The database security (database log, automatic backup of data)
		Application safety (0.7)	User authentication, access control
			The encryption protection of confidential data
			Data integrity verification
			All kinds of the event log
	Security domain management elements (0.3)	Quality of software and hardware (0.2)	Whether Computer brand, operating system and application software is genuine
			The network media brand
		Network plan (0.5)	Whether the network topology structure is reasonable
			Terminal access
			Among network node location and permissions
		Environment management (0.3)	The room environment (fire prevention, lightning electricity protection, dustproof)
			Whether computer room management personnel exists
Dynamic evaluation elements	Technology statistic elements (1)	Operating condition (0.2)	Whether The network system can run normally
		Event log (0.8)	The frequency that the information networks is paralyzed
			The crash frequency of server
			Database error rates
			Leak number
			Emergency treatment frequency
			The virus was testing frequency
			The frequency that the network is attacked

4 The Fuzzy Dynamic Evaluation Example of FDA-NSMS

FDA-NSMS collects all kinds of evaluation factors and data using the combined man-machine method, evaluation experts evaluate the static element with the Delphi method, and the appraisal results have been used in the effective period; The system uses historical experience data and real-time data to make automatic evaluation contrast of dynamic factors, evaluation elements composition and weight is shown in table 1, the score result of secondary evaluation factors at a time point of the FDA-NSMS is shown in table 2. Thus we can obtain the comprehensive evaluation vectors of dynamic and static elements:

$$TD = AD \bullet BD = \{0.36, 0.48, 0.16\},$$
$$TS = AS \bullet BS = \{0.3215, 0.5385, 0.14\},$$

Set up the dynamic and static weight as $C = \{0.4, 0.6\}$, then the final comment vector is $T = \{0.3446, 0.5034, 0.152\}$. Define the comment set as $V = \{100, 80, 20\}$, the security degree of FDA-NSMS is $G = 77.772$. According to the overall evaluation from experts, the data reflects the security situation of FDA-NSMS more accurately.

Table 2. The score result of secondary evaluation factors

Evaluation elements		Fuzzy relation		
		Absolutely safe	Safe	Not safe
The information security degree	Management system elements BS_1	0.25	0.60	0.15
		0.30	0.50	0.20
		0.35	0.40	0.25
		0.30	0.60	0.10
	Technology protection elements BS_2	0.40	0.50	0.10
		0.20	0.60	0.20
	Security domain management elements BS_3	0.35	0.55	0.10
		0.45	0.55	0
		0.40	0.50	0.10
	Technology statistic elements BD	0.60	0.40	0
		0.30	0.50	0.20

5 Conclusion

This paper constructs a network security management system based on the fuzzy dynamic evaluation, in view of present situation of the prevalence of information security hidden danger big, protective low level, the many kinds of technology protection measures lack of regulation schedule in present enterprise or business institution. The system uses the fuzzy dynamic assessment to evaluate the safety degree of information network system, and then guides the operation of the network security management system, and comprehensively consider of the management system, safety domain and various technical protection means (including firewall, intrusion detection, virus detection, and data backup), conduct uniformly security plan, ensure that the information security and normal operation of the system..

References

[1] Xie, Y., Liang, T., Zheng, L.: The design and implementation of network security comprehensive management platform. Communication Technology (September 2004)

[2] Xiang, J., Huang, T., Sun, D.: The design and implementation of the electric power enterprise information network security management system. Automation of Electric Power System 27(15), 71–74 (2003)

[3] Sun, W., He, D.: The fuzzy comprehensive evaluation of information security from the qualitive to quantitative. The Theory and Practice of System Engineering (12), 93–98 (2006)

[4] Baker, A., Oh Navarro, E., van der Hoek, A.: An experimental card game for teaching software engineering processes. The Journal of Systems and Software 75, 3–16 (2005); Clerk Maxwell, J.: A Treatise on Electricity and Magnetism, 3rd edn., vol. 2, pp. 68–73. Clarendon, Oxford (2009)

[5] Conn, R.: Developing software engineers at the C-BOJ software factory. IEEE Software 19, 25–29 (2010)

The Research and Design of IMS Network Management

Xia Ling

Network Center, SiChuan University For Nationalities, China

Abstract. Network management is important means to ensure the normal operation of the network. As an important component, it and other network technologies keep pace. IMS is a core network of service control initiated by NGN. With the separation of control and bearer and the separation of control and service, IMS can support many kinds of access modes, and can achieve the integration of fixed network and mobile network. At present, most of the research for the IMS is forced on the functional entities, interfaces, service quality and so on. As the result, the research on the IMS network technology will do good to the IMS development and reliable operation. Firstly, an introduction of the IMS network is given in this thesis, and the functional entities of the IMS and network management technology, such as TMN network management technology and SNMP network management technology are analyzed in detail. Then based on Net-SNMP and SNMP4J Open Source technology, the program of IMS network management is implemented on the J2EE platform. In the end, IMS network management system is tested on the Open IMS Core experiment platform, which also verified the feasibility of the system.

Keywords: IMS, TMN, SNMP, J2EE, Network Management.

1 Introduction

With the development of communication technology, IP-based business integration and network integration have become an inevitable development trend. Business drivers are a prerequisite for network development, and technology-driven is guaranteed by evolution of the network. Communication needs of users continue to promote the development of communication technology[1]. In the development trend of network convergence, IMS (IP Multimedia Subsystem, IP Multimedia Subsystem) has been the industry's attention. It breaks the existing network to provide a closed business model. It protects the interests of carriers. The current network can smoothly transit to next-generation networks. The future development is mobile multimedia services.

IMS is a global, independent access, and standards-based protocol for IP multimedia service control system. IMS is based on SIP, to support a variety of fixed and mobile access methods, to achieve the integration of fixed and mobile networks, and to provide a rich multi-media business. A high-quality network

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 813–818.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

without an efficient network management system can be said that the maturity of IMS network management system[2],which won't directly affect the actual deployment of IMS.

In this paper, the IMS network technologies, and network management technology has been studied. The demand for IMS network management to TMN and SNMP network management technology design the IMS network management system solutions. With J2EE, SNMP4J, Net-SNMP and other open source technologies, based on object-oriented, modular and hierarchical thinking programming achieve the overall framework of the IMS network management system including performance management module and the network element layer of the S-CSCF of the agent. The feasibility of the system is tested to verify in the Open IMS Core IMS network simulation .

2 IMS Network Management System's Overall Archite

Based on our multi-use classification in network construction of network construction, we design the IMS network management system. A hierarchical approach is adopted to network management so that the designed network structure have a good agreement. The design is simple and easy expansion. We refer to the TMN logical model of the network hierarchy IMS in the design of network management systems[3]. TMN logical model will be divided into a variety of business management layer, service management, network management and element management layer in accordance to the perspective of network management activities. IMS network management system design is based on modular, object-oriented thinking. The use of a layered architecture have a reusable and good scalability of the IMS network management system. Overall structure of the IMS network management system is shown in Fig.1.

3 Designing Experiments

In fig.1, IMS network management system is divided into the Network Element management, network management and databases. Network Element management: Achieve the data management of network elements and reduce the burden of network management. Network management can be well shielded network elements differences and achieve the same type of centralized management of Network Element. Network Management: Include interface module, database connectivity module and network management modules. Network management module is divided into the topology management, fault management, configuration management, performance management and security management sub-module. Database: Provide permanent preservation for network management data.

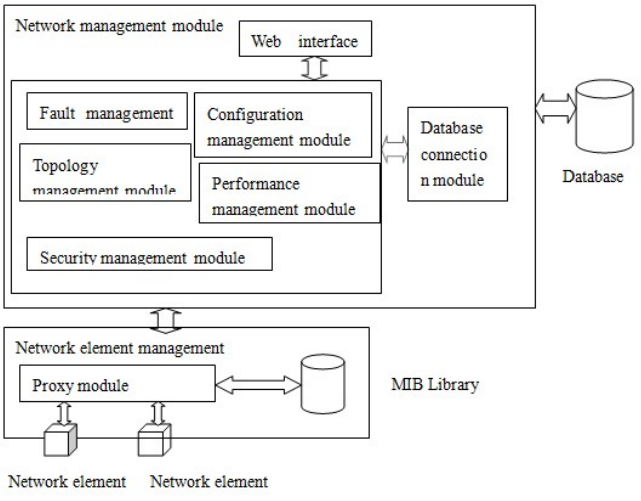


Fig. 1. IMS network management system architecture diagram

4 IMS Network Management Information Model Design

IMS network management system adopts SNMP management framework for network management model. The main content is the design of the MIB. The MIB is designed as early as this relates to the entire IMS network MIB database design and management of various network element, which is associated with scene-specific design of the MIB. IMS network in the design of the MIB of the time takes into account our IMS network management functions focused on the underlying network management and Network Element management. We adopted a bottom-up approach to design.

IMS network in the design of management information is a bottom-up approach. First, design a good feature of these entities the MIB. IETF in this area has been defined for us. A lot of technology associates with the scene-related information models[4], such as MIB-II, SIP MIB library, Diameter MIB library as a reference so we can define a scene related proprietary about MIB library. Then these proprietary MIB libraries expand the formation of the IMS network as a whole MIB library.

IMS network management information tree (MIT) is usually between the managed objects contained in the relationship, including through the layers of management information to form a tree as shown in Fig.2.

Fig.2 shows that any network management parameters can only be determined by the above management information tree. Root of the tree is the subnet number. The next layer of the device is type number. Then the lower number is the device. The node is the last device managed within the parameters of the OID number. Subnet number, device type, number and device number are composed of network management device number parameter, while the network can be defined in the

network and device parameters managed within the OID number compared with proprietary MIB library scene related definitions. This parameter is managed by network device number and device parameters managed with the parameters of OID number.

In the implementations we configure the network devices in the database ID number, the configuration management parameters table and the OID configuration table. The network management and network management agent symbol MIB variables have the mapping conversions. The network device configuration table contains the ID number of the IP address and other information, which is the configuration management parameters form the OID of the management parameters to achieve the symbols in the network management and managed the conversion of the OID parameters.

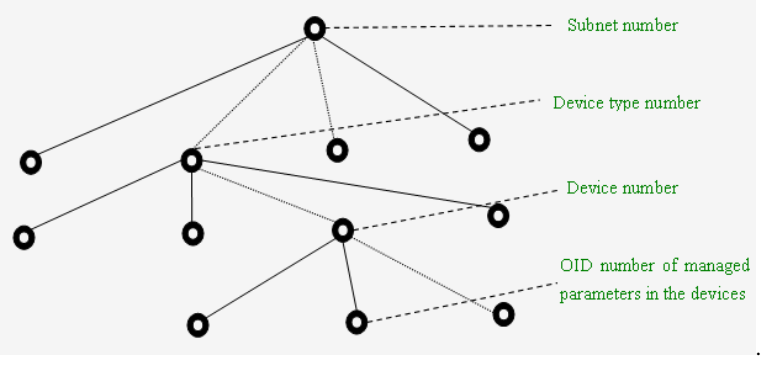


Fig. 2. Tree structure of management information

5 Test Environment

We use Open IMS Core [5] as a IMS simulation network. Open IMS Core is proposed by the IMS core network-oriented open source project on November 16, 2006 in the second International FOKUS IMS Workshop proposed for the IMS core network open source project, which achieves the IMS core Network Element defined in the various entities, such as the CSCF and HSS functional entity functions. The interfaces between the entities follow the R6 3GPP IMS specifications.

Open IMS Core is the 3GPP specification of the specific applications, including two aspects: IMS CSCFS and HSS, CSCFS written by C programming language, which is based on open source projects SER. HSS function (FHoSS) is JAVA-based implementation, and uses a number of other open source projects and frameworks. Open IMS database uses MySQL for data storage. The entire platform is Linux.

Open IMS Core of CSCFs is an open source SIP server, the C programming language with a modular structure can extend the functionality. CSCFs can be used as a SIP registrar, proxy or redirect server with a capacity of thousands per second. It has been successfully run on many Unix-like systems, such as Linux, BSD, and Solaris platforms. In achieving Open IMS Core , each open IMS core sessions control

function entity implemented as dynamically loadable modules. Adding the necessary operation can depend on the 3GPP technical specifications. These modules can also support multithreading. It can maintain state information added.

6 Functional Test

IMS network management system have functional and test. The main functionality of IMS network is verification. The following is described in tabular form.

Performance management function test: S-CSCF network element memory usage parameter query test is as shown in Table 1.

Table 1. S-CSCF network element CPU usage parameter query test

Test case name	S-CSCF network element memory usage parameter query test
Pre-conditions	IMS network management system has successfully been landed. Operators have permission to view the performance parameters.
Steps	In the network management system main interface, click the Performance Management module into the performance management module web pages. In the performance management page, select the "Performance View" function. Select Query S-CSCF network element. Select the CPU usage query parameters. Click OK to submit the query.
Expected output	Enter the performance management module web. Network element query results column shows the query parameter value.

In actual test, S-CSCF network element uses the rate of five percent CPU.

Performance monitoring tasks create a test: Monitoring S-CSCF network element control interface traffic tasks is as shown in Table 2

Table 2. Performance monitoring task creates a test

Test case name	S-CSCF network element memory usage parameter query test
Pre-conditions	IMS network management system has successfully been landed. Operators have permission to view the performance parameters.
Steps	In the network management system main interface, click the Performance Management module into the performance management module web pages. In the performance management page, select the "Task configuration" feature. In the task configuration page, select "New Task" feature. A Web page is set up by task parameters. click OK to submit the new task.
Expected output	Enter the performance management module web. Enter the task configuration page. Enter the new task page. The new S-CSCF network element interface traffic monitoring tasks appear in the Task Manager bar.

Performance Monitoring Performance monitoring task starts and sees the test: the S-CSCF network element interface traffic monitoring tasks are as shown in Table 3.

Table 3. Performance Monitoring mission creates a test

Test case name	S-CSCF network element memory usage parameter query test
Pre-conditions	MS network management system has successfully been landed. Operators have permission to view the performance parameters
Steps	In the network management system main interface, click the Performance Management module into the performance management module web pages. In the performance management page, select the "Task configuration" feature. In the task configuration page, select "Task Manager" function. The selected S-CSCF in the web network element interface traffic monitoring tasks is clicked start in the action bar. Click the monitor after an hour in the performance view.
Expected output	Enter the performance management module web pages. Enter the task configuration page. Web pages appear in the Task Manager on the S-CSCF network traffic charts.

7 Conclusion

IMS core network has been widespread concern in the industry as the future direction of development. The IMS network provides a network management system, and a convenient access to future operations support systems, which will greatly facilitate the operators by deploying the IMS network with good practical sense. The corresponding realization of the program have been achieved by learning the in-depth knowledge of IMS networks, network management standards and trends as well as software development technology. The test results are consistent with our expected output, which verifies the feasibility of the system.

References

[1] 3GPP, TS23.228, V8.10.0, 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects, IP Multimedia Subsystem (IMS): Stage2[S] (September 2009)

[2] Camarillo, G.: The 3G IP Multimedia Subsystem(IMS) Merging the Internet and the Cellular Worlds, 2nd edn. (May 2006)

[3] 3GPP, TS32.101, V8.4.0, Telecommunication management, Principles and high level requirements (2009)

[4] Hwang, J.-H., Cho, J.-H., Lee, J.-O.: A Routing Management among CSCFs Using Management Technology. In: Hong, C.S., Tonouchi, T., Ma, Y., Chao, C.-S. (eds.) APNOMS 2009. LNCS, vol. 5787, pp. 502–506. Springer, Heidelberg (2009)

[5] Blum, N., Magedanz, T., Schreiner, F.: Sebastian wahle From IMS Management to SOA Based NGN Management. Netw. Syst. Manage. (2009), doi:10.1007/s10922.009.9118.4

Requirements of the Times – The Collaboration of E-Government and E-Commerce

Xiao Shufeng

Qinghai University, China

Abstract. With the development of informational technology, e-commerce and e-government bring a lot of convenience to people's lives, more and more attention has been paid, e-commerce and e-government's collaboration is the development of the times. In this paper, based on the concept of e-commerce and e-government, specifically addressed the need for synergies between the two, and its significance.

Keywords: e-government, e-commerce, collaboration.

1 Introduction

Innovation comes from the demand, for each management concept and management model should be based on the social level of technology. If the industrial society promotes the development of the organization by telephone, television, newspapers and the leading information and communication technologies, then, the information society is the leading force to promote organizational progress ,no doubt that the computer network is as the core . Development of informational technology brings new content and challenges to the internal operation and departments' construction of the government and enterprises,therefore, the implementation of e-government and e-commerce will be welcomed. So far, the value of e-government and e-commerce applications for the community is far from the 1 +1> 2 synergistic effect, in other words, the pace of development of e-government and e-commerce is not balanced,neither achieve integration. Now I will interpret e-government and e-commerce's "harmonious development", for further argument, to inspire hope that China's "harmonious society" construction. First of all, the concept of e-government and e-commerce.E-government is the government agency use of technology of modern network communication, with internal and external management and service functions for seamless integration in the government agencies to streamline the work process optimization, integration of government resources, after the reorganization of government departments, government web sites a large number of frequent administration and day-to-day affairs in accordance with the procedures set implemented on the Internet, thus breaking the constraints of time, space, and the separation of departments, a full range of social and one of the standard, efficient, high-quality, transparent, in line with international management standardized management and services . It includes two aspects: First, the government office of internal departments, the information flow between the various affairs with various

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 819–824.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

organizations, including organizations within the flow of funds, the transaction flow and logistics delivery, in order to achieve integration and sharing of resources such as online tax, online social security, Medicare, and so on. In addition,external provision of public services. First of all, within the range of government services, e-government should be fully covered; Second, the public should be given timely and effective response.

2 Experience Based Approach

E-commerce is all kinds entities of commercial activity and demand entities (manufacturers, commercial enterprises, financial enterprises, government agencies, individuals) .To overcome the constraints of time, improve the effect of the business community activities, use of computer network and a variety of digital media technologies and electronic means such as a trade commodity trading and transaction services in the form . In general, e-commerce include the following five aspects: First of all, use of high technology, including the main WAN, LAN, etc.; Second, commodity trading, services, transactions, transaction information, including not only physical and material; third,it is not only between enterprises and customers business activities, but also enterprises and between enterprises is in business activities; Fourth, not only enterprises and external business activities, as well as a full range of internal value chain management; fifth is to enhance the economic effect for the purpose. Of particular note is that today's era of e-commerce is a full range of e-commerce, uniform standards from internal to external resources, communication barriers all cleared, certain aspects of the local electronic extended to various parts of the organizational activities ,there are a lot of business opportunities for enterprises to create and improve operational efficiency.

3 Designing Experiments

Collaborative Commerce (CC) and synergy of the Chief (CG). Related businesses should use IT to interact together in e-commerce, this is the Collaborative Commerce. Businesses or individuals for the many things deal with various government departments, traditional government is to make the enterprise or individual serial to run errands, the government can now use IT to carry out the coordination office ,which will not only improve government efficiency, but also can reduce the overhead of businesses or individuals, which is collaborative government.

4 The Collaboration of E-Government and E-Commerce

The collaboration of e-government and e-commerce is the use of computer and communication technology, government and business integrate in the same database, making the datas’communication timely and effective, information sharing, so as to further improve information technology management and service of the Government during the enterprise developing in many ways to achieve the "harmonious"

development of e-government and e-commerce, the leap in the development of the national economy.

The effect from the macro as follows:

(1) changed the traditional management of the Government to improve the administrative efficiency and level of synergy;
(2) e-government and e-commerce, information sharing, help create new economic growth points, to achieve the government and enterprises dual income.

Performance from the micro:

(1) greatly reducing the communication cost of the traditional economic approach, the government and enterprises, in line with the timeliness of the information, to achieve full utilization of resources and optimal allocation.
(2) improves the visibility of the Government to monitor the work of enterprises can also take advantage of this interactive platform, and mining have their own public information resources, thus creating more value for the enterprise.

Second,the needs of the community to promote e-government and e-commerce collaboration

5 Diminshing Marginal Cost Peinciple

The marginal cost refers to the increase in costs for each additional unit of output. Traditional government and the general business situation to follow the law of increasing marginal costs. According to Simon the limited nature of the doctrine. In industrial society, the number of individuals under the direct management, between 7-13 people, more than you need to stratification, with the economic development and business tasks more weight, more level, the greater the scope of management, corresponding to the higher management costs. In the information society, the rapid development of the network is the development of the traditional pyramid of bureaucracy to the flat, decreasing the marginal cost of government and enterprises, that is, the larger the scale, the yield is more, the increased costs of the increase in the number of unit product gradually reduce, or even 0. For example, after the completion of the electronic social security system, social security department of five enterprises and 500 enterprises to pay the total cost is almost the same, so each additional business, additional investment required will be low. In addition, this synergy is not only to achieve the two sides of the diminishing marginal costs also help value-added re-creation, to improve government management and service level, and the ability of enterprises to collect information.

6 The Government Needs

First, the main functions of government is reflected in the management and services, and enterprise management and service is its indispensable part in the information

technology rapid development of society, the public can not help but ask the government on corporate governance and services embodied in what areas? the need for government and enterprise information systems integration is Visible.

Second, followed by government departments as a physical unit, the inevitable requirement for the procurement of goods the supply of materials to meet the normal operation of government departments, such as office equipment, fuel, electricity, etc., and material procurement and management, in addition, if we put government documents circulation as a special "products", then some of the ways and means of e-commerce can be adapted to apply to the management of government business processes, the overall composition of the government's supply chain management.

Third, the government in order to improve the efficiency and level of social services to the traditional government administration and services, business process re-combination and renovation, the selective use of various methods of e-business process reengineering, and the information society and market economic development.

Fourth, one of the important functions of government is to manage and deliver public services. The Government should establish a good image of the government to provide the capacity and efficiency of public services and e-commerce customer relationship management methods, which can be used to implement improved.

7 The Enterprise Needs

E-commerce transactions double the level of credit is difficult to judge the characteristics of virtual features, integration, coordination and openness, the signing of the contract of e-commerce part of the network is not an effective solution to this problem, which requires the e-government role in the protection of both the businesses or individuals to provide true and reliable information, reduce transaction risks, and at the same time, companies are e-government platform, extends the enterprise value.

First, the enterprises in order to enhance market competitiveness, optimize the allocation of requirements on the enterprise's internal and external resources available to speed up the response to the market, provided in the e-government policy is an integral part of enterprises to make decisions.

Second, the market in order to attract customers, retain customers, and continuously upgrade the customer to provide personalized service, it is necessary to use the relationship between the IT business and enterprise customers to maintain, manage and update customer relationship management, the Government's demand is the enterprises should grasp the important information.

Third, the enterprises in order to meet customer needs, to unite enterprises upstream and corporate partners of the downstream use of information technology management, procurement of raw materials, fuel, etc., as well as product availability and use of e-government, the government will be involved in their supply chains.

Fourth, the interactive information exchange for convenience with the outside world, for quick access to outside information and to improve the efficiency of the enterprises to establish enterprise information portal, which is the enterprise information portal, among the Government of the enterprise The evaluation is the focus of attention in the enterprise and the public.

8 From the Point of View of the Public

As is economic globalization today, the self-employed and entrepreneurial phenomenon is common, but such problems usually have to go through some complex process, the public are to blame. The person e-commerce and e-government collaboration will effectively solve this problem, individuals do not have to go so multisectoral also bring convenience to business and government to achieve a "win-win".

The emergence of the Internet for e-commerce technology base, the majority of enterprises on the Internet hold great hope, take full advantage, to create substantial value for the enterprise, in the face of e-commerce development, the Government must react accordingly. In fact, e-commerce is a driving force behind one of the e-government. It can be seen, e-commerce and e-government will be welcomed.

Third, how achieve e-commerce and e-government collaboration

From the application level, government and business are two entirely different organization, they have a completely different application of target groups, government administration and enterprise management are essentially different, a public affairs management, commercial management. The abstract system level, the two there is a sense of commonality, management and service functions, with this common, theoretically possible for ERP and other enterprise management software to the translation of e-government, which the forefront of the latest trends trends in international e-government.

1 the government supply chain management, is about the procurement of government supplies, the publication of documents using e-commerce enterprise value chain management.
2 independent of government websites, the use of e-commerce, enterprise portal concept to establish the separate page of the Government, to provide true and reliable information for individuals and organizations, "one-stop" service.
3 the government and enterprises sharing database system, the Government has effective control companies better understand the state regulations.

E-government, e-commerce is the product of the development of information technology to promote the rapid development of the national economy, at the same time, collaborative e-government and e-commerce requirements of the development of globalization is the future direction of development of information technology, the only advantage of this opportunity government and enterprises in order to achieve a new leap in social development to a new climax.

References

[1] Jun, J.: Disciple of the Chief of the Introduction to. Peking University Press, Beijing (2003)
[2] Zhang, R.: e-commerce Introduction. Electronic Industry Press, Beijing (2002)
[3] Chen, R., Chun, F., Hsu, L.-P.: Microeconomics. Donghua University Press, Shanghai
[4] Lou policy group on e-government and e-commerce. Central China Normal University (3) (2002)
[5] Shaohua: e-government and e-commerce. Northern Economy and Trade (10) (2003)
[6] Chen, T., Zhang, C.: Collaborative e-government and e-commerce. Business Studies (15) (2005)
[7] Luo, C.: e-government and e-commerce. Business Economy (6) (2003)
[8] Liu, Y.: On the e-government and e-commerce, collaborative development. Yili Normal University (Natural Science Edition) (3) (2009

Study on the Development Path of China Modern Coal Logistics Based on Electronic Commerce

Jiang Lingli and Cen Kuang

China University of Geosciences, Beijing 100083

Abstracts. Traditional logistics and modern logistics are two types in logistics management, which is put forward by Chinese scholars. Modern logistics of coal, compared with traditional logistics of coal, is a type of new integrated management. In the era of prevalent e-commerce, the combination of modern coal logistics and e-commerce promotes the development of coal logistics. Based on e-commerce, this paper will explore the development path of China modern coal logistics.

Keywords: modern coal logistics, electronic commerce, development path.

1 Definition of Modern Coal Logistics

Coal logistics is a small branch of the logistics industry. Ma Qian Jie(1991) is the first person who introduced 'logistics' into coal industry and researched how to manage logistics in coal enterprises from the perspective of logistics, production and management of coal enterprises. To summarize the research from other mainstream coal logistics theory research scholars, it comes to consensus that coal logistics is a systematic logistics. It exists in the whole process including development preparation, production process and marketing activities of coal products. It is a dynamic transfer process of a series of material delivery in producing coal products and managing coal marketing etc.. It is a logistics system consists of supply logistics, production logistics, distribution logistics, recycling logistics and waste logistics of coal (Wang Dezhan (2005), Tang Xifeng (2007), Wen Jun Zhang (2008), An Jifei(2010)). From the definition, it can be seen that 'coal logistics' is defined taking the physical state of logistics flow into account, it does not involve the systematization and optimization of other activities like storage and transport as a whole and is the simplest displacement process. The coal logistics in this definition is only in the early stage of traditional logistics. As a supporting industry in market economy, logistics should be more focus on service, whose tenet should be more efficient, more convenient and more time-saving in the flow of state, which along with transport rationalization, warehousing automation and standardization, loading and uploading mechanization, process and distribution integration, information management networking, etc. According to this concept, this paper redefine the coal logistics: coal logistics is, to meet customers' demand, the whole moving process of coal products from the coal mining in colliery

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 825–830.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

to the hands of clients. During this period, it includes the activities like processing, warehousing, distribution and transport of coal products together with related information collection and processing. It is a strategy measure considering all market conditions like manufacturing, transportation and sales. Its aim is to achieve transport rationalization, warehousing automation and standardization, loading and uploading mechanization, process and distribution integration, information management networking. The coal logistics in this definition is modern coal logistics relative to traditional logistics.

2 The Development Course and Current Situation of Coal Logistics

The concept of "Logistics" was passed from Japan to China in the 1980s. Based on the management characteristics, Chinese scholars put forward two forms: "traditional logistics" and "modern logistics". There is no such concept of traditional logistics and modern logistics abroad for it is endemic to China economic system. Traditional logistics in China is established in the era of planned economy and it adapts to the good storage and transportation in that time characterized by ' little varieties, high-volume, few batch, long cycle', under which circumstance, it develops with simplex function, low quality of service, service passivization. Compared with traditional logistics, modern logistics is a revolutionary breakthrough. It organizes the relatively independent logistics nodes under traditional transport scientific and rationally, so as to provide customers the best transport route, the shortest transport time, the highest transport efficiency, the safest transport security, the lowest transport costs, and is an integrated management model. At the same time, it broke through the capacity-centric perspective in transport service and stressed that the purpose of transport service is customer first -- to design the content and form of transport service according to clients' demands, and to develop professional and personalized service. Compared with traditional logistics, modern logistics is featured with logistics of rapid reaction, logistics function integrated, logistics service serialization, logistics operation standardization, logistics objective systematization, logistics management marketization, logistics instrument modernization and logistics organization network.

To analyze coal logistics development in China from the dual-view perspective of "traditional logistics" and "modern logistics", we are still at the stage of traditional logistics for the present. Dr. Li Min(2009), as a practitioner working in coal industry for more than 20 years, mentioned that the long-standing outdated concept of “emphasis on production, neglect on circulation” has caused slow development of coal circulation in China, which summarized the restriction on coal logistics development in China in the aspect of operation philosophy. In addition to the constraints of the institutional and operational concepts, lag in technology is also the major reason for slow development in coal logistics. With the rapid development of computer networks, communication technology, especially the popularity of internet, a new business mode springs up, via which to realize the coal logistics electronic processing is a prerequisite for modern coal logistics. In the context of modern

logistics developing in full swing in other industries, China's coal logistics is also working on the concept for reform. There have been very few enterprises like TADER COAL SCM, Kailuan International Logistics Co., Ltd. working on practitional operation to realize modern coal logistics employing the concept of coal supply chain management. Therefore, China's coal logistics is still at the stage of traditional logistics for now, but it has begun to reform and transform to modern coal logistics.

3 Promoting Role of E-Commerce on the Coal Logistics

By using computer technology, network technology and telecommunication technology, e-commerce is to realize electronic, digital and network business during the entire process. E-commerce includes several major flows: information flow, business flow, capital flow and material flow, among which material flow is the foundation, information flow is the bridge, business flow is carrier and capital flow is the goal. The emergence of e-commerce provides new strategy timely for logistics industry, helps enterprises to achieve the system integration of logistics management and to create a higher profit value for companies in market competition. E-commerce accelerates the development logistics management in below aspects:

3.1 The Impact of E-Commerce on Transaction Process in Coal Logistics

E-commerce featured with full 24-hour continuous operation mode, sufficient options and automatic transactions helps increase the transaction efficiency and trading opportunities. Via network platform, all kinds of purchase and sales information are gathered , the supply and demand information is shared among all the participants, any dealers could choose the most suitable products and services, options for buyers and sellers increase. At the same time, the transaction can be completed directly through the network platform, which reduces the trade link, shortens the transaction cycle, thereby improves transaction efficiency.

3.2 The Impact of E-Commerce on Distribution Process in Coal Logistics

E-commerce shortened delivery time. In traditional logistics distribution management, the delivery time is very long due to restriction of information exchange. With the intervention of network system, this time amount will become relatively short, information and resource about distribution will be spread to relevant departments by network management.

E-commerce optimized delivery process. Traditional logistics delivery process is composed of many business processes. Due to the lack of IT support, information communication is poor and contact between various links is loose. Logistics delivery process supported by e-commerce is connected by a network. When any terminal received a request, the system will respond within a very short time and be able to design a detailed distribution plan to notify the links to work. That's to say, logistics

delivery business using e-commerce can achieve real-time monitor and real-time decision-making in the whole process.

3.3 E-Commerce Promotes the Development of Coal Logistics Informatization

Management systems, operational processes, information feedback system all requiring informatization, electronization and integration, it's a must to establish a network of logistics information system. To set up good information processing and transmission systems in intranet and extranet is to make clients, consignees and each warehouse, transport company ready to receive goods, to make cargos move fast without delay and to reach customers' specified destination directly. Large distribution companies usually establish the ECR (Efficient Customer Response) system and JIT(Just In Time) system. The establishment of various information systems helps promote the development of coal logistics informatization construction.

4 Development Path of the Coal Logistics in E-Commerce Mode

4.1 Large-Scale Development of Coal Logistics

Economies of scale refer to economic phenomena that with the increase in input, the increase in output over input ratio. Mergers and reorganization, expanding the scale of coal logistics enterprises and increasing the concentration of coal distribution industry are important measures to achieve large-scale, intensive development of coal logistics, among which building strategic alliance in coal logistics is major strategic approach in large-scale operation. Specific form of large-scale operation is Coal Logistics Park. During coal logistics operation, there are several links like procurement, roughing, warehousing, distribution and transportation. By mass procurement, efficient use of storage equipment to achieve joint distribution centralized processing etc. to reduce the average cost and gain economies of scale.

4.2 Specialization Development in Coal Logistics

In modern economic environment of prosperous outsourcing business, coal logistics industry which always dominated by vertical integration model should size up the situation, separate logistics business from production business, allow manufacturers to focus on production and logistics enterprises to focus on distribution, carrying out their duties through specialization. Among them, the third-party logistics is the main pattern of specialization development. The third-part logistics refers to a kind of logistics operation and management that in order to concentrate on primary industry, production and management enterprises entrust the logistics part which processed by themselves to professional logistics service companies by contract, and keep close

contact with logistics enterprises through information systems to control the whole logistics operation.

4.3 Informatization Development in Coal Logistics

Information technology is a necessary condition for the development of modern logistics, also a key support for modern logistics service system. Modern logistics is informationized traditional logistics. The development of coal logistics should be directed by information technology, actively promote the application of the Internet of Things, accelerate the construction of coal circulation informatization, inventory minimized, links fewest, cost lowest and logistics of low-carbon. In practical application more common use of information software includes supply chain management software(SCM), enterprise resource planning software(ERP), client relationship management software(CRM), and other enterprise information management software, logistics and transport management software(TMS), Global Positioning System(GPS / GIS), intelligent transportation system(ITS), and other logistics and transport positioning system softwares, as well as radio frequency identification (RFID) smart recognition system etc.

4.4 Integrated Development in Coal Logistics

Coal logistics integration refers that in the process of coal logistics, breaking industry boundaries, regional boundaries and corporate boundaries, through mutual coordination and harmonization, to create the most appropriate structure of the logistics operation to improve the operational efficiency of the whole logistics process and benefits of business strategy. Implementation of the integration strategy has three main dimensions: internal collaboration, cooperation between the enterprises and the external environment, as well as the unity between the internal and external cooperation. The main form of integration operation mode is supply chain management.

References

[1] Wang, D., He, S.: How to Develop the Coal Logistics. Logistics Technology (03), 49–51 (2006)
[2] Tang, X., Yu, J.: Realizing the Transformation From Contaminative Coal Logistics to Environmental Coal Logistics. Mining Research and Development (1) (2005)
[3] Zhang, W., Xie, J.: Discussion the Coal Logistics in China. China Storage & Transport (4) (2008)
[4] An, J., Pei, H.: Analysis of Shanxi Coal Logistics. Friend of Science Amateurs (12) (2010)
[5] Li, M.: Research on the frame system of coal supply chain management. China University of Geosciences (2010)

[6] Liu, L., Guan, X.: On the Development Modes of Coal Logistics Industry in China. Journal of Shandong Institute of Business and Technology (February 2009)
[7] Chi, Y., Zhang, J.: The Relationship between Traditional Logistics Industry and Modern Logistics Industry and the Transformation from One to the Other. Journal of Chongqing Jiaotong University (Social Sciences Edition) 3 (2003)
[8] Jiang, L.: Thinking about the development of coal logistics to the modern logistics. Coal Economic Research 7 (2010)

Research on Flash Based Online Game

Liu Jinjun

Yangjiang Vocational and Technical College
Department of Computer and Science

Abstract. Online game is becoming more and more popular according to development of internet, especially the high development of mobile internet. Flash web game is one of the most effective way to design web game and it's has become more and more complex and heavy loading. Since all resources of Flash web games are loaded from web server, it would cost a long time if the resource is very large, such as game map or some complex graphics pictures. It's not user interact friendly and the conventional way can't prevent and handle attackers effectively. Therefore, this paper researched on Flash web game system and designed routing algorithm, resource loading strategy and security system. After stimulate with conventional wen game system, the result shows this paper designed Flash web game system can work well with a lower loading time and higher security mechanism.

Keywords: Web Game, Flash, Resource Loading, Game Security.

1 Introduction

According to high development of internet and web technology, Web game is growing faster and faster, which market scale has already exceeded 20 billion in China. The development of online game is the process that multiple user domains transfer to client-side online game. As the descendant of multiple user domains, web game becomes the focus of attention again after a long hard time [1]. According to the creation and development of game production, design idea of game continuously is improved these years. The purpose of game is to let users relax and feel happy in the experience, so relaxed game experience is the fundamental goal of web game designation. Web game greatly deduces the limit of operation environment and makes users conveniently enter the game at any time. Flash web game occurs as a design idea of clientless online game in recent years which combines traditional game development technology with Flash image processing technology. It is the trend of web games development, so study Flash web game technology deeply has a realistic value for development of game production in china [2].

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 831–836.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Flash Based Online Game Synopsis

2.1 Current Situation of Online Game

Traditional online game needs users to download huge client-side and demands a high configuration computer to support, while web game just need users to open website browser and input game address without any client-side downloading and machine configuration problem. Lots of people worry about addict of online game all the time, because game users, especially young people, usually lose their control of game and then trigger the conflict between virtual world and real world. Besides, many people misunderstand online game under the influence of negative reports from some media [3]. A survey shows that the important factor of making users leave away online game in china is objection from family members and friends. For the above situation, web game is more easily accepted by people than other game without conflicted emotions, because its content and way is simple. For web game, multiplatform development comes true easily. The direction of web game development is combining wireless application protocol (WAP) in mobile phone with mobile phone client graphics online game which is cross-platform, not only stay in web performance forms. In other words, a web game can played not only be in website, but also in 3G mobile phone, because game data is shared [4].

2.2 Flash Instance of Online Game

The biggest difference between Flash web game and traditional online game is the organization and uploading ways of game resources. Flash web game dynamic uploads all game resources from the network resources server through website browser after game is beginning, so some big game resources, such as game map needs a long uploading time that makes users wait for several seconds. It is bed for user experience despite normal work. In order to solve this weakness, the gradual game resources loading system is designed in Flash web game [5]. Another weakness of Flash web game is safety, because it operates based on website browser and its file layout is open source. In traditional online game, some operations such as plug-in protection and special encryption format keep its safety, but they can't be applied in Flash web game. Thus, safety technology design according characters of Flash is an important work

3 Research on Flash Based Online Game

3.1 System Architecture of Online Game

The system Architecture of this paper designed online game is based on Flash design and advantages of conventional online game. We added a set of server for resource downloading, which is quite different with conventional online game design. The system can load game resources which are needed for users to launch, according to downloading request from remote place. The whole system architecture, which is designed according to characters of this kind of game, is showed in figure 1.

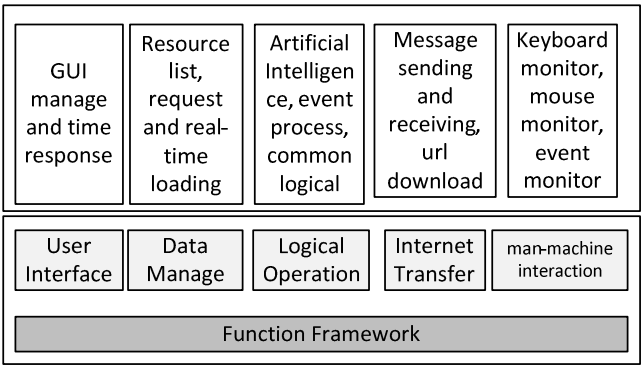


Fig. 1. System architecture of Flash web game system

All sub modules showed in figure 1 consisted a whole game framework and realized the client side design idea into online game research and design process. Modules communicate with each other by interfaces. This paper designed online game without client resource is consisted of online game function system and game server side resources.

3.2 Routing Algorithm Design

φ_{ab} is used to express data of routs, such as node a to node b, during rout searching process, it follow probability distribution as follows:

$$S_{ab} = \frac{\varphi_{ab}}{\sum_{b \in B} \varphi_{ab}}, \quad \forall b \in B \tag{1}$$

If there is no node available for routing, the system will allocate the task randomly. We set the maximum and minimum limited of jump step. During the update process of jump step, routing algorithm includes two steps: routing searching and deduction searching. $y(x)$ is expressed as feasible solution. If $y(x) = 0$, the solution is feasible, if it's not zero, and the solution is not feasible.

$$y(x) = \sum_{b \in B} \max(0, \sum_{a \in A} L_{ab} \bullet X_{ab} - L_b) \tag{2}$$

Routing searching algorithm is followed the below step:

$$R(x) = \left\{ (a, b, a_1, b_1, a_2) \middle| \begin{matrix} x_{ab} = 1, x_{a_1 b_1} = 1, s_b > L_b \\ s_{b_1} > L_{b_1}, a \neq a_1, b \neq b_1 \end{matrix} \right\} \tag{3}$$

In this formula, $s_b = \sum_{a_2 \in A} x_{a_2 b} \bullet L_a$, and $a, a_2 \in A$ as well as $b, b_1 \in B$。

The deduction algorithm is expressed as:

$$D(x)=\left\{(a,b,a_1,b_1,b_2)\middle|\begin{matrix}x_{ab}=1,x_{a_1b_1}=1,y_{ab}=0\\ y_{ab}=1,a\neq a_1,b\neq b_1\end{matrix}\right\} \tag{7}$$

In this formula, $a,a_2 \in A,b,b_1 \in B$.

3.3 Resource Loading Strategy

One of the most different solutions between Flash web games and conventional game is the game resource management and loading methods. Since all resources of Flash web games are loaded from web server, it would cost a long time if the resource is very large, such as game map or some complex graphics pictures. It's not user interact friendly. In order to save loading time of Flash web game, this paper designed a new resource loading system called incremental resource loading system.

Incremental resource loading system support dynamic display the web game content during it is downloading. It loads all images data of web game in the form of binary streams. For example, when loading a map in Flash web game, it takes incremental method, it display map and images at the same time when they are loading. It will display a interleaved fuzziness images at the first time and then it will become more and more clear during more and more resources are downloaded. Because this type of loading method, users did not need to wait for all resources download and can interact with web game first, such as operating event and moving. The game scene will become clearer according to more and more resources are loaded. It would display fuzzy images at the first time and then display clear images when all resources downloaded. Source codes of main part are showed as below:

```
var tmp: Bit;
        var tempdata:BitData;
        for(var y:int = 0; y < Puzzle.rows; y++) {
            imageArray[y] = new Array();
            for(var x:int = 0; x < Puzzle.cols; x++) {
                tmpd = new BitmapData(piW, piH, true, 0x0);
                tmpd.copyPixels ( image.bitmapData, rect);
                tmp = new Bitmap( tmpd );
                imageArray[y][x] = tmp;
            }
}
```

Web game design needs to take security in high important situation in order to keep game operate safe and the fairness among different players. Since Flash web

game is running in web browser, and it's file is open, the security of Flash web game is weaker than conventional large scale online game. Plugin protection and encryption are common methods used by conventional large scale online game, which can't be used on Flash web game; the whole security system is consisted of region detection, obfuscator, compressed as EXE file, password protection, multiple loading SWF and security engine. Compressed as EXE file would make it hard for attacker to break source codes. Obfuscator can obfuscate source codes, even when they are broken; it's hard to understand the logical solutions. Multiple loading means lunch a Flash file and then call another main game from this Flash file, which can highly protect the second main Flash file. Password protection can also protect the Flash web game from attackers. These entire modules combine together by security engine and can protect Flash web game safety much more.

4 Stimulate Experiment and Validation

In order validate the efficacy and ability of this paper designed Flash based web game system; we designed a set of experiment to test it in two aspects, one is the resource loading system and the other is security system. The hardware environment is on a desktop personal computer which installed windows operate system and internet explorer browser. After configure the environment successfully, we test the resource loading system by different size of resources, form 500 KB to 5MB. The block time from Flash web game system lunch to user can interact with interface is showed in figure2.

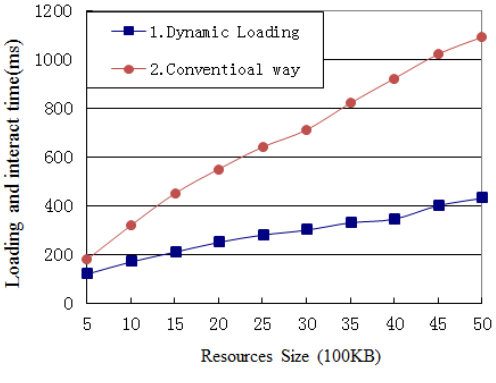


Fig. 2. Loading and interact times of Flash web game system

We also test the security engine of this paper designed Flash web game system. Standard Flash game without security engine is picked as a contract of this experiment. Attackers are randomly picked by attack system and they are executed by script. The result is showed in figure 3.

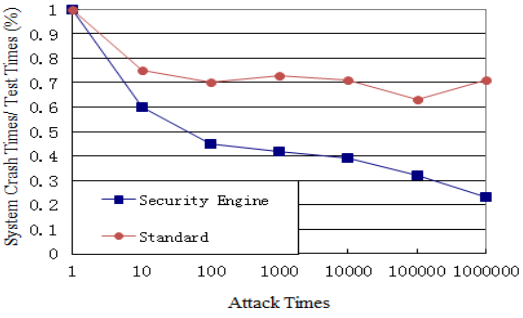


Fig. 3. Crash times of Flash web game system by attackers

5 Conclusion

According to the high development of internet technology, it's more and more popular to play game online. Flash web game is more and more important in daily life and has become a common way to for people relax themself. But conventional Flash web game system can't load game resource dynamically and can't prevent attackers effectively. Therefore, these papers research on a new Flash web game system, designed its routing algorithm, resource loading strategy and security system. After experiment in reality environment, it shows this paper designed Flash web game system works well with high accuracy. The result gave a reference for researching on Flash web game system.

References

[1] Wang, H., Tan, S.: Flash Embedded Linux Device Drivers in the Study and Optimization. International Journal of Web and Game Research, 14–17 (October 2008)
[2] Liu, Y.: The Design and Development of Flash Game-based Physics Courseware. In: 26th Annual Conference of the Society for Electro-Acoustic Music in the United States, Miami (2010)
[3] Johnson, D.W., Jankun-Kelly, T.J.: A scalability study of web-native information visualization. In: Proceedings of Graphics Interface, GI 2008, pp. 4–20 (2008)
[4] Feigin, E.: A Case for Automatic Run-Time Code Optimization. Senior Thesis, Harvard College, Div. of Eng. and Applied Sciences (April 2004)
[5] Schoppa, C.: Browsers revmp for embedded. Electronic Engineering Times (1129), 96–98 (2000)

Applicability Analysis and Classification in Network System Security

Luo Zhuojun

Hunan Mass Media Vocational Technical College, Changsha, Hunan, 410100

Abstract. This article describes five major functions of the hierarchical network, and depicts in detail access control, encryption to protect the data integrity, firewall technology and its applicability, which is the development trend to network security with compositely use of them.

Keywords: hierarchical network, access control, encryption, firewall.

1 Introduction

Network technology are increasingly extensively and in-depth being applied to fields of the society recently, and deeply changing social behavior and looks. However, at the same time, the security of network system is a stumbling blocks in further popularization, especially financial, national defense and other areas application. Therefore, the international standard organization (ISO) puts forward the hierarchical network security system structure, and defines five service functions: identity authentication, access control, data security, data integrity, confirmation service. The confidence degree of the reliable network depends on its security service quality

2 Access Control and Its Applicability Analysis

Access control also named visit control, which realizes the internal security of the network system to some extent. Security access strategies is a kind of rules, which is used to determine whether the main body have access ability to the subject object. The visit includes three levels as follows:

(1) System level control. Its task is implementing safety management from the whole system in order to prevent the users not permitted accessing to host or network and obtaining system resources and services. So it has various control measures of user registration, user login, system security access.
(2) User level control. It classifies its users in the network system, and distributes them different access permission. There are super users, network operators, average consumers, general customers.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 837–840.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

(3) Resource level control. All kinds of resources have their own visited access permission in the network system, and have none relationship with the users in order to protect the safety. There are directory security set and file security set.

3 Ensure Encryption Strategy and Applicability of Data Integrity

Data integrity has two aspects of meaning, data could not be error, missing and destructed for the objective factors, and not be captured, transferred, read, destroyed for the subjective factors in the transfer and storage.

(1) Data encryption. It is an effective method for the data safety with two different encryption strategies to realize safety in communication, which includes link encryption and end-to-end encryption. And it can be an effective way to not analytic information in the line because of these features, and prevent the attacker realizing business data flow analysis. But it does not apply to broadcast network because its subnet has no clear link. If we encrypt the whole PDU, we should not determine the sender and receiver. Therefore, link encryption in the OSI environment will be much limited unless the other measures, it maybe apply only to data protection within the local scope.

(2) Node encryption. It is improvement to link encryption, and can overcome the defect of plaintext attacked between nodes. It is advantageous to choose transmission router that encrypts the message not the header.When the encrypted data transfer into the encrypted component in the node, it decrypts a key and then encrypt the other key, finally transfer into the node, which solves in certain extent the shortage of the link encryption.

(3) End-to-end encryption. It is more appropriate to end-to-end encryption in the large network when the exchange network transfers data in each sending and receiving because encryption finishes between two terminals or two host hardware equipments, and also finishes in the host software, which is good to prevent copying and software leaking while encryption does not decrypt in any way among nodes. This way is beyond of the communication subnet, which has flexibility to choose the end-to-end encryption level. If encryption is in the transport layer, it will be transparent for the safety measures to users, which is not necessary to provide separate security protection for each user but is vulnerable to attack from the transport layer.When we encrypt in the transport layer, the users can choose different encryption algorithm based on their special requirements and will not affect others, which is more easily suitable for different users' requirements. Therefore, the encryption can apply not only to internet environment, but to radio network. We should combine link encryption and end-to-end encryption to obtain better security, where the former encrypts in PDU destination address, the latter provides the end-to-end data protection.

4 Firewall and Its Applicability

Network data encryption can protect data transferring between two ends, but encryption can't prevent the invasion from three attackers, that is, the saboteurs, the furious pride, thieves. Therefore we must adopt firewall technology to realize encryption for the network security.

Firewall is in the bottom of the network security system, and belongs to its technology category, which is responsible for network security authentication and transmission, and provides all kinds corresponding security services for the network application. It has four types based on its different firewall technology: the packet filter, network address translation, agency and monitoring.

Firewall is a safe guard system between internal and external network, and is considered to be an access control mechanism to determine which internal service allows external access, as well as what external service allows internal access. Also it connects internal network with Internet, that is, it is a component restricted access between the protected network and Internet or between any two network. It has the functions that all information flow allowed by the local security strategy from net inside to outside, or vice verse, must pass firewall, and the system has none influence.

Network firewall has the following characteristics. It can strengthen the entire network security strategy as a network security checkpoint, and can effectively, detailed record users' into/out network activities to take security measures, and can hide the details of local network configuration and activities to strengthen the confidentiality of internal network. But it cannot avoid the attack from insider and virus.

Firewall can be effective to prevent the invasion from outside, and its basic principle is that all not allowed is forbidden. Firstly it blockades all the information flow, and then gradually open to the open expected service based on this principle, which can build up a very safe environment because only the service selected by network administrator allow to use. But it abates user convenience because the users' services opened are limited, so the users can't use more network resources. The other king os firewall allow all not forbidden services, which provides a more flexible environment to more services, but it brings up more burden to network administrators especially in the protected network range increased.

5 Conclusion

Firewall is the first line of defense in network system, and identity authentication mechanism provides a necessary foundation for network security, in a word, it is the trend of current network security development that we implement the union end-to-end data flow encryption technology in the network layer with the advanced authentication technology application, and combines the current firewall technology to make the necessary content detection and some other means.

References

[1] Chen, Z.: Network security. Tsinghua University Press, Beijing (2011)
[2] Heggy, T.G., Liu, D.: He network security technologies and solutions. People post and Telecommunications Press (2010)
[3] Liu, J., Li, A., Song, C.: Active Firewall Designc Based on Safety Psychology. China Safety Science Journal, 72–75 (January 2005)
[4] Mo, S.: Network counterwork and information security. In: Fifth China Communication Association Annual Meeting, pp. 820–825 (2008)

Design of Multimedia File Security Communication Platform Architecture Based on Web

Liu Yanshu, Yi Can, and Li Zhiyong

Hunan Mass Media Vocational Technical College, Changsha, China, 410100

Abstract. This paper designs a security communication website system platform based on Web for a variety of multimedia files, and the system can make legal users share media files and protect the digital copyright, thus ensure the interests of the content providers, and promote the multimedia quality and fun.

Keywords: multimedia file, security communication, Web site architecture.

1 Introduction

It is a very urgent problem to protect image, audio, video and other multimedia content with the rapid development of multimedia and network technology and their widespread application. Their piracy causes billions of loss each year due to the digital media easily copied. It becomes very important how to guarantee the communication security, and the intellectual property rights of multimedia file providers and the rights of the payment. Therefore it plays a very important role in the construction of a multimedia file security communication web system.

2 Present Analysis

As us known, multimedia files includes images, audioes and video files. More and more web systems provide the multimedia display and share services, including downloading VOD based on P2P. It is necessary to how to prevent the network multimedia file been downloaded at random. Today the main secure multimedia communication technology is mainly Digital Rights Management(abbr. DRM), that is, it prevents digital information to illegally copy, print and distribute hrough encryption.

2.1 The Abroad Development Status

The United States Computer Association has organized annual ACM DRM conference (ACM Workshop on Digital Rights Management) since 2001, which has the architecture of DRM system, tracking and auditing to DRM of digital

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 841–844.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

content usage, business model and security needs of digital content transactions, digital content business transactions, authentication, cipher key management in DRM system, transmission problems of digital rights, digital copyright description etc.. Foreign DRM system includes Microsoft DAs, Adobe Content server etc..

2.2 Domestic Development Status

Domestic DRM system has Fangzheng Apabi digital copyright protection system. And many webs are using the multimedia security communication technology, such as nipic.com, 9sky.com, xunlei.com etc.. Although the multimedia files in network have been applied in some security protection technology, their contents are easy to copy and modify with them transfering more and more, and still emerge in large numbers of piracy and infringement

3 The Key Technologies

The multimedia file security communication platform based on Web mainly researchs multimedia file encryption and identity authentication technology based on the analysis and study of Microsoft 's DRM system, proposes a website security transmission system adapted to a variety of multimedia files, which purpose is to make legal users share multimedia files in the website system, and protect their copyright, and then copyright the rights of fee users and intellectual property rights of the content provider, to further promote the multimedia file safely communicating and sharing in network. Its key technology is as follows:

(1) To research encryption technology on multimedia file for the transmission mode and the uncontrollable problem of the multimedia files, and overcome the limitation of transmission mode,

The system could release multimedia files in whatever platforms, and directly carry on the jurisdiction management in network to improve the controllability.

(2) To research a variety of identity authentication technology and strictlycontrol user identification according to different grade users difficult to identify and control. The system chould protect the interests of the fee users with password check, hardware identifier, smart cards and other identity authentication technology, and combine these methods with the multimedia file encryption to improve its extremely high safety.

(3) To research other auxiliary function modules in the system to make its functions be perfect, and realize the function modules of multimedia file upload, file format conversion.

(4) To combine the above research results to realize the secure multimedia communication system based on Web.

4 Architecture Design

Multimedia file security communication platform based Web should be simple operation principle in use and management, and view different levels of multimedia files according to different user permissions. The user operation process is shown in Figure 1, and the administrator operation is shown in Figure 2.

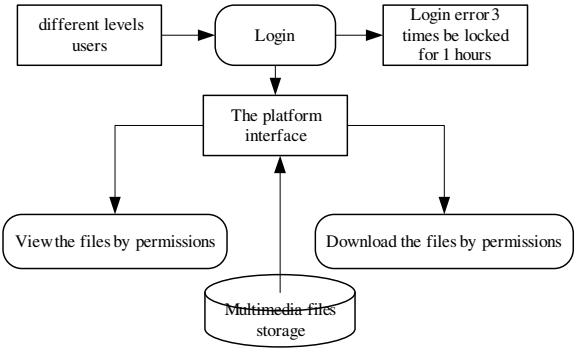


Fig. 1. The user operation process

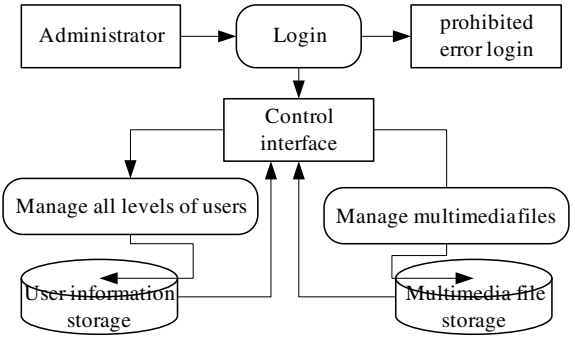


Fig. 2. The administrator operation

The module structure of multimedia file security communication platform based Web should include four modules of user management, message management, multimedia document management and multimedia document security management. Its specific module structure is shown in figure 3. The multimedia security management module is the most important technology module, which the multimedia files can be safely transfer in.

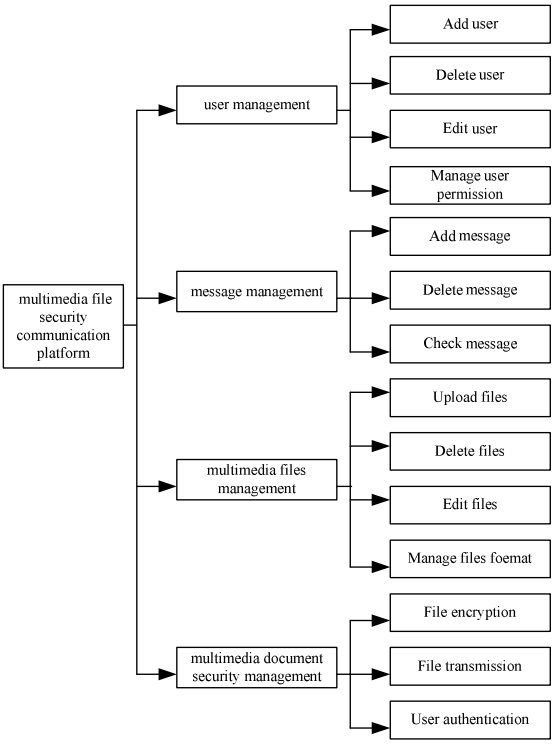


Fig. 3. Module structure of multimedia file security communication platform

5 Summary

Multimedia file security communication platform based on Web in this paper is the core technology of today network dissemination video file and its protection system, which breaks out the limit of the multimedia file transmission mode and achieves to encrypt a variety of multimedia files and identify and control different levels of users. Multimedia file security communication system developed in the framework not only is the multimedia file transmission platform, but is the platform to protect the intellectual property rights of multimedia file provider and the rights and interests of users, which can prevent to download the multimedia files in network communication.

References

[1] Yu, Y., Tang, Z.: Reseach on digital copyright protection technology. Journal of Computers 28(12), 1957–1967 (2005)
[2] Liu, F.: Design and implementation of the security certificate authentication center based on public key. Harbin Science and Technology University, Harbin (2005)
[3] Chen, X.: Research and implementation of Network security and identity authentication technology. Northwestern Polytechnical University, Xian (2003)

Research on the Security Issues of Cloud Computing

Jia Weihua[1] and Sun Shibing[2]

[1] Guangxi Electric Power Institue of Vocational Training, Nanning, China, 530007
[2] Changsha Socail Work College, Changsha, China, 410004

Abstract. This paper defines Cloud and its characteristics firstly, and gives Cloud computing architecture in detail, then introduces the security issues of Cloud computing and gives corresponding security mechanism against these security issues.

Keywords: Cloud computing, Cloud computing security, Access control.

1 Introduction

The produce of Cloud computing has its potential background. With development andpopularrity of the Intemet network application technology, espacially Web 2.0, the amount of Internet users and data grows rapidly which requests for higher data processing. In addition, the imbalance state occurs between the demand and utilization of network resources: some applications require large amounts of network resources, while a large number of network resources have not been fully utilized. Thus, integration and optimization of resource is the inevitable trend of network development, and Cloud computing came into being. Cloud computing is a kind of computing paradigm and means service-oriented architecture. However we should never overlook the fact that Cloud computing applications are accompanied by security risks, such as the risk of enterprise applications, data availability, and data integrity etc. Cloud computing security issues are critical issues in the study of, which affects the use and popularity of Cloud computing.

2 Concept and Characteristics of Cloud Computing

Cloud computing is a kind of large-scale distributed economy-driven computing mode, which provides on-demand to outside users with the abstact, virtual, scalable, manageable computation, storage, platform and service resource pools by Internet. Therefore, it is not a specific technology, but a new computing model, which builds up on virtual technology, distributed computing, utility computing, network technology, web and software services technology, grid technology and others related technology developed during the last ten years. And it distributes the computing tasks into resource pools formed by a large number of computers, so that all applications can obtain calculation capacity, storage space and a variety of software services according to need. In addition, it combines almost all computing forms of IT fields, such as distribution, grid, utility computing, on-demand computing, source; Web; P2P;

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 845–848.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Web2.0; SaaS etc. and it also synthesizes the skinny, fat and mobile client, which allows enterprises, business and service providers to access the cloud resources.

Therefore, cloud computing has the following characteristics:

1) Large-scale: only the massive Cloud computing can achieve its service advantage, especially its the ability of the service and scale economies of service.
2) Virtualization: it encapsulates each layer function to the abstract entities, and provides users with various levels of Cloud services with virtualization technology.
3) Reliability: Its development relies on Cloud services market, and the development of Cloud service depends on its service reliability.
4) Scalability: The size of Cloud can be dynamically expanded to satisfy with the scale growth need of the application and the users. At the same time, Cloud services also support the user application in cloud scalability.
5) Dynamic configuration: It can be customized on-demand and supply on-demand.
6) Economic: It relies on scale economies, which has the low cost advantage.

3 Architecture of Cloud Computing

Table 1 describes the architecture of Cloud computing. It is divided into five levels: physical layer, core layer, resource framework layer, development platform layer, application layer. Among them, five levels are corresponding to seven types of services, and each service can provide services for the user by UI interfaces of Web Services, and all services hava the reliable, secure, scalable, on-demand services, economic and others characteristics.

Table 1. The architecture of Cloud computing

Structure	Service form	Function
Application layer	SaaS (Software as a Service)	To develop various applications in the development platform and provide all kinds of distributed application services.
Development platform layer	PaaS (Platform as a Service)	To build development platform in the resource structure layer and provide all kinds of distributed application services.
Resource framework layer	IaaS (Infrastructure as a Service)	To build computing resources architecture in the inner nuclear layer and provide the distributed computing services.
	DaaS (Data-Storage as a Service)	To build storage resource architecture in the inner nuclear layer and provide the distributed storage services.
	CaaS (Communication as a Service)	To build communication resource architecture in the inner nuclear layer and provide the distributed communication services based on LAN or Internet.

Table 1. (*continued*)

Structure	Service form	Function
Core layer	KaaS (Software-Kernel as a Service)	To achieve the basic distributed resource management in the physical resource layer and provide the distributed application deployment environment through a variety of abstract services.
Physical layer	HaaS (Hardware as a Service)	To be fromed the geographic distribution local resources of cloud backbone and provide a variety of local support resources.

4 Security Issues and Solutions

Although the Cloud computing industry has great market prospects, its service has the following seven potential security risk to the enterprise users.

1) Access by the privileged users. Generally speaking, when the enterprise data is submitted to Cloud computing services, its priority access right will belong to not the corresponding enterprises but Cloud computing services. So it cannot be ruled out the possibility of enterprise data leaked.
2) To be examined. Although enterprise user submited data to Cloud computing service provider hosting, its security and integration issues finally would be in charge of by enterprise owneself.
3) Data location. When the enterprise customers used Cloud computing services, they did not know which server their data were placed in and even which country this server was placed in.
4) Data isolation. Large data of the enterprises user was in the shared environment in the Cloud computing services platform and even if the data had been encrypted, it cannot ensure itself without none danger.
5) Data recovery. Even if the enterprise users understood that their own data placed into which server, and got the service commitment, they must backup the managed data to prevent the occurrence of major accidents, or the enterprise user data can not be restored.
6) Investigation support. If the enterprise users just want to collect some data through legal ways, but Cloud computing service provider may not offer. Thus, if the enterprise user himself is service enterprise, when he need to provide additional user data collection service, he should not turn to Cloud computing services.
7) Long-term survival. If a Cloud computing service business is bankrupted or bought, enterprise customers' service will be interrupted or becomes unstable.

In view of the above seven risks, we put forward computing user security requirements as following.

1) Cloud computing services providers must provide and monitor the s pecific information of the privileged administrator and of control access.
2) They should be willing and able to comply with the relevant laws and regulations.
3) They should provide the contractual commitments, abide by the laws and regulations of relevant data management, and provide the query service of data storage location information.
4) They should provide data isolation storage services according to different data.
5) They should have the ability to recovery the log data fastly and comprehensively.
6) They should support certain kinds of investigation by the contractual commitments.
7) They should provide long-term security measures to risk development, example the user will how to get their own data, and how to get the data be imported into the alternative applications.

5 Summary

Cloud Service is a kind of Web Services based on Internet service, it faces all kinds of security problems because Intemet has many inherent safety defects and also exists in other attacks and threats. Therefore the development of Cloud Service depends on its security deeply, and it is a major significance to consensus on the Cloud Service security.

References

[1] Chang, F., Dean, J., Ghemawat, S.: Bigtable: A distributed storage system for structured data. In: Proceedings of the 7th Symposium on OSDI (2006)
[2] Ohemawat, S., Gobioff, H., Leung, S.T.: The Google file system. ACM SIGOPS Operating Systems Review (2003)
[3] Dean, J., Ghemawat, S.: MapReduce: Simplified data processing on large clusters. In: Proceedings of the 6thSymposium on OSDI (2004)
[4] Gorder, P.F.: Coming Soon: Research in a Cloud. Computing in Science & Engineering (2008)
[5] Vouk, M.A.: Cloud Computing-Issues, Research and Implementations. In: Proceedings of the 30th Int. Conf. on Information Technology Interfaces, ITI 2008, Cavtat, Croatia, June 23-26 (2008)

Alarm Analysis on Intrusion Detection of Network Flow

Liu Yanshu[1], Cao Yujun[2], and Lei Jiping[1]

[1] Hunan Mass Media Vocational Technical College, Changsha, China, 410100
[2] Department of Computer, Xiangnan University, Chenzhou, China, 423000

Abstract. This paper discusses network traffic self-similarity in new field of its security application, and describes Hurst parameter to determine whether the network is abnormal, and timely alarming based on establishing user' normal behavior model of network flow.

Keywords: self-similarity flow, Hurst Parameter, Intrusion Detect System, Alarming.

1 Introduction

Network intrusion early alarm technique researchs on network intrusion detection technology and enhances the ability of intrusion detection with the development of tracking technology. Therefore, it should become very difficult to early alarm, especially the system based on pattern recognition because it not only need recognize the known intrusion mode, but has the ability to deal with unknown ones. It may be more effective to early alarm with two angles, monitoring network abnormal and normal activities.

2 To Model on Network Intrusion Alarm

2.1 Intrusion Detection Model

Intrusion detection is an active security technology, which provides real-time protection for internal attacks, external attack and misuse and intercepts the response to the invasion before harm to network system. It does not completely achieve early alarm by the understanding of invasion process, but must has the appropriate detection model due to the complexity of intrusion activities. There are three kinds intrusion detection model by the way of the obtained original data: the intrusion detection system based on host, network and the distributed.

Intrusion detection can be invided into anomaly and misuse detection by another classification method. The former quantitatively describes acceptable behavior to distinguish the abnormal, potential intrusion behavior.and it relies on the anomaly model and different model constitution method. Common intrusion detection system has statistical anomaly detection method, anomaly detection method based on feature

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 849–853.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

selection, anomaly detection method based on Bayesi, anomaly detection method based on neural network, detection method based on machine learning and detection method based on data mining. The latter applies pattern matching to detect whether an intrusion behavior happens to by the predefined patterns invasion and the results observed the invasion. It assumes that it may be a coded attack in a accurate way, and can confirm that the intrusion behavior is a variant of intrusion method based on the same weakness according to capturing attack and rearranging. Common misuse intrusion detection system has the intrusion detection system based on the conditional probability, expert system and immunity.

Similarly, these models are complementary. Abnormal features model can accurately detect the known intrusion activities and has low false alarm rate, the normal features model has a more accurate system of normal work mode to the determined application environment, and find all activities deviated from the normal pattern, including some intrusion activities in the position.

2.2 Modeling on the User's Normal Behavior of Network Flow

We measure the actual flow in a long time and analyse probability distribution of flow arrival in busy, normal and leisure time, and then build the flow model on user normal behavior. We assume user arrival in the network will be consistent with λ in Poisson distribution, namely, the i user' probability arriving in the time t is

$$P(i,t) = \frac{(\lambda t)^t}{i!} e^{-\lambda t} .$$

The traditional telecommunication network describes user behavior with their call time, and usually is deeply convinced to be consistent with negative exponential distribution. But the user may access different host to obtain relevant information in the computer network. Therefore, the length of each flow (namely packet) depends on the information style in the sight of network flow characteristics because there randomly generates different length of the network flows. Suppose the flow number generated and its length are amenable to the mean geometric distribution, i.e.

$$P_f(n) = p(n \quad flows \quad from \quad single \quad user) = \frac{1}{E_f}(1-\frac{1}{e_f})^{n-1} P$$

$$P_t(k) = p(n \quad flows \quad from \quad single \quad user) = \frac{1}{E_k}(1-\frac{1}{e_f})^{k-1}$$

Suppose time intervals t_f in each flow and time intervals tp in packet of a low all obey parameters respectively μ_f, μ_p negative exponential distribution. That is,

$$F_f(t) = P(t_f < t) = \begin{cases} 1-e^{-\mu_f t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

$$F_p(t) = P(t_p < t) = \begin{cases} 1-e^{-\mu_p t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Packet number N (T) In (0, T) time can be expressed as

$$N(T) = \sum_{i=1}^{u(T)} \sum_{j=1}^{f_i(T)} n_{ij}(T)$$

Among them, u(T) is the number of users in the arrival time (0,T), f_i （T） is the number of flow generated in the time (0,T) by No. i user, and n_{ij} is the number of packets in the arrival time (0,T) of No. j flow of No. i user. The packet number in the arrival time （t,t+τ） is

$$N(\tau) = N(t+\tau) - N(t)$$

User behavior parameters in that mode can be specified according to the need to generate the required network traffic. But the average length of flow and the average intervals time of packets are determined with the actual network environment, and therefore the mode can appoint respectively customer arrival rate of each node according to the needs based on the total network traffic.

2.3 Basic Principle of Network Intrusion Alarm

H_{nor} is self-similarity parameters $Hurst$ of network flow model estimation based on user's normal behavior, H_{now} is self similarity parameters $Hurst$ of network current intervals, $\Delta H_d = | H_{now} - H_{nor} |$, we do calculus within not repeating subsets to ΔH_d, and then its average value is ΔH_{dav}. We describe the network attack surface size by ΔH_{dav}, and we think that the network may be a continuous attack when the ΔH_{dav} values are relatively large in a certain time period.

3 Performance Test

3.1 Network Flow Test

Test result of intrusion detection system is objective and comprehensive only in real network background flow situations. Data packet from flow simulation method of ordinary study to network performance can cause a large number of false alarm in the intrusion detection system because it does not consider data content. Therefore we analysize the actual network flow with the method of divided period in the intrusion detection system testing environment, and then estimate the self-similarity parameters $Hurst$ of network flow with statistical data and calculated data. and we think that network flow is anomalies and alarm timely according to the principle of intrusion when ΔH_{dav} changes greatly to compare the estimated $Hurst$ parameter value of

current actual network flow to the *Hurst* value of the normal behavior in flow model.

3.2 Attack Flow Test

The core of attack traffic test is the key to test on the intrusion detection system. Firstly, we should collect various attack methods as far as possible. The safety sites, forums and mailing lists in Internet have a large number of network security information and some countries and companies have established a fairly complete system vulnerabilities and the attack data stock, which are the attack source. We add the attack flow related to the flow change in the actual network flow and carry concealed treatment to some attack methods in the test. And we analysize the comparison results of the *Hurst* parameter value by capturing the actual flow joined the attack flow with of the normal user's behavior in the intrusion detection system rule database, and determine based on intrusion alarm principle wheather the actual flow is subjected to a continuous attack and has abnormal phenomenon.

3.3 Test Result

We found some common network attacks (such as Fake UDP, Fake TCP (SYN), FakeICMP) and can carry out the effectively and early alarm with the network flow test and attack test. Also there will alarm timely on the aggressive behavior of the flow change by network scanning, denial service.

4 Summary

This article did exploratory research on the network flow self-similarity and its application. It is an important and indispensable link to establish information infrastructure security environment that the intrusion detection system alarms effectively in network users and illegal behavior. And from the experiment and its results, we determine whether the network traffic is abnormal through the comparison of network flow self-Similarity parameter value in the system based on *Hurst* parameter value in the network flow model of user normal behavior. Therefore it is feasible to alarm timely in the apparent attack behavior causing flow changes.

References

[1] White, G.B., Fish, E.A., Pooch, U.W.: Cooperating Security Managers: A Peer based Intrusion Detection System. IEEE Network 10(1), 20–23 (1996)
[2] He, H.-C.: Introduction to Artificial Intellingence, pp. 10–22. Xi-an Technology Press, Northwestern University (1998)
[3] Forrest, S., Hofmery, S.A., Somayaji, A.: Computer Immunology, pp. 10–22. Communication University of Technology Press (1998)

[4] Xie, K., Min, Y.: A network flow model based on measurement. Computer Engineering and Science 123(5), 51–53 (2001)
[5] Yu, F., Zhu, M., Zhu, M.: Improved on feature matching of Intrusion detection system. Computer Engineering and Applications 39(29), 32–36 (2003)
[6] Yu, Z., Liu, J., Lei, Z.: TCP performance analysis on the RSVP. High Technology Letters (July 7, 2002)

The Exploitation and Design of Enrollment and Employment Management System That Based on B/S

Li Bo

Jiangxi Science & Technology Normal University, Nanchang, Jiangxi, 330038
jxnclibo@126.com

Abstract. With the ceaseless perfection of our educational institution, the management of enrollment and employment has been the important component of education management in colleges and universities. It is directly related to the daily operation and management. This article compares the traditional C/S mode and B/S mode, discusses the asp.net advantages, and how to use asp.net to explore and design the system of enrollment and employment management that based on the B/S.

Keywords: management system of enrollment and employment, B/S, ASP.NET, development design.

1 Introduction

With the rapid popularity of information technology, the paperless office has penetrated into the daily life. In the college management, if we do not accept the effective management strategy, we will face the disordered situation when managing the thousands of students. While influencing the normal operation, it will affect the students benefit. In the present enrollment and employment management, there exists the mix condition except the huge information amount. The traditional artificial management will waste plenty of human power and material resources. The exploitation and design of the management system can effectively avoid this condition. Through the perfect and scientific management system, we can depend on the information technology advantage to standardize the information management and improve the management quality of enrollment and employment. Therefore, this article will start from the following aspects and simply evaluate the enrollment and employment management system that aim at exploitation and design questions.

2 System Environment and Exploitation Platform

During the systematic exploitation and design of enrollment and employment management, there needs to combine with the practical development tendency. The

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 855–860.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

pertinency exploitation can provide convenience for the system usage and fully express the advantages. Aiming at the systematic exploitation and design of environments and platform, the detailed evaluations are in the following:

2.1 Environmental System

In the enrollment and employment management system, Visual Studio.NET is the set of utility component for achieving the NET technology. Different from the prior components, this set of component possesses multiple new functions, such as the different types of language programming and the integration of exploited environmental programming. During the software system operation, each part needs the coordination of relevant staff. These people have their own familiar language programming during the operation. If we unified them, we will obtain the barely satisfactory. However, the systematic exploitation can effectively solve this question. In the traditional software exploitation of enrollment and employment system, one system has only one language with great inconveniences during the operation and it will directly influence the management quality. Therefore, we need to exploit the software with multiple languages. It can ensure the management quality and increase the management effectiveness. In the exploitation of this system, NET obtains the popularity based on the advantages. Different operator can select the familiar language component during the operation, use the Visual studio. NET to do the integrated compiling, and put into trial operation in short time.

Table 1. Operation platform and exploitation platform of each system part

Server-side operation system	Windows 2003 Server/Windows 2003 Advanced Server
Client-side operation system	Winddows98/2000 Professional/XP
Server-side data source	SQL Server/ Oracle9i
Development tool platform	Visual Studio.NET
Development tool language	Visual C#,NET,Visual Basic.NET

Through the above table, we can find out during the system operation, the browser component in the program set has not adopts the Vc#.NET language exploitation. However, the part of data management software uses the VB.NET language exploitation. When the system operates, we can compile through the Vi8ual studio.NET platform. This method can save the compilation time and ensure the quality.

2.2 System Mode

In the traditional C/S mode system, it includes client and server. During the entire system operation, the client has the disadvantage in the language input and output that compare with the server. The target is store the amount of data during the operation and integrates the data to the client and server through the programmer software. It can provide convenience to the decrease of the network communication amount and server counting quantity in order to ensure the stable and rapid system operation.

However, during the practical operation, for the limitation of server and data communication quantity, it will lead no space between the two-layer architecture. Even the program exploitation will has the disadvantages of high cost, low effectiveness, single user interface and the expensive maintenance. It will influence the program operation while effect the enrollment and employment management. With the rapid development of information technology, the B/S mode improvement makes up the disadvantages of C/S mode to a certain degree. During the software operation, the business logic perfection can achieve by the web server and the client will manage by the browser. This can decrease the client burden, increase the program operation, and obtain the popularity.

3 The System Design of Enrollment and Employment Management in Colleges and Universities

Different from the traditional management, the system of enrollment and employment management has great relations with the development and inner management in colleges and universities. This needs the designer can integrate the various parts of enrollment and employment in order to express the advantages during the system design. The system design has the following aspects.

3.1 The Design of Data Presentation Layer

During the sequence operation, data presentation layer is also called as the user interface layer. The function is expressing the relative browser interface in the client browser for the further data information selection. In addition, the layer sequence operation can sift the user information and interactive with the business logic layer under the operation. This can ensure the selection and comprehensiveness of the information selection. In the enrollment and employment management, people can through this layer to search the various results, information of enrollment and employment. The operator needs input the simple required information to find out the correct page. The data presentation layer can obtain more than 20 web form pages and each page has the relative storage file. This will avoid the interference during the 20 pages' operation at the same time. This can ensure the operation quality. Aiming at this condition, the system designer needs to support the management on time. It includes the following management methods. In the first place, we can set one EntitData class in the core database. The operation of this class has some functions that different from other classes. It can carry on the DataSet system, which based on the original basic and combine the original method to optimize it into the frame of the whole entity class. Except operating in each entity class, we can ensure the information security during the operation. In the second place, the definition of each entity class needs to determine through XML files and through the file we can ensure the detailed structure of the entire entity class. In the last, the detailed structure determination needs to aim at the set of class constructor and follow the above requirement to finish the formation in order to avoid the program errors during the operation.

3.2 Business Logic Layer

During the enrollment and employment management system, the business logic layer is placed in the middle of the entire system. At the same time, it is the core center of system operation. The basic function of this layer can connect the user and data service, influence other users' transmission requirement, and execute the program operation. In other words, the operation of business logic layer will influence the various functions' achievement in the entire system. The business logic layer can finish the mission through NET component. It includes the definition and requirement of various classes and includes the table manufacture of each background database. Through the business logic layer, we can effectively achieve the business logic function of table operations.

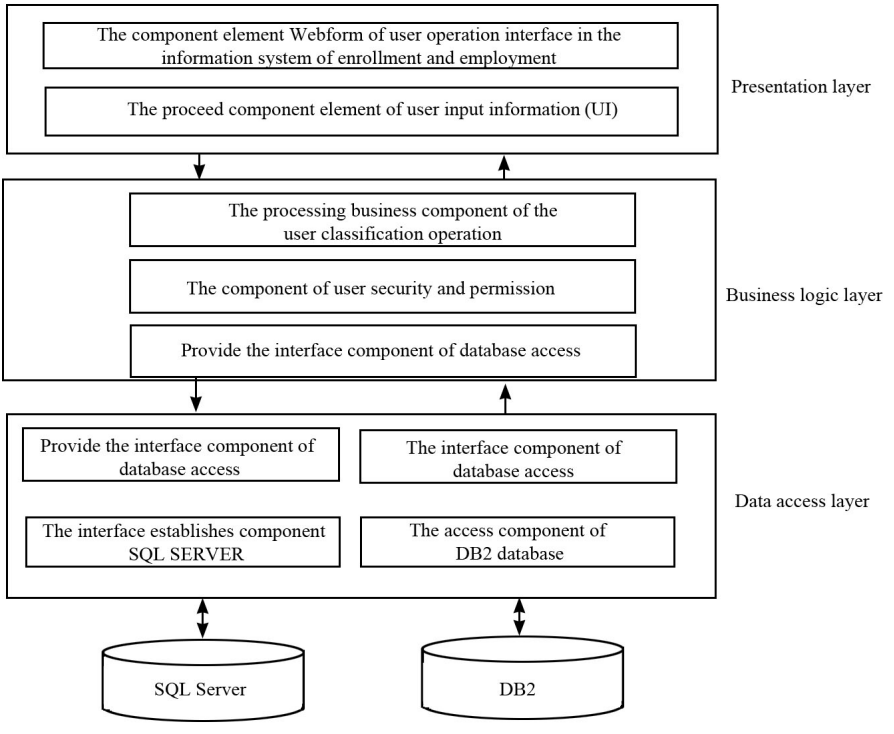


Fig. 1. The applied structural model of enrollment and employment information system

From the above table and figure, we can find out during the operation of business regulation layer, the independent business organization cannot depend on one program. It needs the various business regulation operations in the logic layer. When solving this question, the core center is whether the business regulation layer can establish the unified integration with the data access layer. In order to ensure the program operates, we should develop the system flexibility, encapsulation, and the

importance. The system operation will block some clients to increase the high cohesion and low coupling. The entire operation of the user is not the practical operation of the computer program and it is the interface or other programs. The design of business regulation layer needs to set the modeling work, ensure the program operates, and develop the management quality of enrollment and employment in the colleges and universities.

3.3 The Design of Data Access Layer

In this layer design, there combines multiple data access components and each component will influence the operation. Every component combines with data access operation, user class, resource and other operations. The success of each component needs the core center of satisfying the logic layer requirement. In the entire data access layer, we can depend on the database to adjust the data access component. The biggest advantage is the most effective package of the whole data access layer to avoid the information leakage. However, we can interact through the data access layer during the logic layer operation and achieve the various access requirements. In the above modes, the sub-layer in each access layer can express the application function. Except the effective transmission, it can ensure the operation. The entire design of data access layer has the following evaluation.

(1)Data access interface model

During the interface model operation of data access, the basic function is supporting interface for the up layer users and test the operation condition to avoid the illegality information intrusion. The detailed function is achieved by one category. For the Istudetn, it will achieve by EqManSys, SqlServerDAL, Student class.

(2)Data access factory model

During the program operation, the basic function of the data access factory model can package and expand the data. When extends the data operation range, it can use the abstract engineering model during the access layer design. In the practical operation, the up layer users do not need to guide the background database. They only need to prepare the interface method during the access process. After determining the method, the interface can search the relative data access layer that based on the input data. It can save the access time and greatly improve the program running speed.

4 Summary

From the above, with the rapid development of social economy, the exploitation and design of enrollment and employment management system in college and university, when perfect the entire management, it can standard the management work. Therefore, here we need the relative staff when design the system, they can combine the practical development condition and design the system with pertinency. This can express the system advantage during the operation. It can improve the college management quality and establish the basement for the further development.

References

[1] Liang, X., Yang, Z.: The technology research of enrollment system exploitation. Journal of Xi'An Technological University (02) (2006)
[2] Wu, X., Zheng, J.: The design of enrollment and employment website information management system that based on the ASP. Computer Engineering and Design (11) (2006)
[3] Zhong, Z., Tian, C.: The data extract research and application of enrolment system. Computer Science (02) (2006)
[4] Mei, Z., Jing, Y.: Data access technology in Net– ADO. NET. Microcomputer Information (01) (2006)
[5] Wu, Z., Wan, B.: Information resource establishment model that based on Web. Application Research of Computers (03) (2010)
[6] Mao, D., Luo, R.: The three layer design model of Web application programming that based on ASP. NET technology. Microcomputer Applications (03) (2012)
[7] Wen, C., Lin, Y.: The system model of bidding decision support that based on the data mining. Journal of Xiangtan Normal University
[8] Zhu, L., Zhang, S.: The application technology of computer network. Intellectual property rights press, Beijing (2009)

Research on Network Technology-Based Data Integration Model

Chen Min[1], Wang Jianping[2], and Tan Shaosheng[1]

[1] Department of Information Engineering, Hunan Industry Polytechnic, Changsha 410208, China
[2] Changsha Aeronautical Vocational and Technical College, Changsha 410014, China

Abstract. The traditional data integration technology can't meet the requirements of timely accessing to updated data, while the publish/subscribe mechanism can support the dynamic, asynchronous and many-to-many communication between loosely coupled systems, getting the participants of the communication completely eliminated in space, time and control flow. This paper, through research on data integration technology and information distribution middleware technology, has designed and realized a DDS and XML-base data integration model applied to distributed network. This model achieves a data-centric publish/subscribe communication mechanism with advantages of flexible, loosely coupled and easy scalability, etc.

Keywords: network technology, data integration model, data distribution.

1 Introduction

With the rapid development of computer networks, especially the Internet technology in recent years, the distribution of computing resources is increasingly spreading and with higher interconnection. However, tens of thousands of entities involved in data exchange may be distributed all over the world and their behaviors may vary dynamically with time, causing lack of association between system entities and emerging an "information island". In modern society, information drives all of the commercial activities. Thus, how to access to information and how to ensure the information in real time and effectiveness are essential, which directly reflects the demand for data integration, the purpose of which is to achieve data sharing between heterogeneous data sources to effectively use resources and to enhance the performance of the entire application system. The development of distributed computing environment has brought new challenges to the network application system software architecture and communication mechanisms. On the one hand, people bring out higher demand on the reliability of performance, security, adaptability, and resource utilization efficiency, and on the other hand, in a distributed network environment, the enterprise data integration technology faces many new problems, including the distribution of the data integration technology, asynchronous, dynamic, scalable, and so on.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 861–867.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

These problems, however, are difficult to be fully solved with the traditional data integration technology. For example, such as point-to-point data integration mode is to integrate data in different applications by transforming the data from one context to another, which will cause a complex data integration structure difficult to extend, so it is not suitable for enterprise data integration task within the scope of the WAN; the integrated model of data warehouse, through the definition of a global mode and local mode mapping, provides data interoperability to collect, clean and integrate a series of distributed data, but its integration goals is limited to the known uses, losing flexibility and adaptability, so it is difficult to meet the dynamic and scalability requirements of data integration. Middleware structure-based data integration mode, with its flexible architecture, has become a hot topic, but the traditional middleware-used point-to-point and synchronous communication model has a tight coupling characteristics, which can not meet the requirements of loose coupling communication. In contrast, the message broker mode of publish/subscribe communication model can eliminate the above problem on time, space and the control flow aspects. The publish/subscribe middleware supports asynchronous, dynamic and many-to-many communication between loose coupling systems with good expansibility, which can be used to solve the new problems of data integration technology under in the WAN environment. The data integration technology based on publish/subscribe communication model has gradually become the point and been increasingly used.

2 The Structure of Integrated Model System

The DDS and XML-based integrated model structure is shown in Figure 1. The main function of this model is to integrate data of the various subsystems, use application adapter translate the subsystems' message format into XML format to shield different data formats from disparate systems, making subsystems agreement at the semantically level and enhancing their expression capability. The DDS and XML-based publish/subscribe message bus, using data-centered publish/subscribe communication model to realize the match between the release of XML messages and XPath subscription request, achieves real-time, reliable communications between subsystems. It uses DDs' Qos control and support for real-time to ensure that subscribers receive news of interest timely and accurately. This top-down framework is divided into three layers: application layer, adaptation layer and the message bus layer.

Application layer: the specific application of the various information subsystems, as each information subsystem has different applied logic and message formats. The common point of these subsystems is the need of real-time and reliable messaging transform and higher information expression skills.

Adaptation layer: aims to translate the applied logic and information format in a particular subsystem into a uniform format provided for the data layer and the message bus layer: translate the message format into XML format and the message request format into XPath format. These functions are completed through application adapter.

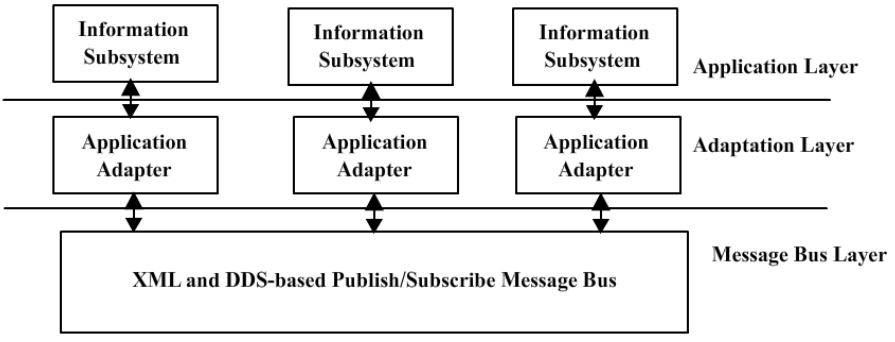


Fig. 1. Integrated Model Framework

Message bus layer: the core layer in the integrated model, comprised with DDS and XML-based publish/subscribe message bus. Its responsibility is to complete message transmission, matching, QoS management and transform message accurately, reliably and real-time from an information subsystem to another.

3 The Detailed Design and Implementation of Application Adapter

Messages exchanged between the various information subsystems are inevitable differences in the syntax and semantics and message formats of the application system maybe differ from that processed by the message broker, so it needs to follow some kind of mechanism for translation and conversion to eliminate the above differences, converting messages between subsystems correctly. Message broker is able to process data from publisher and subscribers. Thus, this paper acts XML as a unified message format in data integration model.

In this model, publish/subscribe message bus and each information subsystem play the role of publisher, subscriber, or both. Each subsystem has its own message format and business logic, but the message bus in this article needs to process the publish message and of XML format and the expression of the subscription request. In order to shield the differences of subsystem message format, we need to complete mapping of XML messages from the subsystems' local news to the message bus, which needs to assemble the subsystems' local news into the message format required by the message bus.

Because each information subsystem has its own specific message format and business logic, the specific implementation of message package by application adapter varies with different subsystem. The processes are:

Information release with XML assembly: obtain data required to be published from the information subsystem and assemble these fields into released information in XML format.

XPath subscription request assembly: obtain subscription requests expressed by string equation or numeric expression table from the information subsystems and convert them into a subscription request expressed by XPath.

Due to differences in the message format, as well as the business logic of each subsystem, the XML messages from the message bus can not be directly used by subsystems, but need to conduct XML parsing and formatting to the messages from the message bus. In the integrated model of this paper, the message from the message bus is the released message in XML format gained by subscribers. The XML message parsing and format conversion completed by adapter is to parse the released message in XML format received by subscriber and convert it into format required by the information subsystem. To parse the message in XML format, we can use DOM (Document Object Model) or SAX (Simple API for XML).

When using the DOM, save the XML document to a hierarchical object node tree matching with the structure of XML documents and information then obtain the necessary information by querying the node tree. It needs to read in the entire document when using the DOM parser, and then build an in-memory tree structure. The advantages of this scheme are: the whole document tree resides in memory, easy to operate; allow operations of deletion, modification and rearrangement and so on; the drawback is the entire document is needed to put into memory (including useless node), consuming memory space and extending processing time for parsing document.

SAX is an event-driven lightweight program. When using SAX, you do not need to read into the whole XML document. One the parser found the elements beginning, end and the beginning or end of the text or document; send the event and store and process XML data through writing code to respond to these events. Its advantage is the analysis of the document can be started immediately rather than waiting for all of the data is processed. Moreover, the application checks the data only when reading the data, but not to store the data in memory. In general, the SAX is much faster than DOM but its disadvantages are also obvious, because it does not store data in any way. Using SAX is not able to access to XML documents at random, and be more difficult to modify the data.

In the integrated data model of this paper, as the number of released message to be processed in a single subsystem, and subscription requests are relatively small, and at the same time messages may be required for assembly, transform and modify, this paper chose DOM as the XML parsing program of the application adapter.

By comparison of indicators of a variety of XML parsers in inspection of the document format, XMLSehema support, namespace support, support for SAX and DOM versions, the C language support and processing efficiency etc. the data integration model in this paper select IBM's Xerees C - as an application adapter's XML parser. Xcerces – C - is a very robust XML parser, which provides two ways of parsing XML document: DOM and SAX, and it supports for Windows, Linux, Solaris, and other systems. Messages in XML format after message parsing, based on subsystem business needs, are converted from an XML syntax format into another. The process is known as the format conversion of XML message, which is the reverse process of XML message specific to the specific subsystem.

In order to achieve the message content-based match mentioned above, and modify as little as possible to the original DDS standard to save the OMG DDS standard's advantages of real-time, data-centric, reliable and powerful Qos control etc. we use a novel approach to express released message and subscription requests in the DDS and XML-based message bus.

Message release:

Release the message topic: we reserve the use of topic, but the Topic used here is compatible with the OMG DDS standard, and use the QoS support to differentiate between each published message instance, rather than be used as a basis matching with subscription requests. We create a string of topic for each released message, which is unique in each released message topic. We take the "XML+ system time+ five random numbers" as the topic of the released message. For example: the topic of the released message "XML200809121045206 5464", 20080912104520 is the system time of 10:45:20, 12th, September, 2008; 65464 is five random numbers. The above released topic, in most cases, can be the unique in a publisher's all released messages.

Released message type support: in the OMG DDS standard, each topic has its own unique data types. This article's data type is defined by IDL, and uses TAO's IDL compiler to process the type definition and dcps_ts.pl script to generate the code needed to transport this type of data. The following message type IDL definition:

```
Module Messenger {
#pragma DCPS_DATA_TYPE "Messenger::Message"
#pragma DCPS_DATA_KEY "Messenger::Message subject_id"
Struct Message {
      String from;
      String xmltext;
      Long count;
};
};
```

DCPS_DATA_TYPE marks for the data types used in subsystem, such as the message used in this article marks for its topic name. The message type body is expressed by a structure whose member types can be numeric (short, long and float etc.), enumerated type, string type, sequence type, array type, structure type and the type of public body. DCPS_DATA_KEY identifies the keys of the topic to distinguish different instances with the same topic.

In the implementation of message bus in this article, the released message content is in XML format. As shown the above, the second field xmltext represents a message body in XML format, and the "from" and "count" field is a reserved field for future use.

The following is the possible message body stored in xmltext field:

```
Bookstore>
<book>
   <title>Harry Potter</title>
   <author>J K. Rowling</author>
   <year>2005</year>
   <price>29.99</price>
</book>
</bookstore>
```

Subscription request:

The topic of subscription request: in the design and implementation of DDS and XML-based publish/subscribed message bus in this article, the subscribers use XPath expressions to indicate their interest in the publication of the message content. Similarly, in order to minimize the modified DDS standard, we still use the topic. The topic of the message request at the subscriber side is a subscription represented by Xpath.

The process of processing subscription is as follows: First use XML parser to parse the released message in XML format and distribute a node id for each element, attribute and text.

The following is the released message after being parsed:

```
Message: id=12
        XML:
        <a(1)>
                <b(2)>
                        <c(3)><d(4)>v1(5)</d></c>
                        <p(6)><p2(7)><q(8)>v2(9)</q></p2c></p>
                        <x(10)@y="v3"(11)>
                                <z(12)>v4(13)</z>
                                <z(14)>v5(15)</z>
                        </x>
                </b>
                        <b(16)><b(17)><x(18)@y="v4"(19)/></b></b>
        </a>
```

For each text node value or attribute value of each released XML message, tagpath, idpath and value generated after SAX parsing are labels and id on the path of from the root node to value respectively. Let's use "/" to split the label on the label path to match the linear expressions of subscribed XPath, and you will find that tagpath does not contain the "//".

4 Summary

The traditional data integration technology can't meet the requirements of timely accessing to updated data, while the publish/subscribe mechanism can support the dynamic, asynchronous and many-to-many communication between loosely coupled systems, getting the participants of the communication completely eliminated in space, time and control flow. This paper, through research on data integration technology and information distribution middleware technology, has designed and realized a DDS and XML-base data integration model applied to distributed network. This model achieves a data-centric publish/subscribe communication mechanism with advantages of flexible, loosely coupled and easy scalability, etc.

References

1. Object Management Group, Data Distribution Service for Real-time Systems Specification. Version1.1 (January 2007)
2. Ives, Z.G.: Efficient Query Processing for Data Integration. Washington University, Washington (2009)
3. Li, Q., Li, Y.: Research on Integrated Frame and Implementation of Federated Heterogeneous Database Application Systems. Application Research of Computers 18(4), 19–22 (2008)
4. Chawathe, S., Garcia-Molina, H., Hammer, J., et al.: The TSIMMIS Project: Integration of Heterogeneous Information Sources. In: Proc. of IPSJ Conference, Tokyo, Japan (2004)
5. Shanmugasundaram, J., Meman, J., Shekita, E.J., et al.: Querying XML views of relational data. The VLDB Journal, 261–270 (2009)

Sentiment Analysis in Chinese BBS

Dai Lu and Ding Lixin

State Key Lab of Software Engineering, Wuhan University, Wuhan, China

Abstract. With the rapid growth of the web text data, BBS (Bulletin Broad System) has already been the popular discussion forum for people exchanging their thinking and mind. If we can mine and analyze these online review data posted by the users, these data can help better understanding of public opinions greatly and plays an important role in government or enterprise's construction, especially in supporting for proper decision-making. In our study, we focus on building attitudinal words weight dictionary which is composed of 1342 words, and introduce the negative words dictionary, degree words dictionary and interjection words dictionary defined by ourselves. Then we create emotion weight calculator to calculate the emotional index of each post and classify the sentence as one of three opinion groups, positive, negative and neutral. The subject is the post we crawled from the famous Chinese BBS social network at random. The result shows that this sentiment analysis method is very effective and inspirational.

Keywords: emotional words, weight dictionary, emotional weight calculator, sentiment analysis, BBS.

1 Introduction

BBS (Bulletin Broad System), which is constructed in internet for providing public information to people, is one type of electric information system and social net works has become a popular media of communication [1]. Due to the accessible and high-speed nature, it is an important way to share and disseminate information and ideas [2]. People from every corner of the world can express their opinions if he/she is interested in or has own opinion for one post. Given all these posts released by the uses for sentiment analysis, we can analyze and mine the text information effectively, recognize the sentiment trend-"happy", "sad" or attitude "positive", "negative". In the traditional sentiment surveys, we used to classify the sentiment analysis into two groups that one is the board and another is the narrow. The board analysis includes speaker's mental attitude. On the contrary, the narrow analysis focuses on the attitude to the point "yes" or "no". The core subject is web text, especially the user's of the subjects and determine the emotional trend of the Posts relevant to the theme. We can identify the emotional subjects [3].

In 2006, Chung-Hsien Wu [4] applied a novel method includes the construction of emotional relevant rules, automatic emotional identification and simplifying the emotion into happy, unhappy and indifferent. In 2007, Kazuyuk Matsumote [5] indentified the attitude of dialogue by the means of constructing emotional dictionary, they offer weight to some words based on the context. In 2008, Alastsin J Gill [6] started

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 869–873.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

to estimate emotion in blogs. He found that when the keywords with strong tendency appear in longer text, the outcome with higher accuracy will be easily reached. For our study, we combine Chung-Hsien Wu and Kazuyuk Matsumote's main idea, and construct weight dictionary to identify the sentiment of text.

2 Methods and Materials

Sentiment mining is based on the index that we offer to every web data. We define it as emotional index. Higher index represents more positive emotion; lower index represents more negative emotion.In our study, the weight dictionary is classified into 4 parts: attitude words dictionary, negative words dictionary, interjection words dictionary, degree words dictionary. The selection of corpus is related to coverage rate. The coverage means corpus distribute in different areas, which are constructed by four dimensional model made up of time axis, space axis, subject axis and style axis [7]. Since the blog has 140 characters and majority of users concern of weather, life, film, feeling and emotions on everyday topics, there is a great deal of no communication among the users makes blog impossible for that four axis . Posts are similar to blogs. They are shot, fast, effective and use a relatively short article to express a sense or a kind of debate. The feeling and perception will be recorded in just a moment by a few words. It is apparent that such instant statement caters to the rush of our life. So we extract keywords form posts and computed it to get the emotional index after the removal of advertisement and recommendation.

In 2008, Xu Linhong did the survey on the construction of emotional corpus and analysis. He collected abundant corpus form textbooks and literature, took the whole text as the object, tag sentences. On the contrary, posts are measured by sentence, tag words. He proves that the words with no emotion are more than other words, which also means that the definition of neutrality is feasible. Additionally, the study does not offer the appropriate solution to division and judgment from degree words and negative words. In fact, negative words have great influence on feeling colors, in particular the statement of any ground for complaint tendentious .

We construct negative words dictionary afterword. In Chinese text, multiple negative words may exist at the same time within one sentence. When the negative words appear odd times, it indicates negative semantic. When those words appear even times, it indicates positive semantic. First, we define the weight of vocabulary, introduce the concept of attitude words and break up the classification of verbs, adjectives and nouns. Attitude words suggest people's attitude, such as joy and sorrow. Most of the attitude words are composed of verbs, adjective adverbs, and a part of nouns. In the meanwhile, not all of the verbs, adjective adverbs can indicate attitude. After filtering the irrelative words, we get the useful attitude words and offer corresponding weight defined from -20 to 20 to each of them. For example, ecstasy is 20, despair is 20. A positive index indicates positive mood, negative index indicates negative mood. The higher absolute index of the word, the more positive emotion can be expressed. We choose 2000 pieces of data at random based on the extraction form micro-blog fanfou then filter the advertisement and promotion. The rest 1433 pieces are artificially marked and weighted on average by three individuals. Finally, we get the initial weight dictionary consists of 524 attitude words ordered by weight. Moreover, the

negative word dictionary contains 13 common negative words. In order to get more complete weight dictionary, we restore to synonym dictionary for the artificial vocabulary expansion. By the means of 7 people making and weighting on average, we get the existing weight dictionary of 1342 words. Therefore, in the following study, we observed that degree words and interjection words work more effectively than emotional tools. In the view of the fact, we add 10 more degree words and 16 interjection words on the basis of previous ones into the dictionary. Since most Chinese interjection words can be divided into two categories: the positive and the negative, some of them can be easily identified usually play an important role in sentence, we offer them 1 or -1 two kinds of weight attribute to determine the index of whole sentence.

In this paper, we use a lot of micro-blogs rather than textbooks, literature as corpus to construct weight dictionary for two reasons. For instance, on the basis of weight dictionary construction, we contribute to the sentiment mining and emotional evolution by analyzing present document and textbooks. It must be pointed out that if the method of building a weight dictionary is different, the weight dictionary will not be same, emotional index calculator will be different either.

3 Experiment

According to the weight dictionary we have built, we compiled emotion weight calculator. The algorithm is as follows:

(1) Read the post a;
(2) Break up this post by punctuation, then get a1,a2,a3…an;
(3) Start from the first clause a1,search the attitude words contained in the weight dictionary, overlap the weight and get v1,delete the matched attitude words in a1, then get a1';
(4) Search all of the negative words included within the negative words dictionary form a1', then delete it, get a''. judge the number of negative words is odd or even. If it is odd, then change the weight of a1 into –v1;
(5) Search a1'' and judge whether it has the degree words included within the degree words dictionary and calculate the number. One more degree word, double the emotional index and delete it, then get a1''';
(6) Search a1''' and judge whether it has the interjection words included within the interjection word dictionary. Make the absolute emotional index of this clause multiplied by character weight of interjection. Since the interjection words can be divided into two categories, positive ones and negative ones in Chinese. For instance, the words "ha ha", "he he", "hey" can play a dominant role in the judgment of emotional index trend. If the whole clause contains negative interjection words, then change the attitude weight into –v1;
(7) Search the next clause a2, jump to (3), start to calculate v2;
(8) After we get vn, add v1, v2,….vn according to specific weight proportion, the emotional index of post a1 is finished;
(9) Read the next post b and repeat (2).

The emotional weight F can be computed as:

$$F = \begin{cases} \sum_{i=1}^{n} ExcWeight(a_i) \mid \sum_{j=1}^{m} A(a_i, W_j) Deg(a_i) Neg(a_i) \\ \sum_{i=1}^{n} \sum_{j=1}^{m} A(a_i, W_j) Deg(a_i) Neg(a_i) \end{cases} \tag{1}$$

$$Neg\ (a_i) = \begin{cases} 1 \\ -1 \end{cases} \tag{2}$$

In our study, we choose the Tianya BBS, one of the most popular discussion forums in China. We calculate 4242 discussion posts at random, and analyze the sentiment by the sentiment calculation we created after the pre-process (figure 2). We observe that the accuracy reaches 80.6%, compared with the outcome we analyze artificially. In order to illustrate how the sentiment calculator works, we crawl some posts relating to house price the broad masses of people concerned with in China and then translated into English. First, we filter the advertisement and promotion. Then we convert the file format into txt.

The house price 6000 per square or 5000 per square are the same to me, I can't afford anyway, raising, come on.

Author: Jinlin Reply on: June 28, 2009, 22:22:03

80s cannot afford so expensive house, we will spend our whole life to work for the bank and developers, whats pity.

Author: luohuo Reply on: June 28, 2009, 22:26:34

It's good to see the house price higher and higher. There are so many people become rich and middle-class because of buying an apartment and selling it. strongly support the rising house price.

Author: zhuangzhuang Reply on: June 28, 2009, 22:28:23

The Chinese house price is always leading the way. I am not sure whether it looks like this. But I do hope the government will notice.

Author: longlong Reply on: June 30, 2009, 07:12:16

Without the subsidies, house price will beyond the purchasing power, then the house will belong to a small part of person forever.

Author: gdeee Reply on: June 29,2009, 23:27:42

Since 2008 the estate industry has already kidnapped the state economy. Government needs to keep the employment and the harmony put a lot of fund to the market, leading to huge bubble.

Author: Vivian Reply on: June 30, 2009, 12:02:26

Fig. 1. The file format txt

4 It's good to see the house price higher and higher.

7 There are so many people become rich and middle-class because of buying an apartment and selling it.

15 Strongly support the rising house price.

Fig. 2. Posts with positive attitude

-21 the house price 6000 per square or 5000 are the same to me, I can't afford anyway, raising, come on.

-8 80s cannot afford so expensive house, we will spend our whole life to work for the bank and developers, whats pity.

-7 without the subsidies, the house price will beyond the purchasing power, then the house will belong to a small part of person forever.

-15 since 2008 the estate industry has already kidnapped the state economy.

Fig. 3. Psts with negative attitude

From the above figures, we can get a clear sentiment analysis view. Nearly half of the posts we crawl are neutral to the house price. The negative users are far less than the positive users, which probably imply that the house price is appropriate for those

people's taste. Only a small account posts with the absolute index higher than 20 indicate the user's reaction to this issue is a little unreasonable.

4 Conclusion

This paper presents an effective way in calculating emotional index and introduces self-defined attitude words dictionary and weight dictionary which makes sense of sentiment analysis in the BBS social network respectively. Despite the result of this study is encouraging, limitations are observed apparently. The characteristic of Chinese is complex sentence type and unpredictable style of expression. Compared with traditional text categories, emotional identification is inherent challenging and difficult. For the further survey, we will expand the weight dictionary and improve the emotional weight calculator to reach a higher accuracy.

Acknowledgement. This work was supported by the Doctoral Program of Higher Education. (Grant No. 20070486081).

Reference

[1] Ding, F., Liu, Y., Cheng, H., Xiong, F., Si, X.-M., Shen, B.: Read and Reply Behaviors in a BBS Social Network (2010)
[2] Zhou, L., He, Y., Wang, J.: Survey on Research of Sentiment Analysis (2008)
[3] Wu, C.-H., Chuang, Z.-J., Lin, Y.-C.: Emoiton Recogniation from Text Using Sematic Lables and Separable Mixture Modles. ACM Transactions on Asian Language Information Processing, TALIP (2006)
[4] Matsumoto, K., Ren, F., Kuroiwa, S., Tsuchiya, S.: Emotion Estimation Algorithm Based on Interpersonal Emotion Included in Emotional Dialogue Sentences. In: Gelbukh, A., Kuri Morales, Á.F. (eds.) MICAI 2007. LNCS (LNAI), vol. 4827, pp. 1035–1045. Springer, Heidelberg (2007)
[5] Gill, J., Gergle, D., French, R.M., Oberlander, J.: Emotion rating from short blog text, Conference on Human Factors in Computing Sytem. In: Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing System (2008)
[6] Zhang, P.: Several Conciderations on the Theory of Large-scale Real Corpus (1999)
[7] Shen, Y., Tian, C.-G., Li, S.-C., Liu, S.-C.: The Grand Information Flows in Micro-blog. In: SEWM (2009)

A Modified Transport Protocol for Heterogeneous Wireless Sensor Networks

Wu Yangbo, Zou Donglan, and Li ShuLiang

Xinyu University, Jiangxi, 338000, China

Abstract. Wireless sensor networks(WSN) are a kind of communication networks with in-dependent sensor nodes which could construct multi-hop ad hoc network to trans-fer data. The transport protocols for Wireless Sensor Network play an important role to promote the performance of the whole systems. This paper proposes a modified transport layer protocol for heterogeneous Wireless Sensor Networks(HWSN). As TCP is not suitable to use in HWSN directly, we design a modified transport protocol(named mTCP) based on TCP by a six-state finite state machine. Experimental results illustrate that our design can improve the communication capability of the transmission layer in heterogeneous wireless sensor networks significantly.

Keywords: Heterogeneous wireless sensor networks, Transport protocol, TCP, Finite state machine.

1 Introduction

A wireless sensor network(WSN) is a self- configured network consisting of numbers of sensors deployed in a sensing field in an ad hoc or prearranged fashion for purposes of sensing, monitoring, or tracking environmental events. WSNs have been widely used in battle field surveillance, environmental monitoring, biological detection, home automation, industrial diagnostics, and so on[1]. A heterogeneous wireless sensor network(HWSN) is a kind of WSN but each sensor may have different capability, such as various transmission capability, different number of sensing units, and so on[2][3].

Transport protocol is very important in HWSN as it establishes end-to-end connections over the network, while offering services such as congestion control, flow control, fair allocation of bandwidth, reliability, packet-loss recovery, energy efficiency, and heterogeneous application support. Transport layer protocols are used to mitigate congestion and reduce packet loss, to provide fairness in bandwidth allocation, and to guarantee end-to-end reliability. However, UDP and TCP transport protocols, which are currently used for the Internet, cannot be directly implemented for WSNs.

For WSN, several various transport layer protocols have been designed to solve various issues such as congestion control and reliability. STCP[4] is a generic end to end upstream transport protocol. It provides both the congestion control and reliability by allocating most of the responsibility at the sink. CODA[5] is an energy-efficient

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 875–879.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

congestion control scheme that can quickly mitigate congestion which has three components: Congestion detection, hop-by-hop backpressure, and multi-source regulation. Congestion control and fairness (CCF)[6] infers congestion based on packet service time, which is a congestion control protocol that uses packet service time to detect congestion in the network.

2 HWSN Architecture Description

The design of sensor node is various in different applications, but their basic structure is the same. The typical hardware architecture is shown in Fig.1, mainly including the battery, the power source management, the sensor, the signal recuperating electric circuit, A/D transformers, the memory, the microprocessor and the radio module. Therefore in the node design stage, we must consider to save the power source. The design of each communication protocol layer should concentrate on energy conservation.[3]

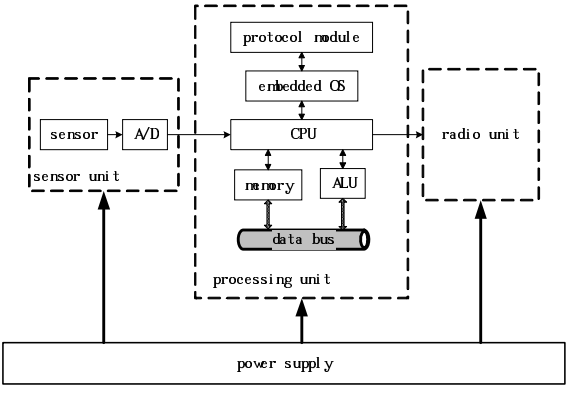


Fig. 1. Framework of WSN node

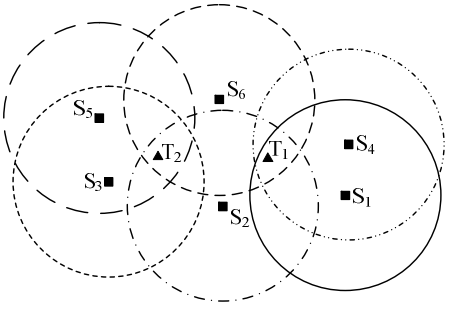


Fig. 2. An example of heterogeneous wireless sensor networks

The wireless sensor networks discussed in this paper are heterogeneous and equipped with multiple sensing units, which means each sensor in the HWSN may be equipped with more than one sensing unit and the attribute each sensing unit can sense may be different as well. As is shown Fig.2, an example is given with two targets and six sensors(with multiple sensing units) to illustrate the architecture of HWSN.

3 A Modified Transport Layer Protocol for HWSN

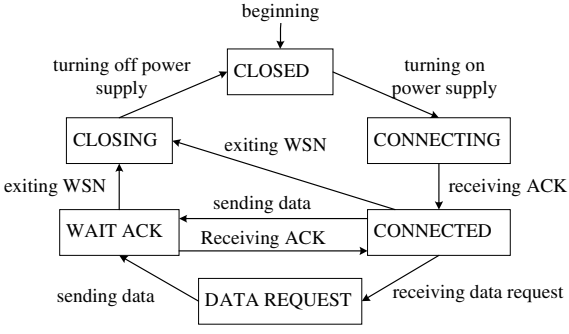


Fig. 3. Structure of mTCP finite state machine

In this paper we mainly consider the transport layer protocol in processing unit. The transport layer is in charge of data stream's transmission control, it can keep high communication quality. Howerer, WSN has limited resources, high error rate, dynamic changing topology. Hence, we modify transport layer protocol to make it suitable to communicate with standard TCP protocol directly. We name the transport layer protocol in HWSN as mTCP, which is an event-dirven protocol. Differing from the wired networks, the design of mTCP's finite state machine should be concentrated on things that nodes can join or exit HWSN dynamicly and freely.

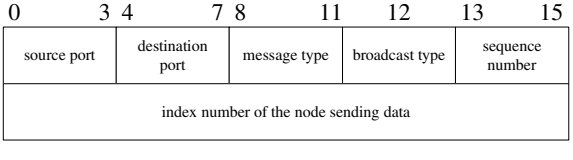


Fig. 4. The modified TCP protocol head

Being different with TCP's complex finite state machine, mTCP's finite state machine only includes six states which are CONNECTING, CONNECTED, DATA REQUEST, WAIT ACK, CLOSING and CLOSED(shown in Fig.3). As mTCP needn't consider reliability and the congestion control of TCP, we can design a simple protocol head, which is shown in Fig.4. mTCP segments are used mainly in transmitting data, and some of them are used for carrying ACK message.

4 Experiments

In the paper we choose OMNet++ as the simulator, which is designed by telecommunication institute of Budapest technology university. OMNet++ can be used as the simulator for communication protocol, distributed communication system, WSN and so on. We design two plans to test the transport layer protocol in HWSN. The simulation environment is a 100m×100m square region, the distance between two nodes is 15m.

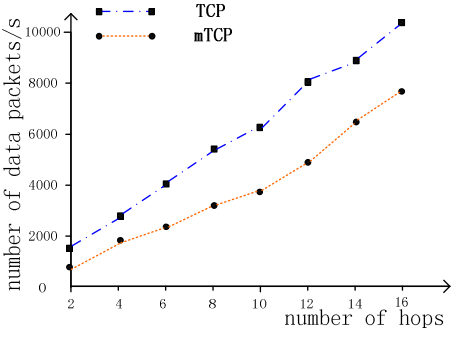


Fig. 5. Number of data packets-number of hops

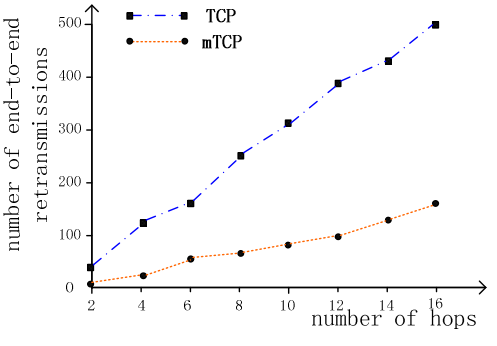


Fig. 6. Number of retransmissions with different hops

In this section, the main work is to test how does data packet quantity and end-to-end retransmission of HWSN change when the hops of the nodes changing.From simulation results(Fig.5 and Fig.6) TCP modified effectively reduces the data packet quantity and the end-to-end retransmission quantity with the node hop value changing. The conclusion is the quantity of end-to-end retransmissions and data packets reduce obviously with the mechanism of segment caching and segment retransmitting from vicinal node. At last, we can see TCP modified with the mechanism we designed is suitable to be used in HWSN.

5 Conclusions

In this paper, we propose a modified transport layer protocol for heterogeneous Wireless Sensor Networks. We modified the standard TCP protocol by a six-state finite state machine. We conduct the experiments in a simulation environment which is a 100m×100m square region, the distance between two nodes is 15m. Experimental results show that our design can improve the communication capability of the transmission layer in heterogeneous wireless sensor networks significantly.

References

[1] Akyildiz, I., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. Computer Networks 38(4), 393–422 (2002)
[2] Sanli, H.O., Poornachandran, R., Cam, H.: Collaborative two-level task scheduling for wireless sensor nodes with multiple sensing units. In: Proceedings of the 2nd Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON), pp. 350–361
[3] Shih, K.-P., Wang, S.-S., Yang, P.-H., Chang, C.-C.: CollECT: collaborative event detection and tracking in wireless heterogeneous sensor networks. In: Proceedings of the 11th IEEE Symposium on Computers and Communications (ISCC 2006), pp. 935–940 (2006)
[4] Iyer, Y.G., Gandham, S., Venkatesan, S.: STCP: a generic transport layer protocol for wireless sensor networks. In: ICCCN 2005, pp. 449–454 (2005)
[5] Wan, C.-Y., Eisenman, S.B., Campbell, A.T.: CODA: congestion detection and avoidance in sensor networks. In: Proceedings of ACM Sensys, pp. 266–279 (2003)
[6] Ee, C.-T., Bajcsy, R.: Congestion control and fairness for many-to-one routing in sensor networks. In: Proceedings of ACM Sensys, pp. 148–161 (2004)
[7] Shih, K.-P., Chen, H.-C., Chou, C.-M., Liu, B.-J.: On target coverage in wireless heterogeneous sensor networks with multiple sensing units. Journal of Network and Computer Applications 32(4), 866–877
[8] Donckers, L., Havinga, P.: Energy efficient TCP. In: AMOC 2002, Malaysia (May 2002)
[9] Dunkels, A.: Distributed TCP Caching for Wireless Sensor Networks. In: Proceedings of the Third Annual Mediterranean Ad Hoc Networking Workshop (June 2004)

Research on the Development and Application of Intelligent Network Teaching System

Li Fachun

Guangdong AIB Polytechnic College, Guangzhou, Guangdong, 510507

Abstract. Along with the rapid development and the widespread application of the computer technology and Internet technology, education and teaching have gradually entered the era of computer networks. The intelligent network teaching system has incomparable superiority of the traditional teaching system, which not only standardize the management of education and teaching, but also reduce the workload of the education and teaching. What's more, it can greatly improve the efficiency of teaching management. Research of the Intelligent Network Teaching System from the current and future development trend analysis and summary will explore the specific implementation steps and the application of this new type of intelligent network teaching system.

Keywords: Teaching system, artificial intelligence, network teaching, development and Application.

1 Introduction

With the rapid development of the extensive application of computer and Internet technology, the traditional mode of teaching was subversive." Online education " as a new teaching model in people's daily learning and teaching model does not require a fixed time and place restrictions and can fully meet the need of self - learning and study at any time.

According to the Ministry of Education' plans, intelligent network education system need to be planned, and to build their online teaching mode and system step-by-step. In September 1998, the Ministry of Education has planned to try the new network of education and teaching in four universities of Tsinghua University, Zhejiang University, Hunan University, Beijing University of Posts and Telecommunications, and immediately after the approval of the Renmin University of China, and Fudan University, including 27 colleges and universities online education teaching test. By 2005, there were 64 colleges to carry out the education and teaching of modern remote network. According to statistics by the end of 2006, all the experimental college culture through modern distance teaching system of the 300 million people, accounting for 12% of the number of college students.

The reason why the intelligent network teaching system can obtain the swift and violent development, because it can provide the rich teaching resources and the full independent study time in it for the student, and has manifested truly by the student teaching thought primarily. Therefore it can achieve the huge success.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 881–885.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 The Present Situation of Intelligent Network Teaching System's Research

The network classroom teaching must be a certain degree of artificial intelligence. Only by this way can provide a more humane teaching platform for different interests and different levels of students. Intelligent teaching system can also make learning assessment, which is more convenient to the students' self- assessment and stimulation.

2.1 Network Teaching System Research

Since the Ministry of Education started a pilot network teaching system to build the current development of China's online teaching system very quickly, there have formed a certain scale of users. However, this teaching system is still having many problems; these problems are mainly embodied in the following areas:

(1) Software system is in relatively low degree of intelligence.
Due to the limitations of computer hardware and software, network teaching system is unable to achieve the use of intelligent and in many cases still need to adjust the students' needs, and also still far from reaching the general standard for artificial intelligence.

2) The design of online teaching model is in the lack of scientific reasoning mechanisms and the use of research
In the absence of these two complete and detailed information, online teaching model is not able to accurately define the students' knowledge level and interests, and not be able to develop appropriate learning or assessment methods according to the different students of various hobbies, so in its development, there have many defects and have rooms for improvement.

(3) The lack of scientific and effective evaluation mechanisms
The process of learning must be a self-assessment and testing from time to time, and write a small summary of their own learning proficiency and progress, and do some simple forecasting and further planning for the further study. However, the detection and evaluation mechanisms of the current network teaching system is not perfect, therefore, to some degree, it hindered the advance of this teaching model.

2.2 Intelligent Tutoring System Research

This teaching system is very common in foreign countries, which the United States, Europe, Canada, Japan are the most outstanding countries in carrying out this. Intelligent teaching systems such as Autotutor intelligent teaching system involves not only liberal arts majors such as psychology, linguistics and pedagogy, but also involves computer science, engineering and physics, science and engineering professionals. The system is very simple in the item selection, curriculum description, the evaluation of students in collaborative problem solution, and populars in students.

We have many domestic universities and research institutes which do researches related to intelligent tutoring system, such as: the Institute of Computing Technology, Chinese Academy of Sciences Institute of Automation, Intelligent Technology and System State Key Laboratory, Nanjing University Artificial Intelligence Laboratory and so on. There are some typical intelligent tutoring systems in designing.

(1)" $Z+Z$ " Intelligent tutoring system.
The system was designed by the Chinese Academy of Sciences Professor Zhang Jingzhong R & D. the system has a smart solver, the basic functions of the human-computer interaction, intelligent reasoning, and dynamic mapping.

(2) Intelligent Chinese teaching system.
This kind of system take the single language as the main application scope's intelligent teaching software is developed by Zhejiang University's Chen Xiaogang, coordinated with Professor Chen Zengwu. The system has the certain extent speech analysis ability, may provide for Chinese learner a simple, relaxed, intelligent study platform. It can also give the analysis and the suggestion through the learners' learning effects. To some extent, it can help Chinese learners study independently in daily life and make a test all by themselves, and as well as the fast enhancement of academic record.

(3) The hypothesized study community and intelligence teaching network as a support platform of Capital Normal University
This intelligence study joint the workflow technology, cscw technology, Data mining technology, intelligent agent technology. It has solute some difficulties of decision-making and many kinds of teaching profession class, and provide a good support of self-learning.

Through the analysis of network teaching system and intelligent tutoring system, the following are mainly to research and analysis the intelligent network teaching system status and application. Research and application of the present stage of China's intelligent network teaching system is the following representative of the technology.

(1) " MAS " technology
" MAS ", short for "Multi - Agents System", is also an " Agent "technology, which originated in the field of artificial intelligence research, to a certain degree of intelligence, and able to independently run and provide the necessary services.

(2) " XML " technology
This Chinese name is " Extensible Markup Language ", " XML " first appeared in the daily use in 1998, and the first application in the field of electronic commerce. Gradually, it has expanded into other sectors. This technique is developed from the basis of the Html technology. It is easier to understand than " the Html, more convenient in data retrieval, data analysis and data exchange.

(3) The GRID (grid) technology
Many of the networks of educational resources were saved in document forms. Due to the diversity of educational resources, the management of the network of educational resources is an urgent need for a more simple and effective management model. the GRID (grid) technology is create on the basis of the method of solving such problems.

Currently used grid category includes two major categories of computing grid and data grid. Data grid is usually to meet data-intensive applications.

(4) NLP technology
“NLP " is a natural language processing technology. According to the degree of difficulty, “NLP "is divided into simple matching, fuzzy matching style and paragraph comprehension of several types. A simple matching is a keyword matching retrieval purposes the system ; The fuzzy matching is to match synonyms and antonyms ; Paragraphs understanding style by the same kind of understanding of the passages retrieval software system. Theoretically, this system is the best so far, can also be called "Full- Intelligent Tutor System.

3 The Specific Implementation of Intelligent Network Teaching System

Combined with the most popular idea of modular design, the preliminary intelligent network teaching system consists of five parts, as follows: the system of users’ register and login, teacher function system, consulting and BBS system, on-line examination system and background management system. The above five modules belong to subsystems of intelligence network teaching system, which follow the B/S model design, taking Microsoft Visual Studio 2005 as a basic language development tool.

The first step is to create user registration and landing subsystem.

This module is a indispensable system to facilitate the students to register and log on. We should design two different interface to provide registered and logged channels respectively for teachers and students, which will make management easier and provide good technical architecture for the following design.

The second step is to create teacher function subsystem.

In intelligent network teaching system, the function of teachers is very important, and a tearcher can provide students with online consultation, information upload and website maintenance.

The third step is to create consulting and BBS subsystem.

This module will convenient the communication between students and students, between students and teachers, between teachers and teachers by providing a virtual community, and convenient student learning to the greatest degree through timely online consulting or leaving a message for the solutions when students have problems on study.

The fourth step is to create on-line examination system

This module is the core part of the intelligence network teaching system. Students can create self-assessment and find their deficiency by online system in order to find out the direction of the study in the future. Students get the paper through the random topic selection system after logging on the on-line examination system, finish the paper in the stipulated time, timely view their scores, and correct the mistake which will help students improve their learning efficiency.

The fifth step is to construct background management subsystem

The background management module is designed special for system maintenance personnel, because a software in operation will meet many problems, which needs a special maintenance personnel to keep system maintenance from time to time. The management of the four subsystems, is to ensure the stability of the system operation.

The realization of the above steps is only the theoretical description for creating intelligent network education system. Intelligent network teaching system with its powerful function and the ability to work effectively, will be widely used in the future.

4 Summary

This article paid attention to the comprehensive analysis on the network teaching system and intelligent teaching system, and did research and comparison on several typical technology, and put forward the construction of intelligent network teaching system and the specific implementation measures, which provides certain theoretical support and technical basis for the construction of the system, and provides the basic research ideas for the future research on intelligence network teaching system.

References

[1] Urusilovsky, P.E.: LM-ART: An intelligent tutoring system on World Wide Web intelligent tutoring system. LNCS, pp. 82–85. Springer, Berlin (2003)

[2] Dorca, F.A., Lopes, C.R.: A Multiagent Architecture for Distance Education Systems. In: Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies, pp. 63–67 (2003)

[3] Chen, T., Zhang, J.: The research status and development of Intelligent teaching system (ITS) in China. China Audio-Visual Education Workers (2), 95–99 (2007)

[4] Xie, Z., Wang, L., Ge, Y.: Intelligent teaching system in cognitive type of student model. Computer Engineering and Applications (3), 229–232 (2005)

[5] Liu, Y., Huang, Y.: Intelligent teaching system knowledge model based on the Web of and control study. Computer Engineering and Design (14), 2640–2642 (2006)

[6] Zhen, Z.: The design and implementation of the frame model of intelligent network teaching based on Web. Modern Education Technology (3), 64–67 (2004)

[7] Luo, M., Wang, W.: Research and implementation of intelligent teaching system based on the Web. Miniature Computer Applications 28(09), 7–9 (2006)

Study on the Computer Networks System Safety Detecting Technology

Hu Liming

ChangSha Hunan Province, Hunan Vocational College of Science and Technology, 41004

Abstract. With the development of science technology, the computer networks systems has been widely used in all fields. However, threat to the security of computer networks systems has been a serious problem at present. So, this paper research and study on the computer networks systems from two main parts--main threat to the computer networks systems and computer networks systems security detecting technology.

Keywords: computer networks, safety technology, detecting.

1 Introduction

At the moment, computer networks came to thousands of families, being in close to our daily-life, people can look up information、 learn knowledge、 search jobs and shopping through computer nets. However, with the spread of the computer networks, security of the computer networks system has been threatened from many aspects, which bringing inconvenience to us. The vocabulary such as "Computer virus"、 "Trojan""Hacker", which threaten the security of computer networks system has lingered us recently, has brought about great economy damage and loss of important data to the people who use the computer networks systems. It would also lead to desperate danger to the county security and social life. To my knowledge from the previous news report, it was just one year in2005, there are 30 virus produced per day, as far as the whole world is concerned, we will find a vicious case that computer virus destroy the computer networks system per 20 seconds. Nowadays, except for traditional virus, the newly arising virus such as spy software, advertisement software, net fishing together with the traditional virus to break the security of the computer networks system, What is worse, the newly ones can cause grater damage than the conventional ones.

2 Main Threat to the Computer Networks System

With the development of the computer networks technology, it brings varieties of convenience to us, as well as kinds of potential danger. Because of the computer networks technology, there once a case that threat to the security of the computer

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 887–893.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

networks system caused by computer virus may transmit to another computer system and main engine, leading extensive computer networks system paralysis and economy loss, The computer networks system are facing following major threat:

2.1 Net Invasion

The so-called net invasion is that people gain the right to visit the information and files in the computer networks system which is unauthorized with illegal means by skilled compiling and debugging systematic procedures.Net invasion performs as a transfer station that get into other systems, and even destroy part of the computer networks system, it caused great threat to our life and work.

2.2 Trojan Horse

Trojan horse is such a matter that hide in the personal and entrepreneur computers, and steal information or the right to use the computers controlled by outer operators. Its full name is Teloy Trojan, English name is “Trojan horse”, the name came from Teloy Trojan, The jeopardize of Trojan horse lies in most of them are vicious, for example, occupying computer system resource, reducing the efficiency of the system, harming the safety of personal information (such as stealing the account of MSN, QQ, Games and bank cards) or take advantage of it as a tool to attack another equipment.

2.3 Worm Virus

Worm virus is a kind of common computer virus, It utilize the networks to reproduce and transmit. The networks and mails perform as the medium. The initial definition of the worm virus is: in the circumstance of DOS, there is a substance that is similar to a worm when the virus breaks out, swallowing the words casually and change them. The worm virus is a self-embodied procedure (or a series of process), which can transmit its copy or part of its functions to another computer systems (the conventional way is net joint).

2.4 DOS Attack

DOS is the abbreviation of the phrase “Denial Of Service”; The attack action which causing DOS is considered as DOS Attack (as the figure one shows). The objective of this action is to make the computers or networks can’t apply normal service. The representative approaches of DOS including: Ping of Death、Tear Drop、UDP flood、SYN flood、land attack、IP Spoofing DOS and so on.

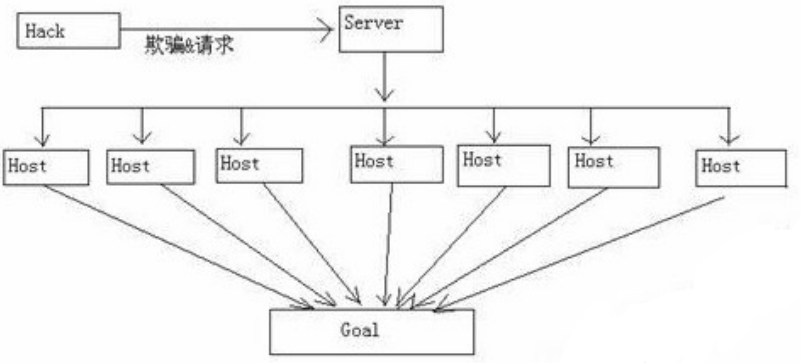


Fig. 1. DOS attack

2.5 Aiming at Attacking Personal Encrypted PC

On the contrary, security protection of personal privacy and data information has been one of the concerned problems by whole people, but some computer program elites usurp these enciphered material through process holes and endanger other's legal rights desperately. So far, the approaches hackers use including plaintext attack、cipher text only attack、chosen cipher text attack、chosen plaintext attack and violent attack.

2.6 Port Scanning

Port Scanning means that some people with special purpose send out a batch of port scanning information, trying to invade some certain individual computers and find out what computer service type the computer applies is (the type in relate to the data of the port). Port scanning (which could help the attackers find out the weak point) is the favorite approach among the computer decrypt essence. In fact, port scanning including sending information to each port, and a piece of information each time. The received response type indicates whether the port is in use and where is the weakness.

2.7 Junk Mails

Any electronic mails without the users' consent are sent to their mailbox in forcing manner is considered as junk mail. Junk mails are characterized by batch sending. The information of earning money、adult advertisement、commercial or individual website advertisement、electronic magazines etc are contained in the junk mails. Junk mails can be classified into two kinds——benign ones and vicious ones. The former kind are some propagating advertisement and have little influence to the receptor. The latter ones are often destructive. In order to spread information extensively, some junk mails sending organizations and illegal message senders often send out lots of mails at the same time through many machines to attack mail service, leading to great

broadband loss of mails service and disturb the normal mails sending progress desperately.

At the moment, many counties have legislated, trying to stop junk mails. Service policies of many networks service appliers contain oppose junk mails and they set up an electronic address for complaint. There are some networks groups, applying mails analysis and help their clients send information to corresponding ISP and make complaint.

3 Security Detecting Technology in Computer Networks System

The conventional concept of network safety is that in the computer networks system, hardware、software and systematic information are not infringed. The networks safety problem means that the change、delete and divulge of personal computer because of accident or vicious attack. The networks security problem has extreme threat on individual、enterprise and even on the country in the information times. Then, how to strengthen security detecting technology on computer networks system has been the key question in computer exploitation and application, the figure 2 displays the usual computer networks security structure.

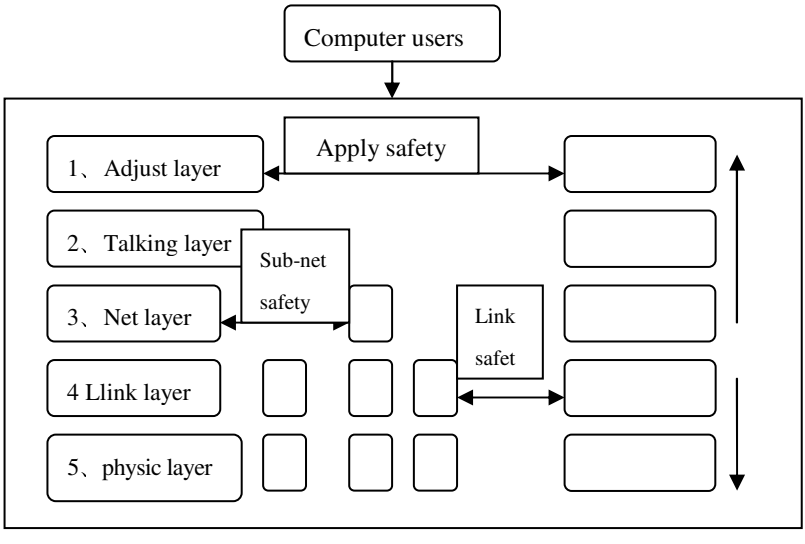


Fig. 2. Simplified structure of computer networks security system

Timely monitoring technology and safety scanning technology are main approaches in networks safety detecting technology. Timely monitoring technology means that the hardware or software check up networks data stream promptly, comparing it with the data that intrude into the system and do corresponding reaction according to the users as soon as the system is under attack. Such reactions may be cutting off the join or inform

the fire wall to adjust the visiting strategy, filtrating the invaded data pack. Security scanning technology (including long-distance safety scanning、fire wall scanning systematically、website scanning and systematic security scanning technology) could scan the safety weakness of partial networks、Websites、main engine operating system and fire wall system, finding out the holes and renovate them, reducing the security risk of the system.

Networks security detecting technology is based on self-adapt safe charge model. The model suggests: neither networks is able to prevent the potential security hazard. The model has two features: one is dynamic and self-adapt, this can be accomplished by escalating the networks scanning software and updating the invasion feature base in networks security monitoring. The other is the widespread of use, which can be used to detect leak of operating system、networks layer and application layer.

Many networks safety scanning software in the early phase are in directed against long-distance networks security scanning, such scanning software can detect and analyze the security holes in long-range lead plane. In fact, as these software can detect security holes remotely, it could also be efficient tool to attack another systems. The networks attackers utilized these software to scan the target computer, detecting the security weakness that can be taken advantage of, the first information through scanning will be the basis for farther attack. Such case still explain the importance that security detecting technology as for realizing the networks safety. The net manager could make use of scanning software to find out the networks holes and restore them before the attack begins and this could also improve the safety of the networks.

Utilizing the networks security detecting technology could achieve networks security detecting and timely attack distinguishes, but it is only important safety components of networks security, together with the fire wall could make up complete networks safety formula.

The basic principle of computer networks security scanning detecting technology is; The networks security detecting system detect the networks in main two parts——this machine and on the networks. The realization of the networks security detecting system is based on the thoughts of the hackers, searching the safety holes on the Internet by imitating attack. In addition, security detect system could also establish the fire wall, protecting the joints which are apt to be attacked easily, monitoring and following the tracks of the users recording in far place. In order to improve the flexibility, the networks security detecting system formulated following strategies: the analyzing part are not established beforehand, but generates automatically together with its corresponding regulation collection when the analyzing process is running. This regulation collection define the true action of each sub-process in standard pattern, as long as people master the normal writing style, simplify the form of the regulation and clear the meaning, any one could add their regulation and change the action of the analyzing sub-process, attaining the goal that control the“detecting——analyzing”procedure. The logical structure of networks security detecting system is shown as following figure 3.

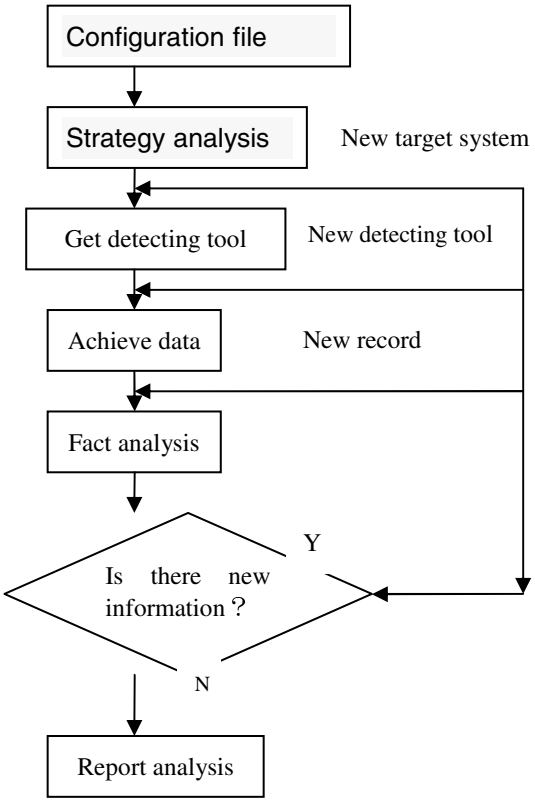


Fig. 3. Logical structure of networks security detecting system

4 Conclusion

With the development of computer networks technology, more attention has been given to the networks security problem, efficient detecting tools and approaches must be put into use, all safety technology should combine organically to guarantee the security of the computer networks system. This paper is based on the state situation the networks faced and spread research in the light of computer networks security detecting technology, studied on the mechanism of the present networks safety detecting technology, lay theoretical foundation and technology support for farther development of networks security detecting technology.

References

[1] Computer Security Institute. 2000 CSI/FBI Computer Crime and Security Survey. Computer Security Issues and Trends (5), 69–71 (2000)
[2] Marcham, T., Payne, C.: Security at the Network Edge. A Distributed Firewall Architecture. In: Presented at the DARPA Information Survivability Conference, vol. II, pp. 256–267 (2001)
[3] William, R., Cheswick: An Evening with Berferd, in which a Cracker is lured, endured and studied. In: Proceedings of the 1992 Winter USENIX Conference, vol. 6 (2001)
[4] Bloedorn, E.: Data Mining for Network Intrusion Detection: How to Started. MITRE Technical Report (8), 213–215 (2001)
[5] Wang, G., Jia, Z.: Research and design on network intrusion detection based on the firewall. Computer and Digital Engineering (5), 127–128 (2005)
[6] Yang, J.: Computer network security and countermeasures. Journal of Changsha Railway Institute (social science edition) (1), 236–237 (2005)
[7] Xu, R., Wu, H., Bi, X.: The key technology network of the information safety. Computer World (18), C7–C9 (2000)

Algorithm Research of Network Course Design Based on Personal Learning

Li Fachun

Guangdong AIB Polytechnic College, Guangzhou, Guangdong, 510507

Abstract. Constructing the network courses based on personalized learning is an important way to satisfy the learners' individual demands and improve network teaching qualities. Taking Applied Technology of Single Chip Processor as the example, this paper analyzes the present problems of network courses, describes the features of network courses on the basis of personalized learning and discusses the basic frames of network courses purposely realizing personal learning. In addition, it provides related algorithms using to design network courses on the foundation of personalized learning, regarding the Web mine technology and ontology technology as the crucial ones.

Keywords: Personal Learning, Network Course, Data Mining, Ontology.

1 Introduction

Network course is a summation of the teaching contents of a certain subject and the implementary education activities reflected by the net, including teaching contents corresponding to the specific instructional objectives and instructional strategies, and web supporting surroundings[1]. Network Course can fulfill advantages of the network teaching to a great extent, with such characteristics as openness, interaction, sharing, coordination and autonomy. At present, the network course has become a way of realizing educational informationization, attached great importance of various types and levels of educational institutions, and has been a major advance project of education teaching reform.

However, production levels of on-line learning are rather spotty. The author's investigation is for the national quality course, finding there are many insufficiencies and its design level has a tremendous promotion. The insufficiencies are reflected by the following: unreasonable organization teaching resources; unchangeable and not timely update of course contents; unmeaning interface results in the difficulty to spur learner' s hobby for studying; uniform course contents, guidances and learning methods to the whole students and the lack of personalized learning. In view of certain disadvantages of the network course construction, the technology needs some improvements. This paper carries out some researches of network courses based on personal learning and puts forward some basic approaches of the realization of personal learning.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 895–903.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Characteristics of the Network Course Based on Personal Learning

The definition of personal learning is adopting proper methods, contents, starting point and assessment way to better develop their independences, reflections, initiatives and creativities, targeting scholar personality traits. By comparison, the so-called traditional personalized study, means teaching plan making varies with the different learners' characteristics. Completing, the teaching activities are always held in fixed classrooms. With respect to the large class with numerous participants and limited condition, the plan can not be carried out well even offered a diversity of learning tasks by teachers[2]. The network courses based on personal learning is a kind of network teaching resources, whose ultimate purpose is to satisfy the learners' personal learning and whose value pursuit is the realization of personality development of the learners. With the network, the learners can be easy to locate learning environment coherent with his or her learning, learning style, learning ability, interest and other personality traits, rather than the stereotyped pattern. In this way, effective learning can be easier achieved due to the dynamic and open system. It belongs to a learning pattern based on network course and adjusts itself according to its circumstance.

Individualized network course also has the following properties:

2.1 Individual Learning Navigation

The traditional network course has equal treatment for all the learners, not only the study contents and forms, but learn order, fixed and can not be changed in the development phase, which making learners lost themselves. However, the individual network course provides individual navigation (including the AutoContent Wizard and the user-interface (UI)), to adapt to the learners' learning situation and their needs.

2.2 Individual Learning Methods

With controlling learners' learning methods, individual network course provides automatic adapted learning environment after analysing learners' knowledge levels and previous learning information, making the personal learning come true according to the learning methods produced by dynamic generation.

2.3 Individual Learning Contents

Individual network course first divides learning contents into several parts (named unit), then combines them with the hyperlink.. As a consequence of this, the learning resources of these parts vary according to the learners' learning background, psychological characteristics, learning styles, learning abilitites, interest and other personal requirements instead of the prevenient assured order and pattern.

2.4 Individual Study Support Service

Except for E_mail for one-to-one communication, an online same support with personalized characteristic should be created to ensure individualized support services, which may make learners participate in with such tools as BBS and learning forum. Simultaneous, it can render services for the learning evaluation and feedback on each phase[3].

3 Key Technology of Individual On-Line Course

A lot of technologies are involved for individual network course exploitations, the most crucial two of which are Web data mining technique and ontology technology.

3.1 Web Data Mining Technique

Web data mining is such a technology that links the traditional data mining technology to the Web, targeting Web page, the Web page structure, users' remaining information and other Web data, and extracts the abstract, potential, useful data with the help of data mining methods. Much information will be left after being visited by learners, which can be got by Web data and personalized network courses will be built according to the mining results.

3.2 Ontology Technology

Ontology is a definite and formal specification of sharing conceptual model. It acts as the rudiments of resource description in an area and guides the users to describe their retrieval requirements accurately and completely[4]. It presents the domain knowledge in a definite and formal way to promote knowledge sharing. Setting up a common understanding of information between computer and the human being to help the directly utilization, so that personal study methods adapted to the learner can be generated intelligently, in accordance with the users' characteristics[5]. Ontology technology can be used to help design knowledge ontology and generate network course in line with personal learning, all of which are built on the basis of the learners' knowledge backgrounds, learning abilities and learning characteristics.

4 The Realization of Individual On-Line Course

A personal network course of "Applied Technology of Single Chip Processor" is established by the author relying on Web data mining technique and ontology technology. The core module of that is shown Figure 1.

By making use of the generation of ontology being edited by the Protégé, an ontology generation module will be read and displayed by the Java and a dynamic tree-shaped navigation structure of knowledge will be produced; The learner generator

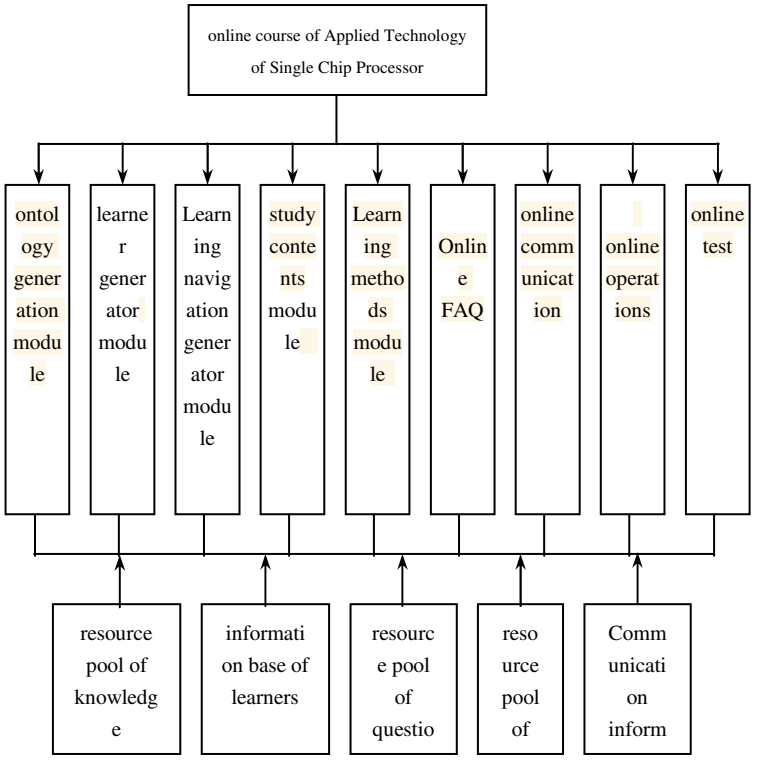


Fig. 1. A schematic drawing that illustrates the structure of online courses

module will classify the learners by the use of cluster analysis so that correct navigations and study contents would be provided according to the learners types; Learning navigation generator module will present relative learning navigation in line with the learners types (including Guide Menu and its Order); The provided module of study contents uses plentiful association rules to choose and to recommend suitable resources for learners; Learning methods module generates learning sequences according to logical relationship and the knowledge mastered by the learners. These modules such as online FAQs, online communications, online operations and online tests are mature and are easy to understand. I need not repeat them one by one here.

4.1 The Establishment of Course Knowledge Ontology by Using Protege

Course knowledge ontology is composed of the course concepts, their relations and formalism descriptions. It can be expressed by a triples[6]: O=<T, P, R>

The O stands for ontology of course knowledge, the T means its concept set, the P is its attribute collection and the R represents its relationship setting up on the foundation of T, mainly on behalf of relationship of various concepts.

The algorithm steps of creating course knowledge ontology:

4.1.1 The Choose of Applied Purpose and Range of the Ontology

The establishment of related domain ontology or the process ontology should be based on the domain or task you researched. The larger the domain is, the larger the ontology will be. That's why the analysis of ontological requirement is needed before establishing ontology. At the same time, the analysis can help clear purpose, scope, application, and users.

4.1.2 Establishment of Important Concept Set of Ontology

The first thing is to determine the key concept of the course if you want to construct a course ontology. The teaching procedure and teaching regular patterns can be the reference substance to assist the analysis of course content. You can regard the most important and the most common concepts as the crucial parts and construct concept sets covering the whole concepts of this course. For example, the following 15 core concepts are defined in Applied Technology of Single Chip Processor: hardware structure of MCS-51 single-chip, C51 data type, C51 statement, C51 function, Keil C development environment, Proteus software, timer/counters, interrupt system, serial port, parallel interface technology, memory expansion, display connector, keyboard interface technology, A/D Interface Techniques and D/A Interface Techniques.

4.1.3 The Establishment of Concept Structure Relationship

After defining the core concept, a whole concept model in line with it should be set up from above to below, reflecting the paternity among the concepts. For the convenience of restructuring and personal learning of the knowledge, we have to do the following things: to analyse and decompose the knowledge itself, and to define their relationship, which including is-prior and is-next, without is-a and instance-of.

4.1.4 Properties of Definition

Concept structure relationship just refers to the basic framework of ontology, a definition of the conceptual attribute can not be overlooked. Attribute is the description of conceptual characteristics or nature, which can be divided into two parts: static properties and dynamic one. The static attribute means the descriptions of the concepts, characteristics, without behavior support; while the dynamic one is the dynamic characteristics of definitions, operations, processes and the call. So we should specify the field of definitions and rang of the attribute when you are defining the attribute.

4.1.5 Instance Creation

An instance is an essential element of a certain concept, that is, conceptual specific entity. The instance is also the information showed to the user.

4.1.6 The Realization of the Ontology

Express language with formalization means changing the defined ontology from the natural language form to the one that can be understood by the machine. After this

process, we can use computer to store, process and utilization. At present, the formal description language of ontology is teeming, such as RDF and RDFS, OIL, DAML, OWL, KIF, SHOE, etc. In this article, we take the RDF/OWL language as ontology description language. RDF (Resource Description Framework) is a model through the "properties-value" describing resources, used for representing information resources; while OWL(Web Ontology Language)is a semantic Web ontology language, using to meet the needs of the language of Web ontology, to realize the llogic and to prove its function. In order to improve the efficiency of development, you can use Protege, an ontology editor tool, to design course knowledge ontology, developed by Stanford University. With the help of Protege Owl, Owl ontology development will be much easier.

Ontology can not be built well once, so the continuously revises vary with specific applications. That's why we say ontology creation is a constant repeat and circulation process. "Applied Technology of Single Chip Processor" network course first use Protege tool to develop the central concept and their decomposition to produce OWL files, then read them by Java and finally generate tree-shaped navigation structures of knowledge.

4.2 Construct Personal Learners Model with the Help of Cluster Analysis

Individual network course respects personality, and personal learning style and learning habit and presents these characteristics in the network course. First of all, individual network course picks up personality traits from databases such as learning preference, knowledge degree and previous learning information, then divides the learners into groups by the data mining to build personalized model of learners. Among them the database includes the static information and dynamic one. The static information is acquired by their registration information and the dynamic one is the result of following learners' tracks and recording information related to learning. The frequency of learners visiting network course, the address and content, the residence time, the frequency of visiting, task completion and knowledge test condition are the aspects.

The steps of K-means algorithm, a way of constructing personal learners model using cluster analysis:

(1) Data Collection: The collection includes the following points: registration information, visited pages and retention time and mastering degree. And pick up some specific statistical magnitude to measure the similarities of different students and form a data set.

(2) The Initial Clustering Center: Based on statistics, randomly selecting appropriate amount of data sample as the center of the initial clustering state is the case.

(3) Classification: Selecting anyone from the data set, then counting the distance between it and the center, and finally classifying it into the the nearest clustering center. In this way, all the learners can be classified into a certain group and those with high similarities can be found in a fixed place.

4) Rearranged Clustering Center: Amending the equalizing value of all the data sample of each cluster computer and all the clustering center value, which is implemented in line with the results of the last step.

The steps of the third and fourth won't stop until the whole clustering keep still. Only in this way can the learners remain his own clustering.

4.3 Using Interpretive Structural Model to Produce the Personal Learning Method

With the defects of superabundant and complex learning resources, the traditional network courses always make the learners lost himself. Comparatively speaking, the ISM (Interpretative Structural Modeling Method) is much better. It can generate personalized learning methods to control the learners and provide an automatic adjustment for the learning environment.

Interpretive structural model is an effective way of analysing and hinting complex relationship. It decomposes the complex relation of different component elements into multilevel hierarchic structures, which are clear[7]. Based on ontology and learners' cognitive structure, we can use interpretive structural model to generate methods related to learners personalized learning and guide the effective learning.

The algorithmic steps are listed here:

4.3.1 Adjacency Matrix of Forming Relation among Knowledge Points

It refers to the knowledge points of c_i and c_j of the ontology. If the c_j is the must grasp before learning c_i , we can conclude that there is a "sequence relation" between c_j and c_i. The adjacency matrix A of direct relation is founded on the "sequence relation". $A= (a_{ij})_{n\times n,}$ when the "sequence relation" exists, $a_{ij=}1$; otherwise, $a_{ij}=0$.

4.3.2 The Realization of Reachability Matrix

Reachability matrix is represented with P, $P= (p_{ij})_{n\times n..}$When there is an access between c_i and c_i, $p_{ij}=1$,or , it equals to 0. As to adjacent matrix of n-order A and identity matrix of n-order E, the Boolean operation should be used in line with P= (A+E) k-1. Only when the $(A+E)^{n-1}\neq(A+E)^n=(A+E)^{n+1}$ holds, can the reachability matrix P be calculated.

4.3.3 Defining Matrix of Learners' Learning Background

Assumed knowledge point set K={c1,c2,…,cn}, matrix of learning background $B= (b_{ij})_{n\times n,}$ the definition of B can be summarized as follows: When i=j, and the knowledge point has a long distance from his objectives, $b_{ij}=1$, or $_{bij}=0$. Obviously, for those who are not ready for learning, the b_{ij} has only one value , that is 1.

After finishing learning each knowledge, the students can use online tests to check himself or herself. The tests often show the knowledge that have been mastered by learners. In this situations, b_{ij} is equal to 0. Thus, a matrix of learners' learning background is shaped.

4.3.4 Relational Matrix of Calculation

For matrix of learning background B and reachability matrix P, using R=BP to calculate, structural matrix R can be got, which is available to the learners' personal levels.

4.3.5 The Formation of Leaning Sequence

Count structural matrix R, gradually determine the hierarchies of different knowledge point and bring about personal learning methods.

4.4 Using Association Rules to Excavate and Provide Individual Learning Content

4.4.1 Recommend Learning Materials for Learners with the Help of Association Rules

In individual network course, using association rules to construct a conversation page set and to dig the potential relation of these pages. We can limit according to the pre-setting support and confidence coefficient to recommend learning materials in line with his personality traits. For example, the users visit both page A and page B concluding such anassociation rule：if the user visit page A, he will be likely to visit page B. This means that the learner will recommend the learning resources on page B when he is browsing page A. In addition, a prediction for the learners' demands will be made with the help of association rules to provide a possible page.

4.4.2 Compounding Learning Materials Dynamically with the Help of Association Rules

Using association rules, those pages often skimmed by the learners together can be found, without sequence relation. After association analysis, the frequent visitor can be found. According this, we can arrange learners, especially those of the same interest and needs, which is satisfied by linking frequent resources together.

In short, making use of association rules can recommend personal learning contents for learners and the concrete algorithm of which can be seen in reference part[8].

5 Conclusion

This paper mainly discusses something about individual network course. For the course of Applied Technology of Single Chip Processor, the major contents are listed as follows:

Providing the algorithms to design relative module with the help of Web data mining technique and ontology technology.

Designe network courses based on personal learning, offering personal learning environment for Web learning.

The author believes that individual network course will spark an new revolution in design and development of the network course, which is the only road for the the development of network course.

References

[1] Zhou, C., Yan, H.: Foundation of Network Course Manufacture, vol. (10), p. 7. Huazhong Institute Press, Wu Han (2006)
[2] Wang, R.: Educational Innovation: The Research of Personal Learning Based on 1:1 Digital Environment. Software Guide·Pedagogical Techniques 2(11), 52–54 (2009)
[3] Huang, F.: Development Mode and Strategy Research of Network Course Based on Personal Learning. Education Research (10), 77–81 (2009)
[4] Cheng, N.: Development System of Personal Network Course Based on SCORM. Computer Systems & Applications (9), 13–17 (2009)
[5] Wang, J., Qiao, H.: Application and Research of Integrational Virtual environment Based on Ontology Technology. Shandong Science (4), 47–52 (2009)
[6] Liu, G.: Construction and Realization of Ontology in E_Learning System. Journal of the China Society for Scientific and Technical Information (8), 499–507 (2009)
[7] Wang, X., Liu, Y.: Using Interpretive Structural Model to Analyse Mechanism of Transformation of Subjective Implicit Knowledge. Information Science (3), 475–480 (2006)
[8] Gao, L.: Research of Suggested Module of Personal Learning. Heilongjiang Researches on Higher Education (3), 135–137 (2008)

Observation on the Wireless Sensor Router Algorithm Based on Mobile Agent

Pei Fang and Zhang Jie

Hunan Mechanical & Electrical Polytechnic, Changsha, Hunan 410151, China

Abstract. This paper begins with the basic theoretical concepts of mobile agent and wireless sensor network, gives some views on the node access path algorithm of the wireless sensor, wishes that it will benefit the later design.

Keywords: Wireless sensor, Mobile agent, Algorithm.

1 Introduction

Wireless Sensor Network (WSN) is the focus which attains particular attention in the modern information times and attracts lots of exerts and scholars all over the world. It is a entirety that is integrated by sensors, embed calculation、 information dispersed operation and wireless communication technology, it is a new technology for information operation transmission and is widely used in the economy、military、medical and industrial fields, having achieved satisfying progress, so it gains more and more attention of the humans. It is no doubt that the development of WSN will bring profound significance to our daily-life and promote social development to a new stage. However, with the widely application of the WSN, there are more and more node the presented router connect exchange technology seems exhausted to handle the WSN, so, new algorithm has been put on the schedule.

2 Basic Theory of Mobile Agent and Wireless Sensor Network

2.1 Theory of Mobile Agent

Mobile agent is a computer procedure which is used to help the connected users accomplish their certain tasks. With the abundant resource and casual visit right of the mobile network, the procedure could settle the data resource with the most proper resource with the cheapest way with particular algorithm to help the consumers complete their works. Whether the technology is excellent or not is relate to the network situation、communication agreement、operating system of the service and the automatic level, if one wants to own a perfect mobile agent, following problems must be settled:

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 905–909.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

(1) Agent execution, mobile agent is a process that can run automatically, it will search for needed information in other network circumstance through wireless network, then collects information and calculate it. The essence of the program is that the procedure looking for a proper place, then moves as a entirety to the found place, thus, a problem is involved——there must be a platform that can transfer to ensure the procedure can work successfully after the mobile agent process moved.

(2) Migration, the operation process of mobile agent system requires the mobile agent can run smoothly in various systems which ask for the migration criterion can adapt different systems to achieve good harmony. The migration contains three steps: code migration、data migration and executive circumstance migration, obviously, the code means the procedure code of the mobile agent system which is utilized to finish particular task; Data indicates the needed data parameter in the executive program; Executive circumstance refers the needed platform to guarantee the code could work successfully. The move of the mobile agent system includes two types——powerful move and weak migration, the former kind includes the three ones above, however, only the first two are contained in weak migration.

(3) Communication language means the special linguistic form in mobile agents, in contains the communication language between same agents in different systems and the language in different agents in the same system. The information of mobile agent language contains resource information、algorithm share and so on. The presented communication languages incorporate RPC、RMI and agent communication language, the last one is advanced, it helps communicate in different mobile agents and has nothing to do with the open communication because it has no constraints for communication.

(4) The model the WSN adopts is called the distributed calculation model, different with the traditional models, it has joints handling which is useful to deal with information promptly. The mobile agent referred above is a calculation procedure that moves on the network and this is a complement to the multi node characteristics, so lots of mobile agents are used in wireless sensors, in addition, mobile agent is characterized by low energy consumption、allow balanced load、high reliability and calculation accuracy and so on. The application of mobile agent wireless sensor network makes them complement each other to achieve more efficiency.

2.2 Theory of Wireless Sensor Network

WSN is the abbreviation of Wireless Sensor Network, it is a new technology that is widely used in social produce such as information resource store、gaining and handling field, it incorporate many advanced technologies such as sensors、connect calculation、network communication and data settle, it handles the data information by distributed calculation. Taking advantage of wireless network technology makes the handling of the information will not be constrained by time and place, it is a open free state. In addition, the cost of making up the wireless sensor network is low and low energy consumption、simple system structure and easy to master are also the merits of the WSN, then it is widely used in military、industrial produce、traffic manager,

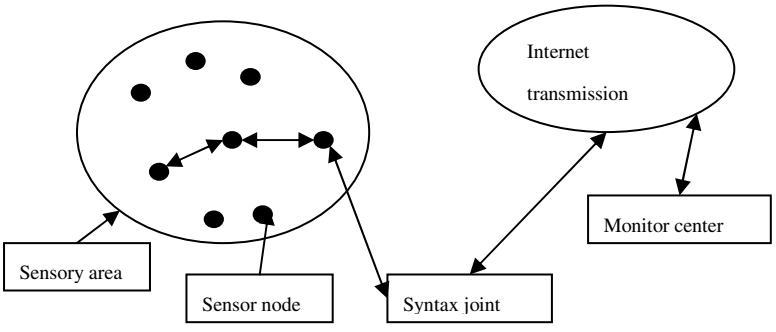


Fig. 1. System structure of wireless sensor network

medical and many other fields. Conventional WSN is made up by four parts——sensor node、cluster joints、data transmission network and long-distance monitor center, its detail structure is shown on fig 1:

WSN is a kind of network that is designed for special demand, it has much unique compared with other common network, its characteristics are displayed as follows:

(1) Large scale: The object of WSN is to coat all the territory by distributed nodes, so, it has special requirements for the sensor joints. The sensor joints must be on large scale, each of the sensor has limited power to handle information, then, in order to gain satisfying outcome, distributed density of the sensors must be improved through which the accuracy of the information and handle speed for message can be assured.

(2) Open dynamic structure: The characteristic of WSN referred above is large scale, but it is difficult to ensure that the property will meet our requirements all the time after the network is established. Even the strength is kept same, the sensor joints may not be perform well forever, troubles may not be avoided, now, sensor joints need to be replaced and even change sensor joints, for the sake of the function of the whole system, the network structure is required to be able to adapt change and complement dynamically.

(3) Self-adapt network: In order to ensure the whole-coat of WSN, a large number of sensor joints need to be installed, some of the joints are not easy to be safeguarded for which reason people can't make up network work between every two joints. Therefore, some joints are not used often are needed, the sensors are needed to adapt automatically, making up the network with the round nodes automatically to finish transmission of data, its essence presented are dynamic.

(4) Reliability: For the whole-cover of the WSN or the space that people can't reach, we should collect the data and calculate analysis all the time, the conventional means is dropping the sensor joints to the target territory by helicopter to finish task. Most of the territory is in bad environment, so, maintenance and repair for the nodes seems impossible, however, the work must be completed successfully, then the sensor joints are required with high reliability.

(5) Energy restrained: Most of the power that the joints use in WSN is provided by batteries, a problem occurs——the electricity will be run out sooner or later. During the applying period, the electricity becomes lesser and lesser, so does the emit strength, for the sake of the transmission of information, the distance between the joints must be controlled. Therefore, energy supply strains the development of WSN.

3 Observation on the Wireless Sensor Router Algorithm That Is Based on Mobile Agent

The object of the research on the router algorithm of WSN is to find a better visit path to the nodes, not only the exact and rapid transmission of information can be ensured, but also the transmission distance is the shortest and cost less time, only in this way, the energy consumption of the network can be reduced. The are three main algorithm for the joints visiting path of WSN:

3.1 Static State Model

In the static model, visit order and joints were preset, they will not change automatically, in thin model, there are two methods in joints path algorithm:

Elicitation method: it is easy, it chooses a sensor node as the connected joint, then gives axle around the joint and look for the nearest sensor joint round it as the next joint, by analogy, the procedure stops when the target joint is found. The approach is the most popular when there is only one visit path, however, it is not the best when put into the whole field network transmission, so, in practical application, it is used to compare with another designed path to judge whether another is excellent or not.

Inherited method: The inherited method takes the whole situation into account, it is a modified visiting rout compared with the elicitation method, it collects the information of all the sensor joints in the network and calculates the visiting order of the joints in this method. However, there is a problem in the algorithm——if the nodes are in trouble, the initial rout will be broken, then another information collection of the joints is needed which will reduce the efficiency of the network as for the complex system.

3.2 Dynamic Model

The access order in this model is not unchangeable. It collects the information of all the joints in whole network, then calculates it according to the algorithm in the elicitation method, achieving the best access path. If some joints breakdown in the operation process and leads to the information can't be transferred, then dynamic algorithm is applied, it will provided the second best visit path to ensure the network run successfully. The calculated access order of the joints in dynamic method is based on the energy consumption and time in the data resource transfer process to make sure the rout save more energy and time.

3.3 Mixed Model

It is easy to understand this model, both elicitation method and dynamic method are included in this model. In the model, syntax of the nodes is calculated in elicitation method, the access path between the joints are handled with dynamic algorithm.

4 Conclusion

The wireless sensor network are used extensively and become more and more important in social produce. This review researched the wireless network that is based on mobile agent, starts with basic theories of the wireless agent and wireless sensor network.

References

[1] Manges, W.I.T.: Time for Sensors to Go Wireless. Sensors Magazine 32(5), 4–5 (1999)
[2] Kahn, J.M., Katz, R.H., Pister, K.S.J.: Next Century Challenges: Mobile networking for "Smart Dust". In: ACM International Conference on Mobile
[3] Garten, J.E.: Ten Emerging Technologies that will Change the World. Technology Review 106(1), 22–49 (2003)
[4] Akyildiz, I.F., Weilian, S., Sankarasubramaniam, Y., et al.: A survey on sensor networks. IEEE Communications Magazine 40(8), 102–114 (2002)
[5] Wang, C., Luo, H.: SPIN protocol for wireless sensor networks. Journal of Shiyan Technical Institute 17(2), 60–63 (2004)
[6] Peng, G., Liu, R., Wang, F., et al.: Overview on research of wireless sensor networks. Journal of Guangxi Academy of Sciences 23(4), 369–372 (2007)

Higher Vocational College Computer Culture Foundation Online Examination System Design

Yonghua Zhang, Lizhe Wang, and Ji Li

Changchun Vocational Institute of Technology, Changchun, Jilin, 130033, China

Abstract. This study takes the Changchun Vocational Institute of Technology" Computer Culture Basis" course of the online examination system that is based on ASP.NET technology as an example to introduce the design and realization of online examination system development process. In the Windows environment using ASP.NET and COM programming, designing" Computer Culture Basis" course in Higher Vocational Colleges of paperless examination and scoring system.

Keywords: Fundamentals of Computer Culture, COM Technology, Automatic marking, Markup language, On-line examination.

1 Introduction

As everyone knows, in the traditional way of examination, an examination contains at least three links, first before attending the examination, teachers design test papers and prints them. Second is to organize students' examination link, then evaluation links including teachers marking, assessing and test results analysis. Obviously, with the increasing types of tests and examination requirements, the workload of teachers will be more and more heavy, the work in the traditional examination becomes more and more complicated and more error-prone and is not adapt to the information age examination in the assessment of needs. Convenient paperless online examinations to evaluate alternative to the traditional examination method is the reform of examination approaches and a grand goal, the former involves examination of the various operations with the aid of the computer network platform to complete, flexible, efficient management, in order to make the test process become convenient, fast、 efficient、 fair、 is the subject of the study objective.

In order to improve the students' employment and social adaptation ability, higher occupation college education has already put forward by the examination-oriented education to quality education, and the present traditional examination system still has the drawbacks of exam-oriented education. The reform of examination method is to realize the real meaning of quality education, highlight the practice ability cultivation and evaluation, can effectively promote the teaching content, teaching method reform.

The twenty-first century is the information age of network, with social networking、 informatization and modern trends continue to advance, information technology as the

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 911–916.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

main indicator of the science and technology has become one of the teaching contents in school education that cannot be ignored. Computer culture foundation is a public foundation courses in higher vocational colleges for non-computer specialty students, for the modern college students, this course will undertake to popularize the basic computer knowledge, cultivating students learning and mastering the ability of acquiring information、 information processing task, directly affect students' information consciousness, the ability to control information. So, with the help of computer and network technology, the full realization of the examination paperless、 network、 automation、 and can fully with the traditional paper advantages, has far-reaching realistic meaning and practical value.

2 System Requirements and Technical Analysis

2.1 System Demand Analysis

The computer online examination system we developed is mainly to inspect the students to grasp the basic course of computer culture, including the basic computer knowledge, basic Word operation, basic Excel form processing operation content. After investigation and analysis, online examination system should solve the following problems:

1. Complete paperless 、 no floppy online examinations. The test results are presented directly and stored to the server, the examination paper of examinee is automatically downloaded and displays on the examinee's machine;

2. Easier supervision. Since the examination system rejects other candidates to register for the examination with the same number, candidates can guarantee uniqueness; And when the time had run out, the system will automatically reject candidates to register for examination, the effectiveness of the examination is guaranteed. So the problems in traditional exam that teachers need strict control become easy in online examination system;

3. Timing control to hand in the papers. As long as the time of candidates runs out, candidates will not continue to answer, the system will prompt examinee to save all the answers, and the results are uploaded to the server, realizes real-time hand in.

4. Personalized examination presented in automatic test paper. Online examination system will generate different examination paper to each candidate individually, so students in their machines to see the roll surface is different, the cheat probability is too low. If the students understand the examination system, students will cancel the idea of cheating before examination, it can even promote students to review and prepare comprehensively and fully, teaching effect will be greatly improved.

5. Convenient database management that is easy to update and maintain. For the types involved in question database, through the database they can be maintained conveniently, knowledge quantity and content can also be maintained conveniently,

effectively control the paper types and volumes of candidates during the examination process.

6. It is convenient and quick to inquiry achievement. After the online marking work is ended, examinee can check their grades respectively according to the login name and student ID.

2.2 Technical Analysis

ASP in abbreviation of Active Server Pages, is a spool support Web system, provides good programming environment, using the web server to complete all script processing, can produce and run a dynamic、interactive、high performance WEB service application program, is dynamic webpage language developed by Microsoft, and it has inherited the tradition of Microsoft products-- server products can only run on Microsoft, running in the IIS (Internet Information Server) (windows NT) of the server side script execution environment, flexible、simple programming. When operating users are scanning the webpage, they can operate according to their actual needs; the site will produce the corresponding operation in response to the user, and will run the results back to the user's browser. The Unix also has Chili Soft plugins to support ASP, but ASP's function is limited by itself, combination of ASP+COM must be realized to expand and it is very difficult under Unix COM environment. It is scripting languages that used in ASP, ASP can build powerful applications, its all codes are embedded into HTML code, so when the complex functional webpage is compiled, it may cause the problem that the program code readability is low. ASP.NET is a new technology of Web design of Microsoft and higher efficiency compared with ASP. It is programming framework that is built on the common language, reusability is very high, a much smaller code amount to realize the same function compared with ASP. In addition, ASP.NET adopted new programming environment, representing the major direction of technical development. In security aspect, based on the windows authentication technology and each application configuration, the ASP.NET has been removed injection vulnerability in SQL, enhance safety performance of the examination system we are about to develop.

3 Overall Design of the System

The project is to accomplish a paperless examination system on computer network that is used for the theory of knowledge and computer operation test for Higher Vocational "Computer Culture Basis" course. Based on this goal, consult and learn some relevant information and some relevant auxiliary examination software made by commercial companies before designing and analyzing the function of the test system. The design thought this study put forward was shown in Figure 1:

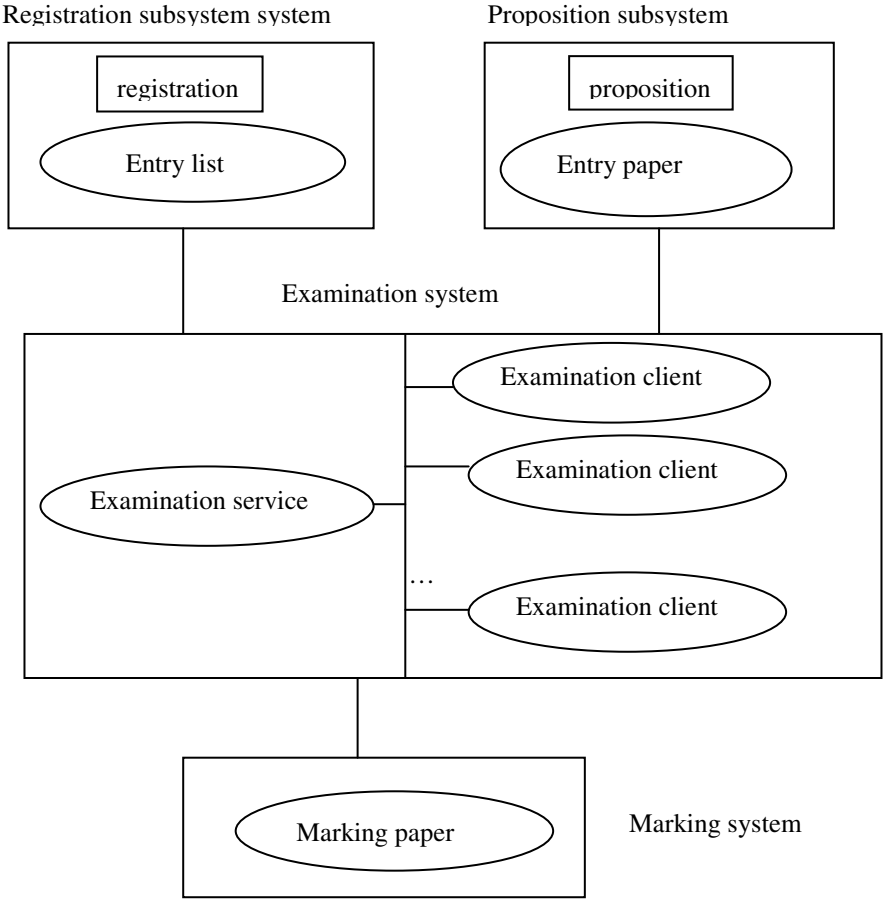


Fig. 1. Framework of examination system

Based on the above analysis and specific requirements of online examination, we need to design and accomplish such a test management system: the system structure is designed to contain registration subsystem, proposition system, examination system, marking system four components. The registration subsystem provides examinee data file, proposition subsystem generates the database file, examination system to generate candidate paper documents, marking system to generate results database.

4 The Systematic Function Module

The test system is mainly used in the computer test, as described previously, it is decomposed into four subsystems according to its function: registration system, the proposition system, examination system, marking system four parts. The function of each subsystem is as follows:

Registration system: its function is to complete the examinee registration work, at the same time the examinee register the basic information, examinee complete the online information input work.

Proposition system: complete the library establishment of the examination system, manage all of the data, provide the function to establish the database and proposition. In the present examination system there are two types(multiple-choice questions and operation) operation problem set for the two general questions, respectively Word and Excel.

Examination system: when the exam invigilator program will be responsible for the candidates, enter the name and student number input candidate registration database to identify the information of the examinee in validation, and then take papers from a server randomly, this exam interface takes effect prompt and the system will start time automatically, examination time ends or when the candidate click to submit papers, the program will submit the candidates' answer to the examination to the test server and stored in the corresponding database.

Marking system: the part can be said to be relatively independent part, when an exam is over, the answer of the candidates will be exported from database, then input it into marking database, run the marking subsystem, the multiple choice can be dealt with in real time.

After the analysis to the function of the system, we design the function module diagram as shown in Figure 2, after student users login the system, first browse test instructions, and then test. This paper is mainly on the teacher module. It includes teachers' proposition and teachers marking. The teacher proposition module includes two big modules, one is multiple choice module, the other is Office operation module. In Office module, it contains Word and Excel operation proposition. In the two modules comprises their sub module. We in this paper introduced the Font module proposition. Because the Word examination in general including font, paragraph, table three cases study. This system is mainly to achieve the three modules. First of all, when teachers in the Font module provide proposition, the attributes of Font need to be surveyed are selected and set the score, the remain work will be completed by the system. The system then began opening its storage of correct answers in the document, read the Font attribute selected by teachers, the system then regards its attribute as the correct answer and inserted into the database.

Summary

The research is based on the B / S mode, the computer makes mathematical analysis on random group algorithm, in the subjective operation test marking part, through technology research to the Microsoft Office and COM software, we realize automatic scoring operated by the Office operation. Conquer the difficulty that subjective questions could not be marked in previous examination, in the design process broadens my computer applications in the research field, the software system development process in the technical implementation did relatively thorough discuss.

But because of time constraints, the system still has a few places to be improved, such as tightness, exception handling problems. We also plan to apply other products of Office in this system, it can broaden the observation range to students and better system adaptability.

In conclusion, the realization of this system can be applied not only in Changchun Vocational Institute of Technology" Computer Culture Basis" course examination, but is also extend to other Vocational Colleges which are using " Computer Culture Basis" published by us. The system is practical, database is still need to be increased, after the system becomes comprehensive perfect, we may envisage that promote the application of the system among higher Vocational Colleges in Jilin province .Making efforts to prompt information management process of higher vocational" Computer Culture Basis" examination.

References

[1] Shi, Z., Liu, D., and other applications of Beijing Tsinghua University press, pp. 154–178 (2005)
[2] Wu, C., Sun, S., Leaf, G.: ASP.NET database navigation project cases, vol. (12), p. 143. Tsinghua University Press, Beijing (2004)
[3] Zhu, J.: One based on. NET framework online examination management system implementation. Electronic Science and Technology University Master thesis (07), 4–5 (2009)
[4] Zheng, A.: SQL Server practical guide. Publishing House of Electronics Industry, Beijing (2001)
[5] Zhou, X., Bai, H., Guan, L.: SQL Server 2000entry and improve, vol. (05), pp. 1–341. Tsinghua University Press, Beijing (2001)

MVDR Method for the Whole Conformal Arrays

Lin Zou[1] and Zishu He

[1] University of Electronic Science and Technology of China, No. 2006, Xiyuan Road, Chengdu 611731, P.R. China
weight_0@163.com

Abstract. The minimum variance distortionless response (MVDR) optimization method is studied on the whole conformal array. we use non-isotropic antenna element pattern as in practical engineering. Three-dimensional beamforming pattern for 2×4 and 4×2 cylindrical conformal arrays are presented. Finally, simulation results show that the method is effective and can be utilized for conformal arrays.

Keywords: Mvdr, Whole Conformal Arrays, Signal Processing.

1 Introduction

Conformal arrays have received considerable attention because of their flexibility in attaching to arbitrary surface of vehicles and aircrafts to save space and capability of offering wide angular coverage and avoiding boresight error, etc [1]. However, conformal arrays also present many challenges to the designers, such as manifesting themselves with different element patterns and orientations because of the varying curvature.

Beamforming is one such technique that is the use of adaptive or smart antennas to produce a movable beam pattern, which can be directed to the desired coverage areas and minimize the impact of unwanted noise and interference, thereby improving the quality of desired signal. Conventional approaches of Beamforming for ULA are relatively mature, such as least mean square(LMS) algorithm[2], recursive least square(RLS) [3] algorithm, MVDR[4][5] algorithm, and other combined algorithms[6]. The LMS or RLS are two commonly used algorithms for beamforming. The former has good tracking performance with low computational complexity, and is robust against numerical errors [2]. On the other hand, the RLS algorithm can achieve a faster convergence that is independent of the engine-value spread variations of the covariance matrix [7]. MVDR beamformer is one of the key algorithms in signal processing. In this paper, we will show that MVDR can be used for directional elements in three-dimensional cylindrical conformal array as well.

2 Conformal Array Data Model

In this section we consider a D×E cylindrical conformal array shown in Fig. 1.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 917–921.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

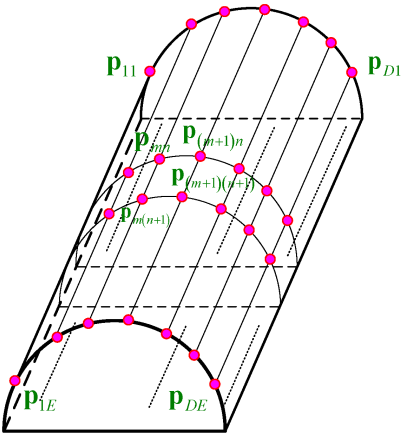


Fig. 1. Geometry of a D×E cylindrical conformal array

The far field of a system of simultaneously excited cylindrical conformal array of antennas may be written, omitting the time phase factor $e^{j\omega t}$, as

$$F(\varphi,\theta)=\sum_{i=1}^{D}\sum_{j=1}^{E}\left[a_{i,j}^{*}f\left(\varphi_{i,j},\theta_{i,j}\right)e^{-\mathbf{k}^{T}\mathbf{p}_{i,j}}\right] \tag{1}$$

Through Euler rotation matrix [8] [9], and we consider that the pattern function of each element is $f\left(\varphi_i,\theta_i\right)=\cos\theta_i$, we can rewrite (1) as

$$F(\varphi,\theta)=\sum_{i=1}^{D}\sum_{j=1}^{E}\left\{a_{i,j}^{*}f_i(\varphi,\theta)e^{-\frac{2\pi}{\lambda}\left[R\sin\theta\cos(\varphi-\varphi_n)+z_n\cos\theta\right]}\right\} \tag{2}$$

3 Mvdr Method for the Whole Conformal Arrays

Consider a conformal array consisting of M antennas. Assuming that the K arriving signals are narrowband signals, we denote the array steering vector as

$$\mathbf{V}(\theta_k)=\left[1\quad e^{-j\phi_k}\quad\cdots\quad e^{-j(M-1)\phi_k}\right]^{T} \tag{3}$$

We know the output of linear array at time n is given by

$$y(n)=\mathbf{w}^{H}\mathbf{x}(n) \tag{4}$$

where $\mathbf{x}(n)=\left[x_1(n)\quad\cdots\quad x_M(n)\right]^{T}\in C^{M}$ is the array observation vector, $\mathbf{w}=\left[w_1(n)\quad\cdots\quad w_M(n)\right]^{T}\in C^{M}$ is the complex vector of beamformer weights. and $(\cdot)^{H}$ denotes the Hermitian transpose.

The observation vector can be written as

$$\mathbf{x}(n)=\mathbf{s}(n)+\mathbf{i}(n)+\mathbf{v}(n)=s(n)\mathbf{V}(\theta_1)+\mathbf{i}(n)+\mathbf{v}(n) \tag{5}$$

where $\mathbf{s}(n)$, $\mathbf{i}(n)$ and $\mathbf{v}(n)$ are the desired signal, interference, and noise components, respectively. $\mathbf{V}(\theta_1)$ is the presumed desired signal steering vector. $s(n)$ is the desired signal waveform. The desired signal and interferers are assumed to be uncorrelated.

The MVDR beamformer [10] minimizes the output interference-plus-noise power while maintaining a distortionless response to the desired signal. The MVDR problem is given by

$$\min_{\tilde{\mathbf{w}}} \tilde{\mathbf{w}}^H\tilde{\mathbf{R}}\tilde{\mathbf{w}} \qquad st.\ \tilde{\mathbf{w}}^H\tilde{\mathbf{V}}(\theta_1)=1 \tag{6}$$

where

$$\tilde{\mathbf{R}}=E\left[\tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}^H(n)\right]=\tilde{\mathbf{A}}E\left[\underline{\mathbf{s}}(n)\underline{\mathbf{s}}(n)^H(n)\right]\tilde{\mathbf{A}}^H+E\left[\mathbf{v}(n)\mathbf{v}^H(n)\right] \tag{7}$$

and $E(\cdot)$ denotes mathematical expectation.

In many cases, $f(\theta_k)$ satisfies the conditions that $f(\theta_k)\neq 0$ $(1\le k\le K)$ and the family of vectors $f(\theta_k)\mathbf{V}(\theta_k)$ $(1\le k\le K)$ is linearly independent. $\tilde{\mathbf{R}}$ is positive definite matrix all the same.

The closed-form solution to (6) is

$$\tilde{\mathbf{w}}_{opt}=\frac{\tilde{R}^{-1}\mathbf{V}(\theta_1)}{f^*(\theta_1)\mathbf{V}^H(\theta_1)\tilde{R}^{-1}\mathbf{V}(\theta_1)} \tag{8}$$

4 Simulation Results

In the experiment, the performance of MVDR beam formers with directional elements for 2×4 and 4×2 cylindrical conformal arrays are studied.

The simulation setup is as follows: D×E cylindrical conformal array with D elements spaced at a half wavelength apart in the same ring, and there are E rings spaced at a half wavelength apart too. The additive noise plus interference is assumed to be a zero mean complex Gaussian noise. We use 128 snapshots.

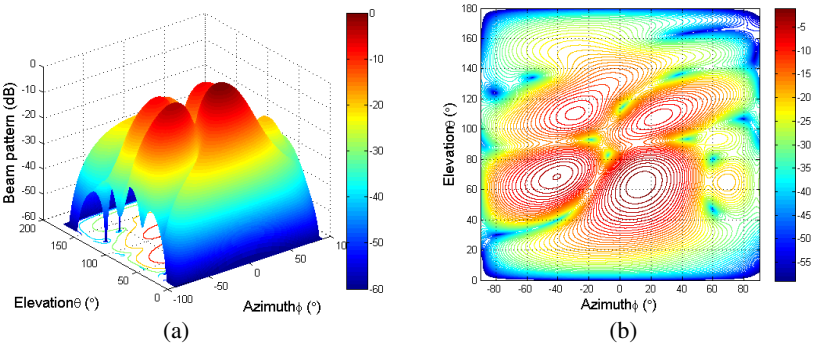
There are three signals impinging upon all arrays:

(1) The signal of interest $s(t)$ with angle of arrival $\varphi_0=20^\circ$, $\theta_0=60^\circ$,

(2) The first interference signal $i_1(t)$ with angle of arrival $\varphi_1=-40^\circ$, $\theta_1=40^\circ$,

(3) The second interference signal $i_2(t)$ with angle of arrival $\varphi_2=60^\circ$, $\theta_2=80^\circ$,

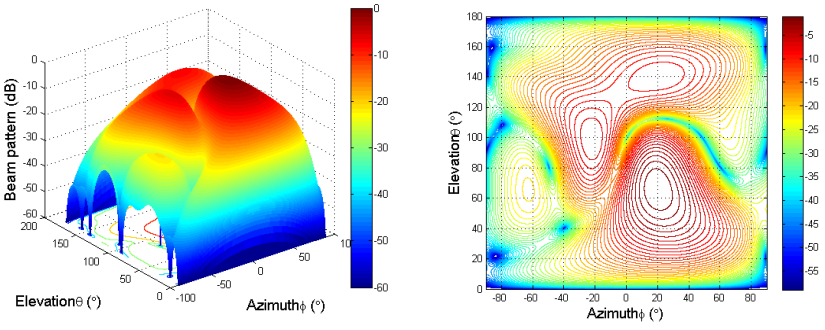
The received noise variance $\sigma^2=1$ and the amplitude of $s(t)$, $i_1(t)$, and $i_2(t)$ is determined based on the SNR and INR.



(a) (b)

Fig. 2. (a) 3D pattern of 2×4 cylindrical conformal array; (b) Contour plot of 2×4 cylindrical conformal array

The simulation results are shown in Figs 2-3. It can be seen that in the sidelobe region the optimum beam pattern has a deep notch at the locations of interferers and the other lobes are well behaved. The MVDR method for whole conformal arrays with directional elements performs well in terms of the main beam width and sidelobe level. As for the main beam, the 4×2 conformal array point to the desired signal more accurately.



(a) (b)

Fig. 3. (a) 3D pattern of 4×2 cylindrical conformal array; (b) Contour plot of 4×2 cylindrical conformal array

5 Simulation Results

The MVDR method for whole conformal arrays with directional elements is proposed. The optimization method in this paper is suitable not only for the cylindrical conformal array, but also for other conformal array.

Simulation results indicate that this method is an effective tool for practical conformal array beamforming.

Acknowledgements. This work is supported by the National Science Foundation of China under grant No. 61032010, the National Nature Science Associate Foundation of China under grant No. 11076006.

References

[1] Josefsson, L., Persson, P.: Conformal array antenna theory and design. John Wiley & Sons, Inc., New Jersey (2006)
[2] Kabilan, A.P., Meena, K.: Performance comparison of a modified LMS algorithm in digital beam forming for high speed networks. In: Conference on Computational Intelligence and Multimedia Applications, vol. 4, pp. 13–15 (2007)
[3] Srar, J.A., Chung, K.-S.: Performance of RLMS algorithm in adaptive array beam forming. In: 11th IEEE Singapore International Conference, pp. 493–498 (2008)
[4] Puska, H., Saarnisaari, H., Iinatti, J., Lilja, P.: Synchronization probabilities using conventional and MVDR beam forming with DOA errors. In: IEEE 18th International Symposium on PIMRC 2007, pp. 1–5 (2007)
[5] Pal, P., Vaidyanathan, P.P.: Frequency invariant MVDR beamforming without filters and implementation using MIMO radar. In: IEEE International Conference on ICASSP 2009, pp. 2081–2084 (2009)
[6] Eldar, Y.C., Nehorai, A., La Rosa, P.S.: A competitive mean-squared error approach to beamforming. IEEE Trans. Signal Processing 55, 5143–5154 (2007)
[7] Lei, W., Lamare, R.C.: A novel constrained adaptive algorithm using the conjugate gradient method for smart antennas. In: Conf. ACSSC 2007, pp. 2243–2247 (2007)
[8] Burger, H.A.: Use of Euler-rotation angles for generating antenna patterns. IEEE Mag. Antennas Propag. 37(2), 56–63 (1995)
[9] Milligan, T.: More applications of Euler rotation angles. IEEE Mag. Antennas Propag. 41(4), 78–83 (1999)
[10] Haykin, S.: Adaptive filter theory, 4th edn. Prentice Hall, New Jersey (2001)
[11] Strunk Jr., W., White, E.B.: The elements of style, 3rd edn. Macmillan, New York (1979)

Research and Implementation of SOA-Based E-Commerce Platform

Hu Liang

Jiangxi Science and Technology Normal University, Nanchang, Jiangxi, 330013, P.R. China
hliang1972@126.com

Abstract. SOA (Service-Oriented Architecture) is an architectural model that can deploy, compose and use loosely coupled, coarse-grained application components through the network. In this paper, taken some travel site as an example, it analyzes the SOA technology applied in the development of e-commerce platform based on the needs of e-commerce with the characteristics and the advantages of SOA architecture of the system.

Keywords: SOA, E-commerce, Web Service.

1 Introduction

E-commerce is to transfer the original traditional sales and shopping channels into the Internet by means of electronic and its technology, where production enterprises break the national, regional visible and invisible barriers to realize globalization, network, intangibility and individuality.

With the rapid development of computer hardware technology in recent years, e-business environment has been improved immensely that network bandwidth, network speed and network usage costs is no longer a stumbling block to development of electronic commerce; but in e-commerce software environment, the relevant standards always lag behind the pace of the development of related applications, which has resulted in wide variety of inter-enterprise e-business solutions and information storage formats. Construction of traditional technology-based business environment, there are two prominent problems: first, the participants did not use the e-uniform technical standards or system architecture as a way to contact each service. The lack of a unified commercial release and discovery mechanism, make the mutual exchange of information and cooperation very difficult; second is the system itself. The problems include long cycle of development, poor reconstruction, upgrade and maintenance problems and so on. These have greatly limited the further development of e-commerce.

To address these problems, we follow-up study in related fields at home and abroad based on the latest research results, build a service-oriented architecture based (Service Oriented Architecture, SOA) e-commerce platform for dynamic e-business application framework. E-commerce platform developed with SOA technology, can improve the reusability of existing software. It can also improve the adaptability of e-commerce platform. SOA e-commerce application platform makes e-business applications unified and open platform so that the e-commerce system in

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 923–928.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

the enterprise has the quick service, fast reconstruction, safe reliability, reusable and free expansion to provide a new application program for the further development of e-commerce.

2 Service-Oriented Architecture (SOA)

At present Service-oriented architecture (SOA) is no uniform definition. The more influential SOA is defined as: "essentially a collection of services. Mutual communication between services may be simple data transferred; it may be conduct two or more coordinating services active. It needs some method of inter-service connections. The so-called service is a function with a precise definition, perfect package, independent of other the environment and state serving. "

Service-oriented architecture (SOA) is an organization based on service computing resources, with loosely coupled services and indirect addressing capability of the software architecture. In essence, SOA is service-oriented software architecture, to design and build a loosely coupled software solutions approach. The basic elements of SOA architecture is service for business processes as reusable components that simplify the information services or the state of the data migration process, to respond to customer requests and provide high quality services.

2.1 SOA in Three Roles

According to the conception of Service-architecture. corn, SOA is essentially a collection of services. Between services communicate with each other, this communication may be simple data transmission, or it may be two or more services coordinating certain activities. It needs to connect between services in some way. So-called service is a function with precise definition sound packaging, and independent of other services, of the environment and the state.

In the SOA architecture, there is collaboration in three roles, namely service consumers, service agents and service providers as shown in Figure 1. Service provider (provider) register and publish through a service agent (broker); service consumers (consumer) can find the required service through the service agent (find), and bind to their applications.

Service providers is to provide services and to post information to the service registry, such as service interface, service access address and other information services, service description and other related meta-information (information such as service providers, quality of service features).

Indirect addressing function provided by service registry provides is to help consumers find and locate the appropriate "intermediary services" .The existence of registry services, stripping the service consumers and service providers rely on direct addressing between the services so that the address change of services will not affect consumers. In addition service registry enables service consumers to realize dynamic service configuration, among many of which matches the selection criteria best, or form service composition.

Service requester can have indirectly or directly obtain service description through registry or through the service provider, and follow the service description and address of the interface and service providers to achieve the interaction.

2.2 SOA Characteristics

SOA has a loosely coupled, coarse-grained services, standardization of interfaces and message-based communication features. The service does not involve low-level communication between the programming interface and communications model.

Loosely coupled is a service interface independent of the service to achieve the hardware platform, operating system or programming language. Loosely coupled SOA is a "loosely coupled" component services, isolation of the service users and service providers to use in the service implementation and client services. Service providers and service users loose coupling between the key point behind the service interface and service implementation as a separate entity exists. If the module changes in the system, it only modifies the module's internal procedures, without impact on other modules, which makes the service implementation not be affected in case that service users were modified to be more flexible to meet business changing needs.

Service granularity refers to the function of a service included the complexity of the services. The greater size means that the service functions. The great size service is, the more complex function is. SOA advocates loosely coupled coarse-grained services, particularly in the design phase, with a reasonable coarse-grained effectively reflecting the new business processes. SOA service interface includes the standardization of message formats, transport protocols and location, which the interaction between services need a variety of criteria. But the hidden details of service implementation allow the interface independent of the hardware platform, operating system or programming language service implements.

SOA services communicate through messages between the uses of pre-defined message. When the service requesters and providers communicate, they do not have to understand the environment in which service providers. To the caller service is transparent, which the client calling the service on the SOA is not necessary to know the concrete implementation of the service, because SOA has encapsulated these services, and publish all services together. When the client calls through the SOA platform, these services architecture, SOA platform invoke these services with the agency and the needs of the business according to the actual input and output parameters for services, transmission agreements to be amended accordingly.

3 The Complexity Demand of e-Commerce Platform

E-commerce platform has developed into an established business platform on the Internet, where the different businesses (buyers and sellers) are gathered in a virtual space for business. The development of E-commerce platform is rather complex with the following features.

E-commerce platform serves for strong dynamic e-commerce platform which is also gathered at a central site of this network, including not only e-commerce businesses, customers, including financial institutions, and tax and administration and

other government departments. It is a very common for e-commerce business clients to change. Fixed though a target is, the company's business may also change with the market demand and e-commerce services business. Only by constantly improving the flexibility of management, e-commerce services businesses respond to the change of user needs.

The electronic market on the e-commerce platform supports all the different enterprises and trade-related activities, but also allows the supply chain aspects of the implementation of cooperation between each other in the design, development, production and distribution. E-business supply chain is almost unable to appear the role of the main chain, which is determined by the nature of e-business services. The more successful e-commerce businesses do, the more diversified clients face, the more extensive services, but also difficult to manage A. Only by constantly improving the management flexibility, it can be brought to the new challenges of the enterprise e-business information management.

E-commerce platform faces more and more severe the security situation. B2B e-commerce transfer the electronic market up to e- market, which requires some of the resources enterprises provided for the development of IT infrastructure applications and internal processes. So a higher level of supply chain integration gradually improved as business systems are connected to an electronic trading platform; electronic market liberalization leads to an increased risk of privacy (such as leakage of sensitive business information), and many transactions improve the complexity of system integration, the cooperation of business process and so on.

It is not difficult for enterprises to understand the analysis above. Fully absorbing the existing knowledge of e-commerce businesses, enterprises might rapidly improve their management level, and actively adapt to the requirements of clients, e-commerce business to survive. For the relatively stable business enterprise, the enterprise e-commerce platform focuses on constantly improving the management of existing business, while the most e-commerce businesses, will have to emphasize flexibility and synergy between objects and services. The development of the traditional idea of e-commerce platform, often confined to certain specific functions resulting in a serious obstacle to the improvement of management efficiency, such as quite conspicuous "information island ", information processing ability, cross-platform exchange of information barriers. Faced with these problems, enterprises need to introduce new software architecture, software re-construct the elements of the external visible properties and relationships between elements in order to adapt to external multilateral business needs., it is SOA service-oriented architecture technology precisely provides a new perspective solution to the problem.

4 E-Commerce Platform Based on SOA Design (with a Travel e-Commerce Platform as an Example)

Some travel website need to build a cross-boundary, cross-platform, easily expand e-commerce system, by which using different hardware and software platform of travel agencies, hotels, transportation departments to establish a unified data platform and achieve their arrival, pre-sale tickets, intelligent group of business strategy. During the system architecture design, the SOA is selected based Web Service model

to solve the business problems, such as the heterogeneous system interconnection, low cost, highly scalable demand.

Running on multiple distributed servers across the country, platform from the structure and function can be divided into: travel portal, Web Service application services, systems data services UDDI, data services, travel agencies, hotel data services, data services, shipping stations, Online banking settlement services and so on. Each type of service consists of different parts of different platforms servers.

The main part of the system and center of gravity respond to the user's request accordingly in the tourism portal and Web Service application service travel portal, while data services are all from the Web Service applications. Cooperation with travel websites, travel agencies, hotels and transportation units publish interfaces to communicate through Web service, according to the principle of proximity to the business data to travel data server cluster. It is unified billing to settle the balance sheet data from the travel gateway while temporary data is stored locally. The system architecture is a distributed computing architecture as well as a typical multi-layer structure:

The user interface layer is tourism web portal to draw with the end-user interface and interaction, and is some aspx pages and related to the page code in the application; business logic layer is to encapsulate business logic and rules, and to call each server's Web service on each server; to provide for the provision of user requests units of response and communication and cooperation; to provide data for the co-operation with uniform access, in a heterogeneous between systems providing support services in the application for the. NET components.

Web services proxy Web service layer calls the production of each proxy class to generate proxy object; data access layer is to complete the interaction between the bottom and the database, in which the underlying data for all Web applications have adopted the same set of operating the business code in order to achieve the purpose of reuse; physical data layer consists of scattered clusters and database services co-operating with the database over the composition, in which uses abstract data access interface to make good use of the class data resources. UDDI Services is all the found and release of Web services in the system to achieve better cooperation between partners in the system.

5 Conclusion

The development of E-commerce platform is a complex process from which ideas, methods and technology are applied to practical management. This paper introduces the idea of serving the development of e-commerce platform, through the SOA framework to deal with the problems existing in the current e-commerce systems, such as integration process, lack of flexibility, scalability and business agility. It has a loosely coupled, business agility, scalability and other advantages, greatly improving the IT capital in the enterprise value.

References

[1] Bi, X., Tan, W.: SOA-based Framework for DDoS Attack Defense Network. Computer Engineering (13) (2009)
[2] Bi, X.: Boundary Feedback Based on Autonomous Distributed DDoS Defense Methods. Computer Engineering (11) (2009)
[3] Zhang, X., Hu, H.: SOA-based Enterprise Application Integration and Implementation. Computer Knowledge and Technology (Natural Science) (17) (2007)
[4] Mei, J.: SOA in the B2B E-commerce Application. Dalian Maritime University (2007)

Production Dynamic Scheduling Method Based on Improved Contract Net of Multi-agent

Ling Qin and Shulin Kan

School of Mechatronics Engineering and Automation, Shanghai University, China
kanshulin@126.com

Abstract. To solve the dynamic and complex problem of production scheduling, this paper proposed a new production dynamic scheduling method based on improved contract net of multi-agent. In this method, a production dynamic scheduling framework model based on multi-agent was built, and each agent function was analyzed in detail. An improved contract net was investigated by means of spirit parameters. By means of improved contract net, the negotiation procedure between scheduling negotiation agent and resource agent was constructed by improved contract net multi-agent. Finally, the simulation system of production dynamic scheduling based on improved contract net of multi-agent was demonstrated and validated by QUEST software. It has shown the proposed method can improve the benefit of production scheduling, and provide a support for adapting to complex and dynamic production scheduling.

Keywords: Dynamic, Scheduling, Multi-agent, Improved contract net.

1 Introduction

Because the mutual effect and relationship among workpieces, machines, fittings, and so on in manufacturing time, assembly time, manipulation order, the production scheduling of workshop is complex. At the same time, because production scheduling is optimal problem in equation or non-equation restraints, the problem scale and calculation quantity production scheduling is tremendous. Therefore, traditional scheduling methods do not adapt to complex production scheduling.

Because urgency and emergency usually offer, workshop state, workpieces, machines, fittings, and so on need to be adjusted. When workshop state is changed, production scheduling need consider operation information, and has dynamic characteristic in the case of machines failure, workpieces order change, fittings adjustments, and so on. Thus, traditional scheduling methods do not adapt to dynamic production scheduling.

There were some researches of production plan and scheduling aspects, such as D. Ouelhadj, et al. [1], T. N. Wong et al. [2], J. Sun and D. Xue [3], N. Aissani, et al. [4], M. K. Lim and Z. Zhang [5], M. K. Lim and D. Z. Zhang [6], J. Reaidy et al. [7], and so on. However, few studies have been devoted to dynamic and complex problem of production scheduling. Furthermore, few researches have investigated the problem by improved contract net of multi-agent.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 929–936.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

The purpose of this paper is to provide a new production dynamic scheduling method based on improved contract net of multi-agent. In this method, a production dynamic scheduling framework model based on multi-agent is built, and each agent function was analyzed in detail. Contract net was improved by means of spirit parameters. By means of improved contract net, the negotiation among agents was investigated, and Production dynamic scheduling algorithm and procedure were designed by improved contract net multi-agent. The objective of the study is to provide guidance for the benefit improvement of production scheduling and support for solving the fluctuation problem the dynamic and complex problem of production scheduling.

2 The Production Dynamic Scheduling Framework Model Based on Multi-agent

In multi-agent, each agent is independent and capable of making decisions. An agent has its own objectives to achieve and its autonomous behavior is a consequence of its observations, knowledge and interactions with other agents [8-10]. The production dynamic scheduling framework model based on multi-agent is composed of agents from different layers in hierarchical structure. The agents of this model have cooperation characteristic to upper layer, and have autonomy characteristic to lower layer, so that they can adjust the factors involving quality, assure the stability of production scheduling. Production can be dynamically close-loop scheduled by the precaution and process treatment integration in the model. The production dynamic scheduling framework model based on multi-agent can be shown as Figure 1.

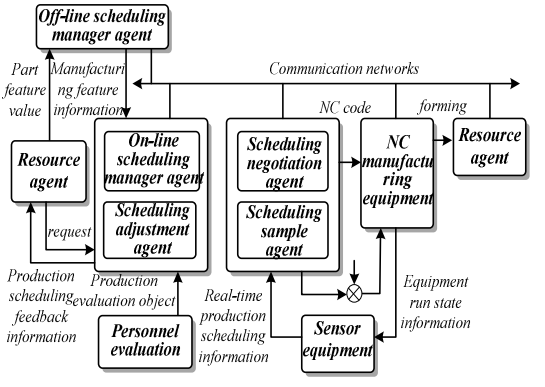


Fig. 1. The production dynamic scheduling framework model based on multi-agent

This model is expressed as MA:

$$MA = \{RA, SA, SNA, SAA, OFFA, ONA\} \tag{1}$$

2.1 Resource Agent (RA)

It is a mapping by inspect equipment, inspect hardware, fitting, manufacturing cell, manufacturing equipment, and so on. It includes all manufacturing resources.

2.2 Scheduling Sample Agent (SA)

The production information can be auto-sample from inspect equipment, inspect hardware, PC, or control networks, and it includes equipment state, running state, and so on.

2.3 Scheduling Negotiation Agent (SNA)

The adverse trendand of production problem can be identified in the deficiency process capability case. The key influence factors of production feature value and optimal factors combination can be found. It can provide a support for production improvement.

2.4 Scheduling Adjustment Agent (SAA)

Based on the key influence factors from scheduling negotiation agent, some methods can be adopted to adjust and change the influence factors, and production scheduling improvement scheme can be designed for assuring stable quality. By networks, the scheme can be transmit to upper layer to analyze and evaluate quality.

2.5 Off-Line Scheduling Manager Agent (OFFA)

Based on quality standard and process value, manufacturing resource including equipments, raw material, personnel, and so on, can be reasonable allocated for maximum use and optimal configuration. The production dynamic scheduling procedure can be stable operated.

2.6 On-Line Scheduling Manager Agent (ONA)

By means of the statistical result of quality feature value, quality information can be analyzed and evaluated. The adverse trendand of production dynamic scheduling problem can be identified and diagnosed by historical data. Production dynamic scheduling scheme can be established for changing or adjusting key influence factors.

Depending on key process quality information, and equipment running state information, or manual input quality information, by means of resource agent, scheduling sample agent, scheduling adjustment agent, off-line scheduling manager agent, on-line scheduling manager agent, production dynamic scheduling information can be transmit to upper layer to evaluate production dynamic scheduling, and evaluate result can be transmit to low layer to adjust and change manufacturing resources. Consequently, production scheduling procedure can be dynamically operated and can be close-loop controlled by the precaution and process treatment integration.

3 Production Dynamic Scheduling Based on Improved Contract Net of Multi-agent

3.1 Improved Contract Net of Multi-agent

Contract net is a cooperation agreement of all node commutation and control in distributed problem solution environment. Tradition contract net can solve a task allocation in multi-agent, and particularly is appropriate for sole-task, sole bid, sole-round bid environment. In contract net negation of multi-agent, the efficiency of task implementation is influenced by the complexity of task, resource consumption is grand, and task quality do not evaluated.

For solving the tradition contract net problem, the spirit parameters are introduced.

Spirit parameters includes as following:

Trust Degree: The probability of scheduling negotiation agent publishes task t to a resource agent, and successfully complement task is called as the trust degree of scheduling negotiation agent, and it is denoted as $Trust(m,i,t)$. $Trust(m,i,t)$'s scope is [0, 1].

Trust Threshold Value: The minimum trust degree of task t allocation in resource agent is called as acknowledge threshold value, and is denoted as Tr. $Trust(m,i,t) \geq Tr$ is necessary condition of task entrusted by scheduling negotiation agent.

Acknowledge Threshold Value: The minimum trust degree of a capability in resource agent is called as acknowledge threshold value, and is denoted as Ts. $Trust(m,i,t) \geq Ts$ is necessary condition of capability saved by scheduling negotiation agent.

Based on spirit parameters(trust threshold value, trust threshold value, and acknowledge threshold value), contract net can decrease bid request work in later stage, the efficiency of task implementation is not influenced by the complexity of task, and task quality can be evaluated. By introduction of spirit parameters, contract net is improved. The negotiation procedure between scheduling negotiation agent and resource agent by improved contract net is shown as Figure 2.

3.2 The Negotiation Procedure between Scheduling Negotiation Agent and Resource Agent

In production dynamic scheduling procedure, time is key factor. Time depends on negotiation cost and is evaluation value of negotiation proposal. Therefore, strategy based time is often used in production scheduling procedure. Depending on strategy based time, the production dynamic scheduling multi-agent negotiation mechanism among factories is established. Scheduling negotiation agent has explicit time demand, for example, scheduling negotiation agent launch a request at before deadline. When the time nears to deadline, scheduling negotiation agent will increase concession degree to obtain consistency at before deadline.

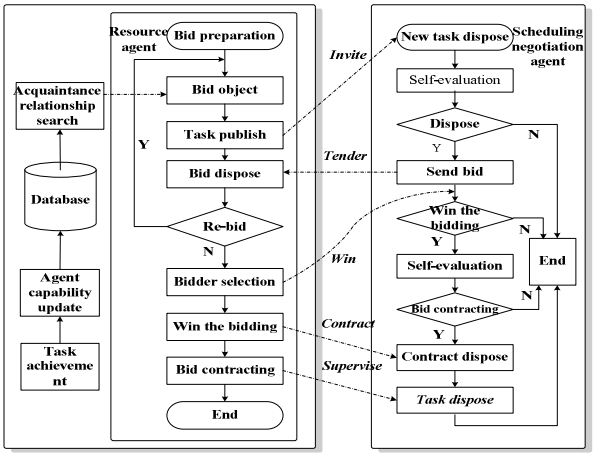


Fig. 2. The negotiation procedure between scheduling negotiation agent and resource agent by improved contract net

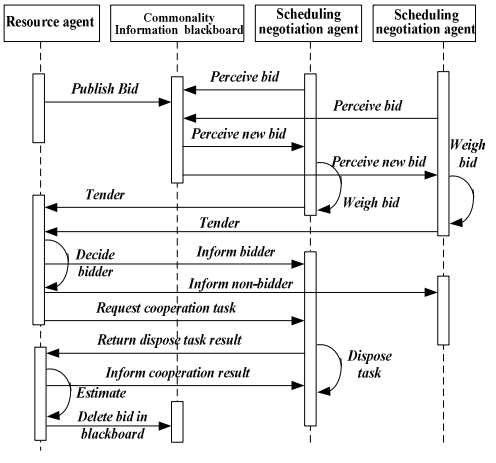


Fig. 3. Bid negotiation procedure between scheduling negotiation agent and resource agent

There is a proposal value generation function based on strategy based time:

$$V_{a_i}^{x_j} = \begin{cases} \min_{a_i}^{x_j} + (1 - \alpha_{a_i}^{x_j}(t))(\max_{a_i}^{x_j} - \min_{a_i}^{x_j}) \\ \min_{a_i}^{x_j} + \alpha_{a_i}^{x_j}(t))(\max_{a_i}^{x_j} - \min_{a_i}^{x_j}) \end{cases} \tag{2}$$

where, $\min_{a_i}^{x_j}$ and $\max_{a_i}^{x_j}$ are maximum and minimum retention value of each proposal a_i. To each proposal a_i, each proposal value $V^{x_j} \in [\max_{a_i}^{x_j}, \min_{a_i}^{x_j}]$. $\alpha_{a_i}^{x_j}(t)$ is independent variable function of time parameter t. $\alpha_{a_i}^{x_j}(t)$ has two fundamental representation form:

Polynomial form:

$$\alpha_{a_i}^{x_j}(t) = k_{a_i}^{x_j} + (1 - k_{a_i}^{x_j})(\frac{\min(t, t_{\max})}{t_{\max}})^{\frac{1}{\beta}} \tag{3}$$

Exponential form:

$$\alpha_{a_i}^{x_j}(t) = e^{(1-\frac{\min(t,t_{\max})}{t_{\max}})\gamma^{\beta} \ln k_{a_i}^{x_j}} \tag{4}$$

where, k is a constant. Proposal value is depended on interval size and k. β is the parameter of strategy based time, and decide the change law of value selection. Finally, there are three production dynamic scheduling multi-agent negotiation rules:

1) ***Boulware rule.*** When $\beta < 1$, scheduling negotiation agent makes a little concession in negotiation begin, and scheduling negotiation agent makes a big concession in negotiation end to obtain negotiation consistency.

2) ***Conceder rule.*** When $\beta > 1$, scheduling negotiation agent makes a big concession in negotiation begin, and PDSNA makes a little concession in negotiation end.

3) ***Linear rule.*** When $\beta = 1$, scheduling negotiation agent keep unalterable concession degree by changeless velocity and scope.

On the basis of the value range of β, The concession size and type of scheduling negotiation agent can be identified in production dynamic scheduling procedure.

Depending on production dynamic scheduling multi-agent negotiation rules, there is production dynamic scheduling procedure between scheduling negotiation agent and resource agent. Figure 3 shows bid negotiation procedure between scheduling negotiation agent and resource agent.

4 The Simulation System of Production Dynamic Scheduling

QUEST/Delmia is simulation software of digital factory and discrete event. It is the solve scheme of CIMS, which could be used to perform process flow design, visually simulate and analysis the veracity and efficiency of process flow [11-12]. With QUEST, users can visualize and prevent potential problems and improve existing processes. QUEST is also a powerful tool to design, analyze and visually represent complex manufacturing processes and data, for those who are not familiar with the manufacturing process.

The simulation system of production dynamic scheduling was consisting of an arrival process, a production process, and a queue. The arrival process is such that whenever the buffer of the group of machines of the very first process is free, a new part will enter. In this simulation, the buffers monitored for arrival correspond to the buffers for the 32ed Lathe group, the 42ed Vertical Lathe group, and the 18th Lathe group. Queuing discipline used in this simulation is first-in first-out (FIFO). The simulation system of production dynamic scheduling is shown in Figure 4.

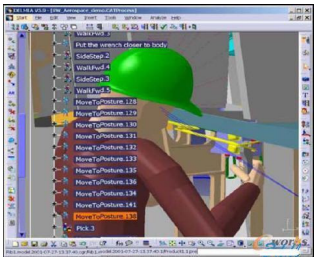


Fig. 4. The simulation system of production dynamic scheduling by QUEST

In Figure 4, the simulation system was run for batch sizes ranging from 1 to 10. Several additional runs were done to determine total distance traveled by laborers, and the possible impact on the model when an additional de-bur machine and milling machine were added. One of the most distinguishing features of QUEST is the ability to calculate distances traveled in production dynamic scheduling. To make use of this useful feature, it was decided to change the arrangement of the plant. After checking trails, it was found that the new layout, if implemented, can save about 4.33% of total distance traveled in production dynamic scheduling. Using QUEST could result in improvements of benefit of production scheduling, and solution of complex and dynamic problem of production dynamic scheduling.

5 Conclusion

A production dynamic scheduling method based on improved contract net of multi-agent is proposed in this paper. In this method, a production dynamic scheduling framework model based on multi-agent is built, and each agent function is analyzed in detail. An improved contract net was investigated by means of spirit parameters. By means of improved contract net, the negotiation procedure between scheduling negotiation agent and resource agent is constructed by improved contract net multi-agent. Finally, the simulation system of production dynamic scheduling based on improved contract net of multi-agent is demonstrated and validated by QUEST software. It has shown the proposed method can solve the dynamic and complex problem of production scheduling, and improve the benefit of production scheduling.

References

[1] Ouelhadj, D., Petrovic, S., Cowling, P.I., Meisels, A.: Inter-agent cooperation and communication for agent-based robust dynamic scheduling in steel production. Adv. Eng. Inform. 18, 161–172 (2004)

[2] Wong, T.N., Leung, C.W., Mak, K.L., Fung, R.Y.K.: Dynamic shopfloor scheduling in multi-agent manufacturing systems. Expert Syst. Appl. 31, 486–494 (2006)

[3] Sun, J., Xue, D.: A dynamic reactive scheduling mechanism for responding to changes of production orders and manufacturing resources. Comput. Ind. 46, 189–207 (2001)

[4] Aissani, N., Beldjilali, B., Trentesaux, D.: Dynamic scheduling of maintenance tasks in the petroleum industry: A reinforcement approach. Eng. Appl. Artif. Intel. 22, 1089–1103 (2009)
[5] Lim, M.K., Zhang, Z.: A multi-agent based manufacturing control strategy for responsive manufacturing. J. Mater Process Tech. 139, 379–384 (2003)
[6] Lim, M.K., Zhang, D.Z.: An integrated agent-based approach for responsive control of manufacturing resources. Comput. Ind. Eng. 46, 221–232 (2004)
[7] Reaidy, J., Massotte, P., Diep, D.: Comparison of negotiation protocols in dynamic agent-based manufacturing systems. Int. J. Prod. Econ. 99, 117–130 (2006)
[8] Csáji, B.C., Monostori, L.,, Kádár, B.: Reinforcement learning in a distributed market-based production control system. Adv. Eng. Inform. 20, 279–288 (2006)
[9] Chen, K., Chen, C.: Applying multi-agent technique in multi-section flexible manufacturing system. Expert Syst. Appl. 137, 7310–7318 (2010)
[10] Wang, L., Lin, S.: A multi-agent based agile manufacturing planning and control system. Comput. Ind. Eng. 57, 620–640 (2009)
[11] Wang, C., Qiu, C.: Virtual simulation of the job shop scheduling system based on delmia/QUEST. In: 2008 Asia Simulation Conference 7th International Conference on System Simulation and Scientific Computing, ICSC 2008, pp. 1129–1132 (2008)
[12] Bzymek, Z.M., Nunez, M., Mu, M.L.: Simulation of a machining sequence using delmia/quest software. Comput. Aided Design. Appl. 5, 401–411 (2008)

Research on Information Integrated Platform for Gear Enterprise Networked Manufacturing

Xiang-tao Chen and Xiang-yang He

Modern Education Technology and Information Center
Henan University of Science and Technology
Luoyang, China
cxtxz119@126.com, hexy@mail.haust.edu.cn

Abstract. To deal with the problems of information integration and data sharing among directed heterogeneous application systems in gear enterprises, a functional architecture of information integrated platform is proposed after detailedly analyzing the information integrated operation mode of gear enterprise networked manufacturing. The construction and exploitation of the integrated platform are implemented with the technology of Service-oriented Architecture (SOA) and Web services. Finally, example of actual hypoid gear networked manufacturing is introduced to demonstrate the practicability of the proposed platform.

Keywords: gear enterprise, information integrated platform, integrated operation mode, networked manufacturing, Service-oriented Architecture, Web services.

1 Introduction

Lots of information systems consist in gear manufacturing enterprises. However, because the system architecture, development language and data type are different among application systems, the information transmission and integration become difficult in the process of gear manufacturing. Common Object Request Broker Architecture (CORBA), Distributed Component Object Model (DCOM) and other distributed technologies are generally used to implement the integrated manufacturing by many scholars [1], [2], [3], [4]. However, they have become more and more difficult to fit the demand of enterprise information integration for tightly-coupled interface, protocol-specific, etc. Web services [5], based on SOA architecture model, have characteristics such as cross-platform, loosely coupled, etc, and can commendably meet the requirement of information integration for gear enterprises. In this paper, the information integrated platform for gear enterprise networked manufacturing (IIPGENM) is constructed by deeply analyzing the information integrated operation model (IIOM) of gear enterprise's application systems. The application integration and software architecture of IIPGENM are implemented based on SOA and Web services. Finally, an example of actual hypoid gear networked manufacturing is given to validate the practicability of the IIPGENM.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 937–944.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 IIOM of Gear Enterprise Networked Manufacturing (GENM)

In the process of GENM, the integration and transmission of gear-related manufacturing information are implemented by the integration of application systems (e.g., ERP, CAD, PDM, MES, and DNC) in gear enterprise. The IIOM of the GENM is shown in Fig.1. ERP implements the supply and demand management of gears with the SCM and CRM. It can also establish procurement and production plans according to market forecast, customer orders, materials inventory of gear enterprise, etc. The design systems i.e., CAD, CAM, CAPP, and PDM, according to the design task of gear, can jointly carry out the product design and processing planning of gear to complete the design of machining drawings, technics, programs, etc. PDM mainly manages varieties relevant information in the process of gear networked manufacturing, and sends gear manufacturing bills of materials (MBOM) to ERP. ERP constitutes the gear manufacturing task according to the MBOM and production plan, and then sends the task and preparation information of processing resources to MES. By combing the gear design information (e.g., drawings, technics, and programs) from PDM, MES executes production scheduling and mission planning according the production order, device status, assignment status, etc, and then dispenses the Dispatching Singles(e.g., production task, NC programs, and tool information) to bottom machining equipment through DNC. At the same time, MES obtains the real-time status of bottom equipment and production schedule of gear, etc through DNC, and feeds back the production capacity of current equipment, execution status of production order, etc to ERP. PDM obtains the gear manufacturing information (e.g., dynamic information of equipment, machining programs). ERP and PDM achieve effective order assignment, production scheduling, technics adjustment, etc so as to dynamically guide the gear manufacturing and realize information management and networked manufacturing of gear enterprise.

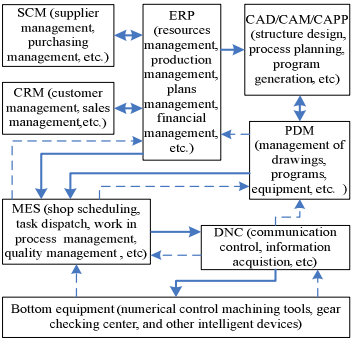


Fig. 1. IIOM of the GENM

3 Functional Architecture of IIPGENM

The functional architecture of IIPGENM is constructed by combining the IIOM of the GENM. It mainly includes user access layer, function services layer, application enabled layer, data resource layer and system support layer, just as Fig.2 shown.

User access layer, providing all kinds of services to different users of IIPGENM, mainly includes users inside and outside gear enterprises.

Function service layer, providing various networked manufacturing function services to the internal and external gear enterprises, mainly includes gear production management, information management and system management modules. Thereinto, gear production module mainly implements the information integration and services application of the full manufacturing process of gear; system management module, including users management, roles management, etc, mainly achieves the basic management of the IIPGENM; information management module implements all subsidiary function integration of gear machining, and mainly includes human resources management, document management, etc.

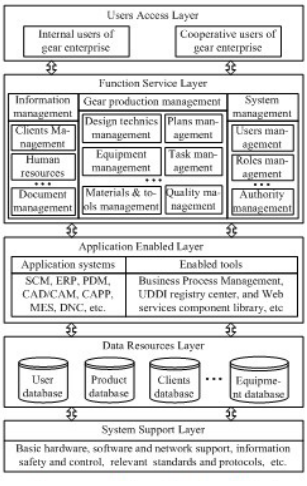


Fig. 2. Functional architecture of IIPGENM

Application enabled layer, mainly including application systems and enabled tools, provides support for the development, operation and various services of IIPGENM. Application systems, providing software systems for gear networked manufacturing, implement the data transmission and information sharing both inside and outside gear enterprise by encapsulating and publishing, so as to complete the full process of GENM. Enabled tools, providing methods and tools for the construction and operation of IIPGENM, mainly includes UDDI registry center and Web services component library, etc.

Data resources layer, being responsible for business data storage and sharing of integrated platform, mainly manage the relational manufacturing resources of gear networked manufacturing.

System support layer, is the system-foundation to implement the IIPGENM, and mainly includes basic hardware, software and network environment, etc.

4 Construction and Realization of IIPGENM

4.1 Application Integrated Architecture of IIPGENM

Based on standardized service interface, Service-Oriented Architecture (SOA) [6] provides across-platform and loosely coupled application framework for information integration of gear enterprise. As the means to implement SOA, Web services [5] are described with standardized XML messages, and can effectively implement dynamic integration of heterogeneous applications for gear enterprise. The integrated framework of IIPGENM based on Web services is shown in Fig 3.

Users within enterprise can directly access the platform through Intranet, and the outward and cooperative users of enterprises can remotely access to the platform through Internet by traversing the firewall. Integration Services Engine provides unified Web service interfaces by encapsulating various application systems exploited on J2EE, .NET, and CORBA, etc as Web service components through standard interfaces, and then publish them to UDDI registry for users calling. Applications are either existing applications or newly developed Web services applications. For the former (e.g., ERP, MES, and PDM), we generate the WSDL files of applications to describe the system functions and calling methods, etc, package them into Web Services components; then exploit adapters of legacy systems; and finally publish the WSDL files to UDDI registration server which can be divided into private and public registration center through UDDI API. SOAP client-side, i.e. application Server, after receiving request SOAP of users, finds relevant Web services in UDDI registry via SOAP, and then sends a XML-based SOAP service request encapsulated by SOAP proxy to SOAP server-side. SOAP server-side finds the corresponding information by receiving and interpreting the received request information, and then returns the responding SOPA information back through the SOAP proxy to implement services binding and calling.

4.2 Software Realization of IIPGENM

According to the aforementioned integrated framework of the IIPGENM, the software architecture of the IIPGENM is constructed in **.**NET environment. It mainly includes

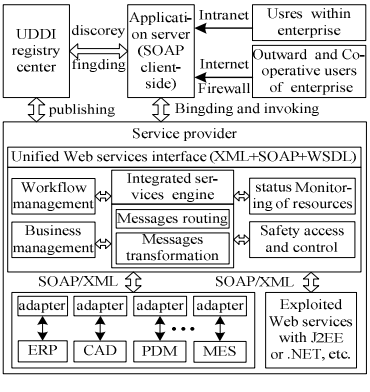


Fig. 3. Integrated framework of IIPGENM

presentation layer, Web services layer, bussiness logic layer, data access layer and data layer, as shown in Fig 4.

Presentation layer is mainly responsible for interaction with the end-users, which are directly implemented by ASP.NET controls and HTML.

Web services layer, according to the functional demand of the IIPGENM, encapsulates logic processing classes and methods of business logic layer to relevant Web services in accordance with common standards, different classification and granularity, and defines the names, methods, etc of the Web services. The Packaged Web services can be invoke and combined each other through the service mechanism i.e. description, publishing and discovery.

Business logic layer, used to achieve varieties business rules and logics of the IIPGENM, accesses the database by calling the database execution logics of the data access layer, thereby completing the necessary business rules. The layer is the core of the entire software architecture and includes all business processes programs (e.g., ERP, MES, and PDM). In this paper, business data and corresponding operations are encapsulated into the business entity classes, and then these entity classes are compiled as .NET components to complete the further encapsulation, which paves the way for Web service layer calling.

Data access layer, used to interact with database, mainly extracts or modifies data of the database according to requirements of business logic layer, and we access the database with the ADO.NET.

Data Layer, including various databases (e.g., SQL Server, Oracle, and DB2), mainly stores business data and control data that related to application programs.

5 Application and Experiment of IIPGENM

The presented IIPGENM has been initially applied to the networked manufacturing engineering of "Henan Key Laboratory of Modern Mechanical Design and Transmission System". It mainly includes Web server, CAD/CAM workstation, ERP server, numerical control lathe CKJ6142, numerical control mills YK22100F and gear

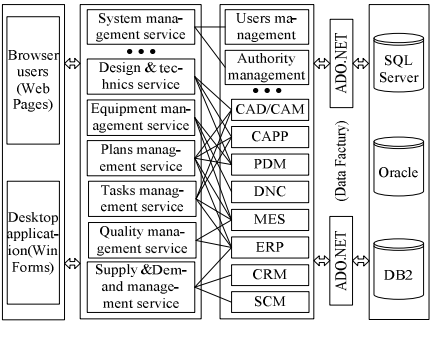


Fig. 4. Software architecture of IIPGENM

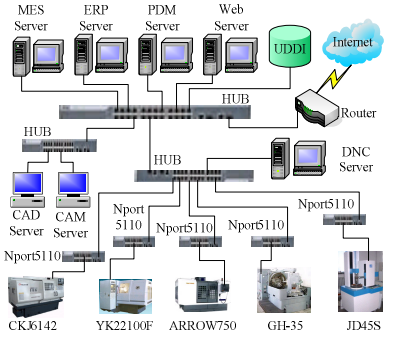


Fig. 5. Overall communication architecture of laboratory

measurement center JD45S, etc. The overall communication architecture of the laboratory is shown in Fig. 5. Interactive Ethernet is used in the laboratory, and the bottom numerical equipment is connected to DNC server by serial port server NPort5110. Various functional systems, integrated based on Web Services, transfer data and share information by coordinating each other through laboratory's Intranet, and provide functional services to the entire gear production process from order to product. Simultaneity, the platform provides UDDI registration service to both internal and external systems, and connects with exterior network (e.g. Internet) by Router. Moreover, the Router also works as a filter to protect the information in system.

Experiment on networked manufacturing of a pair of hypoid gears is given to validate the usefulness of the IIPGENM. Taking pinion for example, according to the production and design plan, the geometric and machining parameters computation of gear, tooth contact analysis (TCA), machining setting cards generation, parametric drawing, machining programs compiling, etc are orderly done on CAD/CAM workstation based on the design task and basal roughcast parameters of gear pairs, which are shown in Tab.1. Thereafter, according the gear task and design information

Table 1. Roughcast Parameters of Hypoid Gears

Items	Pinion	Gearwheel
Number of teeth	7	40
Pitch diameter(mm)		340.00
Pinion offset (mm)	30	
Tooth width (mm)		47
Pinion midpoint helix angle (°)	45.08	
Crossed axes angle (°)	90	
Average pressure angle (°)	22.5	
Tooth contracting model	Normal contracting	
Rotation direction	Left	Right

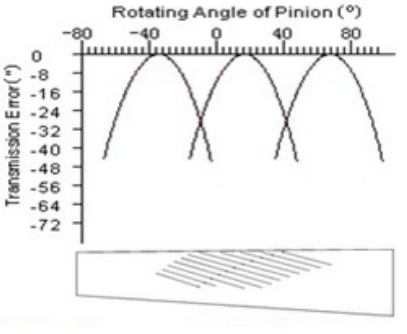


Fig. 6. Gearwheel convex transmission error curves and contact lines

provided by the integrated platform, we machine the gear roughcast of pinion on the CKJ6142, make use of the Modified Roll Method [7] to mill the pinion on the YK22100F. After milling on the YK22100F, we check the tooth surface errors of the pinion on the JD45S by lay outing measurement areas and gridding [8], and then establish the error correction model and find correction values of machining parameters. According the found correction values, we feedback and modify the machining parameters of YK22100F through the laboratory network, and finally make tooth surface errors meet the specified requirements through several measurement and feedback. At the same time, the DNC, PDM, and MES, etc can obtain the relevant manufacturing information of gear and implement the information sharing. For simplicity, the pinion experiment photos i.e., the TCA of gear pair, initial and final gear tooth surface measurement errors are only given in this paper, just as Fig. 6, Fig. 7, and Fig. 8 shown.

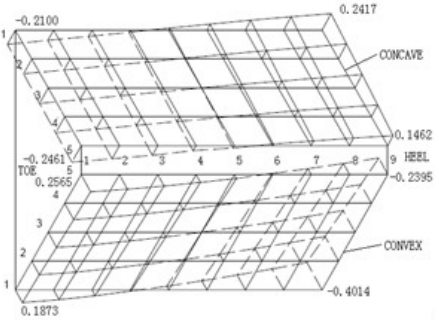


Fig. 7. Initial tooth surface measuring error of pinion

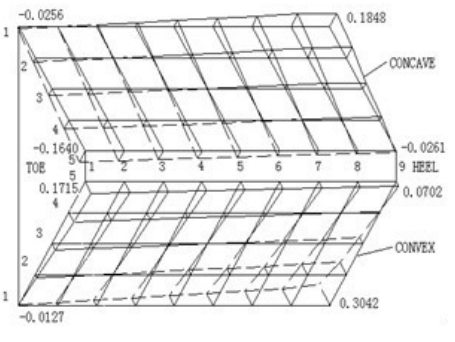


Fig. 8. Final tooth surface measuring

6 Conclusion

The proposed platform i.e. IIPGENM effectively achieves information integration and sharing in the process of gear manufacturing, and also plays an important role for improving production efficiency, reducing production costs of gear enterprise, etc.

Acknowledgment. This topic of research is supported by Natural Science Research Program of Henan Province Education Department (No: 2009A520011).

References

[1] Shin, J.H., Park, S., Ju, C., Cho, H.: CORBA-based integration framework for distributed shop floor control. Computers and Industrial Engineering 45, 457–474 (2003)
[2] Chan, F.T.S., Zhang, J., Li, P.: Agent- and CORBA-based application integration platform for an agile manufacturing environment. International Journal of Advanced Manufacturing Technology 21, 460–468 (2003)
[3] Inverardi, P., Tivoli, M.: Deadlock-free software architectures for COM/DCOM Applications. Journal of Systems and Software 65, 173–183 (2003)
[4] Zhou, B.H., Li, L.F., Yu, C.M.: DCOM and MMS-based control software architecture for automated manufacturing system. International Journal of Advanced Manufacturing Technology 27, 951–959 (2006)
[5] Estrem, W.A.: An evaluation framework for deploying Web Services in the next generation manufacturing enterprise. Robotics and Computer-Integrated Manufacturing 19, 509–519 (2003)
[6] Erasala, N., Yen, D.C., Rajkumar, T.M.: Enterprise application integration in the electronic commerce world. Computer Standards and Interfaces 25, 69–82 (2003)
[7] Lin, C.Y., Tsay, C.B., Fong, Z.H.: Mathematical model of spiral bevel and hypoid gears manufactured by the modified roll method. Mechanism and Machine Theory 32, 121–136 (1997)
[8] Wang, J., Wang, X.C., Jiang, H., Feng, W.J.: Measurement and Compensation of Deviations of Real Tooth Surface of Spiral Bevel Gear. Chinese Journal of Aeronautics 16, 182–186 (2003)

A New 3D Statistical Method for Mobile Channel Model

Yihuai Yang[1], Gengfa Fang[2], Eryk Dutkiewicz[2], and Yangqing Li[1]

[1] School of Automatic Control and Mechanical Engineering, Kunming University, China
[2] Department of Electronic Engineering, Macquarie University, Sydney, NSW, Australia

Abstract. 3-dimension statistical method is used to simulate the popular micro-cell propagation model: COST231-Walfish-Ikegami model[6][7]. The extended model allows for improved path loss estimation by consideration the character of the urban environment parameters are lognormal random variables rather than deterministic. The numerical calculation results and trial measurement data is presented. Based on the measurement and modelling results, we find that the 3D COST231-WI model fits very well with the measurement result in the metropolitan scenario. The conclusion is useful and helpful to network planning in the future.

Keywords: propagation simulation, path loss, COST-Walfisch-Ikegami model, wireless systems.

1 Introduction

Unlike the wired channels that are stationary and predictable, radio channels are extremely random and do not offer easy analysis. From now on, the mobile radio propagations models are usually studied from two ways. The one method is using deterministic models[1][2] which are based on the electromagnetism (such as the ray tracing methods), these models are suitable for indoor and micro-cell circumstances. However, being the fixed geometry (such as buildings, streets···) and time-consuming computations, this method are not use popular. Another method is using statistical models[3][4][5] which are based on amount of measurement data. They linked to the environment and the parameters of the measurement campaign. Such models are proven technique and have accepted by ITU, ETSI and many famous communications companies, such as Motorola, Samsung and Nokia. Recently the statistical models, such as Okumura-Hata model, COST231 model and Egli models are popular used.

For the case of multiple diffractions the complexity of the propagation circumstance increases dramatically. The COST231-Walfisch-Ikegami (COST231-WI) model[6][7] is one of the most successful model for the case of multiple diffractions propagation prediction. This model is a combination of the models of J. Walfisch and F. Ikegami. It was enhanced by the COST 231project. It distinguishes LOS and NLOS, and considers the buildings in the vertical plane between the transmitter and the receiver. The model allows for improved path loss estimation by consideration of more data to describe the character of the urban environment, namely the average rooftop level h_B, the average street width w, the average separation distance between the rows of buildings b and the road orientation with respect to the direct radio path φ. Because the three parameters

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 945–951.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

depending on the buildings (street width, building heights, building separation) are not identical for all locations in the cell, and these parameters are not uniform in the real circumstance, so it is not easy to use.

The statistical approach[8] can solve above problem, and let the COST231-WI model be used in different areas, but only one parameter is variable in each simulation. It is not compatible with the actual situation. In this paper we try to solve the above problem through using bidimensional statistics analysis method[8][9] to make propagation simulation in a metropolitan NLOS environment. The conjunction of the three paramters (h_B, w, b) are estimated at the same time. It will be more close to the real circumstance.

2 Propagation Models

2.1 COST231-WI Model

COST231-WI model[6][7] is one of the most popular models for multiple diffractions propagation prediction. The model has been accepted by the ITU-R and is included into Report 567-4.The estimation of path loss agrees rather well with measurements for base station antenna heights above roof-top level. The important parameters for this model are the base station (BS) antenna height h_b , the mobile station (MS) antenna height h_m , the mean building height h_B , the street width w , the separation distance between the rows of buildings b and the the distance between BS and MS d. A typical situation of is model given in Fig 1. The angle of incidence from the BS to the first diffraction edge φ as shown in Fig 2.

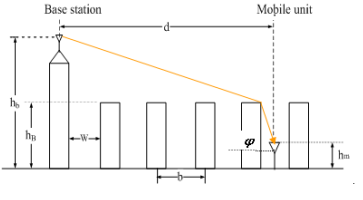


Fig. 1. COST-Walfisch-Ikegami Model

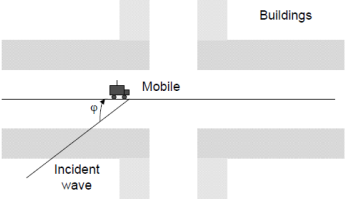


Fig. 2. Defnition of street orientation angle

The model distinguishes between line-of-sight (LOS) and non-line-of-sight (NLOS) situations. For the NLOS situations, the path loss of COST231-Walfisch-Ikegami Model is composed of the terms free space loss L_{fs} , roof-top-to-street diffraction and scatter loss L_{rts} and the multiple screen diffraction loss L_{msd} :

$$L_{NLOS} = \begin{cases} L_{fs} + L_{rts} + L_{msd} & (L_{rts} + L_{msd} \geq 0) \\ L_{fs} & (L_{rts} + L_{msd} \leq 0) \end{cases} \tag{1}$$

The free-space loss L_{fs} is given by:

$$L_{fs} = 32.44 + 20\lg f + 20\lg d \tag{2}$$

where, f is frequency in **MHz**, d is distance in **km**.

The term L_{rts} describes the coupling of the wave propagating along the multiple-screen path into the street where the mobile station is located.

$$L_{rts} = -16.9 - 10\lg w + 10\lg f + 20\lg \Delta h_b + L_{ori} \tag{3}$$

where, $\Delta h_b = h_b - h_B$, w, Δh_b is in m, and the street-orientation fading L_{ori} is defined by:

$$L_{ori} = \begin{cases} -10 + 0.354\varphi & 0^\circ \le \phi \le 35^\circ \\ 2.5 + 0.075(\varphi - 35^\circ) & 35^\circ \le \phi \le 55^\circ \\ 4.0 - 0.114(\varphi - 55^\circ) & 55^\circ \le \phi \le 90^\circ \end{cases} \tag{4}$$

The heights of buildings and their spatial separations along the direct radio path are modeled by absorbing screens for the determination of L_{msd}.

$$L_{msd} = L_{bsh} + k_a + k_d \lg d + k_f \lg f - 9\lg b \tag{5}$$

where

$$L_{bsh} = \begin{cases} -18\lg(1 + \Delta h_b) & \Delta h_b > 0 \\ 0 & \Delta h_b \le 0 \end{cases} \tag{6}$$

$$k_a = \begin{cases} 54 & \text{for } \Delta h_b > 0 \\ 54 - 0.8\Delta h_b & \text{for } \Delta h_b \le 0 \text{ and } d \ge 0.5 \\ 54 - 1.6d\Delta h_b & \text{for } \Delta h_b \le 0, \text{and } d < 0.5 \end{cases} \tag{7}$$

$$k_d = \begin{cases} 18 & \Delta h_b > 0 \\ 18 - 15\dfrac{\Delta h_b}{h_B} & \Delta h_b \le 0 \end{cases} \tag{8}$$

$$k_f = -4 + \begin{cases} 0.7\left(\dfrac{f}{925} - 1\right) & \text{for medium sized city and suburban} \\ & \text{centres with medium tree density} \\ 1.5\left(\dfrac{f}{925} - 1\right) & \text{for metropolitan centres} \end{cases} \tag{9}$$

2.2 3-Dimensional Statistics Analysis

The paper presents a new extension method[9] to apply the popular COST231-WI model. Through this method, the model can valid for generalized conditions in mobile communications. We assume the important parameters h_B, w, and b are variables rather than deterministic, and these three parameters are vary at the same time. The Fig 3. can illustrate this scenario. It is more close to the realistic.

Being the metropolitan centres network planning and optimization works are mostly concentrates on metropolitan areas and mostly NLOS environment, the analysis presented in this paper is limited to these circmustance.

Because of the higher probability of the following situation of metropolitan areas: the scatter loss L_{rts} plus the multiple screen diffraction loss L_{msd} are large than zero, the base station antenna located above the roof-tops, the height of receiver antenna h_{m} lower than the h_B.

The COST231-WI model can be simplified to:

$$L_{rts} = 161.82 - 19\lg(2w) + 20(h_B - 2.5) - 18\lg(91 - h_B) \tag{10}$$

where, ***w=b/2*** and is 90 degrees, the default values of ***w*** and ***φ*** are recommended by COST231, and h_m=2.5, h_b=90m.

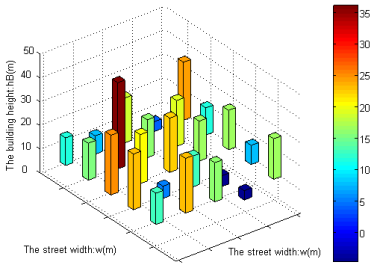


Fig. 3. Characterisation of buildings

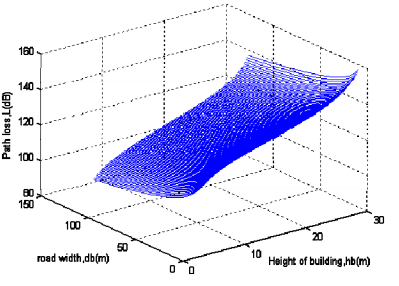


Fig. 4. Pass loss of the COST231-WI Model

Form the Fig 4. we can clear see that the three parameters: h_B, ***w***, and ***b*** are very importance for calculation of the path loss for COST231-WI model.

Implementing this model on a computer involves generating two matrix of N×N uncorrelated lognormal variables, which with appropriate means and variances to represent the h_B and ***w***, and then applying the simplified path loss model, which is based on the distance between the MS and the BS.

Being the random, the results are different in every time. Following is a sample with N=5.

Uncorrelation values for h_B:

7.5906	35.9674	23.1406	25.6818	28.0461
13.9070	28.3163	21.5799	25.3525	19.0907
15.4403	3.7944	23.3956	3.5841	24.9439
14.3268	16.8520	16.7982	21.0314	39.1045
4.3247	36.0386	8.3481	23.8262	19.4281

Uncorrelation values for ***w***:

23.9325	15.4515	28.7873	16.3016	29.8130
24.5284	27.5845	20.9359	16.2000	5.5428
11.2880	20.3295	23.6189	18.1027	10.1704
10.0183	4.2887	29.7963	16.7756	12.2598
10.2346	18.6792	14.1194	18.8333	14.1853

The bigger the N value the more precise, and correspondingly, the computing time will be longer. Compromise considering, we select N=100 in this paper.

The lognormal distribution functions are selected because the physical characteristics of the buildings and the center-center spacing of the rows of buildings have these distributions[8].

3 Simulation Results

The effect of the variation of the three parameters: h_B, w and b, in the same time, are analyzed by using the probability density function. Fig 5. shows the pdf of the path loss when the mean of the h_B is 18, and variance is 12, meanwhile, the mean of the w is 20, and variance is 10. Fig 6. shows the pdf of the path loss when the mean of the h_B is 18, and variance is 4, meanwhile, the mean of the w is 20, and variance is 5.

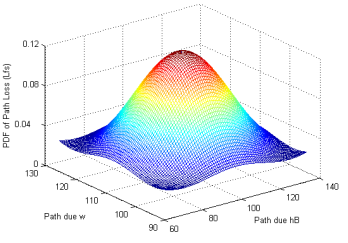


Fig. 5. pdf of L_{fs} (1)

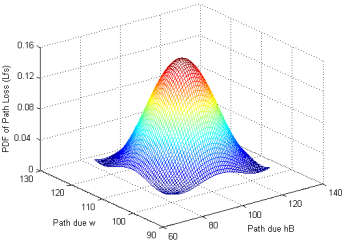


Fig. 6. pdf of L_{fs} (2)

4 Predicted Results

The measurement was performed at Yunnan Province, Kunming City, metropolitan centres with density buildings, NLOS circumstance. The measurement is carried out at a frequency of 2GHz. The equipment for propagation measurement consisted of two parts: the transmitting terminal and the receiving terminal. The transmitter composed with a signal generator, a power amplifier and a horn antenna. The receiver composed with a GPS device, a portable computer and a spectrum analyzer. The transmitting power is 30dBm, the receive antenna is an isotropic antenna with 10dBi gain. The GPS receiver can get the geographical coordinate of the sample points. The transmitter was located on the 15th floor of the building. The h_B of the building is about 45m. The maximum, minimum and average heights of the buildings in the area are about 35 m, 5m and 18 m, respectively (see Fig 3.).

Fig 7. shows the path loss of the extended COST231-WI model when three parameters vary at the same time. The simulation is done under three sort of situations.

Fig 8. shows the comparison between the extended 3D statistical model with the measured data.

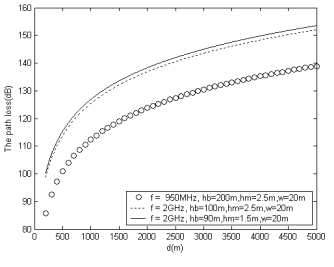


Fig. 7. Path Loss of the extended COST231-WI model

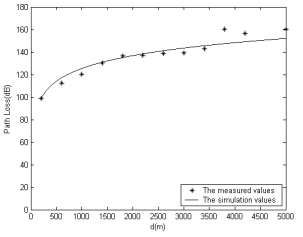


Fig. 8. Comparison between extended model and measured data

Fig 8. indicates that our path loss prediction is fit well with the simulation results. This shows the accuracy of the statistical model.

5 Conclusion

It is very important to do network planning before doing network deployment. During the planning, the study of the mobile propagation models can provide interference prediction and calculation of the wireless propagation environment, it is the vital and the base work of the network planning. Only doing it better can we finish link and coverage estimate. The propagation models' accuracy will affect the reasonable of the network planning.

In this paper, a typical example of COST231-WI channel model is developed, where the three important parameters: the building height (h_B), the street width(w) and the separation distance between the rows of buildings (b) are not deterministic variable as in the original model, but statistical variables, and three parameters are varying at the same time.The performance of the model is studied by simulating probability density functions (PDF), marginal communication probability and the rate of the valid service area.

Measurement work was also done in this paper to verify the validity of the extended statistical model. The measurement system is established for 2GHz based on the sophisticated measurement instruments and the virtual instrument technology. The comparison between the extended WI model and measured data indicates that the proposed model provides relatively well agreement with measured data. The proposed method is easy to operate with 3-dimensional metropolitan scenario.

References

[1] Rossi, J.P.: A ray model for decimetric radiowave propagation in an urban area. Radio Science 27(6), 971–979 (1992)

[2] Oestges, C., Paulraj, A.J.: Physical model for broadband wireless channels. Electron. Lett. 37(3), 1195–1197 (2001)

[3] Chun, L.: A Statistical Model for a Land Mobile Satellite Link. IEEE Trans. Veh. Technol. 34(3), 275–281 (1985)

[4] Lutz, E., Cygan, D., Dippold, M., Dolainsky, F., Papke, W.: The land mobile satellite communication channel- Recording, statistics, and channel model. IEEE Trans. Veh. Technol. 40(2), 375–386 (1991)
[5] Corazza, G.E., Vatalaro, F.: A statistical model for land mobile satellite channels and its application to nongeostationary orbit system. IEEE Trans. Veh. Technol. 43(3), 738–742 (1994)
[6] Walfisch, J., Bertoni, H.L.: A theoretical model of UHF propagation in urban environments. IEEE Trans. Antennas and Propag. 36, 1788–1796 (1988)
[7] Ikegami, F., Yoshida, S., Takeuchi, T., Umehira, M.: Propagation factors controlling mean field strength on urban areas. IEEE Trans. Antennas and Propagation 32, 822–829 (1984)
[8] Cavdarih: A statistical approach to Bertoni-Walfisch Propagation Model for Mobile Radio Design in Urban Areas. IEEE VTC 53rd 1(3), 279–283 (2001)
[9] Neto, A.P., Rozal, E.O., Pelaes, E.G.: Bidimensional Statistics Analysis of the Bertoni-Walfisch Propagation Model for Mobile Radio Design in Urban Areas. In: International Microwave and Optoelectronics Conference (IMOC 2003), vol. 2, pp. 801–806 (2003)
[10] Rappaport, T.S.: Wireless Communications Principles and Practice. Prentice Hall, USA (2002)
[11] Hata, M.: Empirical formula for propagation loss in land mobile radio services. IEEE Transactions on Vehicular Technology 29(3) (September 1980)
[12] Damosso, E., Correia, L.M.: COST Action 231 – Digital mobile radio towards future generation systems. Final report, European Communities (1999)
[13] Nakagami, M.: The m-Distribution, a general formula of intensity of rapid fading. In: Hoffman, W.C. (ed.) Statistical Methods in Radio Wave Propagation, pp. 3–36. Permagon Press (June 1960)
[14] COST 231, Urban transmission loss models for mobile radio in the 900- and 1,800 MHz bands (Revision 2). COST 231 TD(90)119 Rev. 2, The Hague, The Netherlands (September 1991)
[15] Yang, Y., Xie, Y., Shen, D.: An Improved COST-WI Model in metropolitan Areas. In: Proceedings of ICMTCE 2009, pp. 284–287. IET Press, Beijing (2009)
[16] Yang, Y., Shen, D., Xie, Y.: Statistics Analysis of the Sakagami Model. In: Proceedings of 3rd IEEE ICCSIT 2010, vol. 2, pp. 445–448. IEEE Press, Chengdu (2010)

An Optimized Organizational Structure Model of Internet Security Management

Ran Lia[1] and Shoulian Tang[2]

[1] School of Economic and Management
[2] Beijing University of Posts and Telecommunications Beijing, China
liran1983a@gmail.com

Abstract. Internet symbolizes a splendid achievement of human civilization in the twentieth century. It has been evolution from special academic and military network to important message based facilities which almost to all kinds offields, such as political, economical, trade, culture mediums, and education fields and it produces large influence.

Internet has been improving society revolving efficiency and level of productive forces. It also gives people great convenience *for* their work and life. Obviously it has been necessary part for human being. But the new technique-Internet is called two-sided sword which means bringing people a lot of negative influence except bringing people's benefits. The internet security problem is became a important issue for our government.

Based on the existing internet management organization, this paper focus on the current internet security situation, proposed an optimized organizational structure model of internet security management. We show that the *current* national internet management organization have some problems to affect its oral operation. Some internet issues occurred suddenly and couldn't be proposed well. So we research the existing internet security management system, and through analysis its units' responsibility and right to find the main problems. And proposed the optimized organization structure model of internet security management to improve the exist problems.

Keywords: Internet Governace, Organization Structure, Control Subject, Control Object, Control Aim.

1 Introduction

In recent years, national statistics and studies show that the most direct and significant impact of Internet on today's society is brought by the network 's huge, small, complex information dissemination and interaction. Proof of a large number of social hot events, the information capacity of the increasingly powerful social reality, and human communication in daily life and social order are closely related. We have to pay attention on its impact of this new information dissemination phenomenon.

Chinese *government* as the leader in Internet governance, in the process of Internet governance play an important role.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 953–958.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

However, it has its own unique complexity unlike other government Internet governance management. From the government point of view, it is involved in all aspects of Internet governance. Open, grassroots-style development, industry blurred the characteristics that make the Internet is not a single department keep things under control, the Internet governance have a relationship with the various government departments. All departments have the responsibility to rule it. Our government does not have a unified authority of the department to monitor the Internet, a number of government departments responsible for this, the Information Industry Department, the Public Security Bureau, Security Bureau, Department of Cultural Affairs, the Central Propaganda Department, the State Council Information Office, the State Press and Publication Administration and the State Security Bureau and other departments in accordance with their own functions, the introduction of regulations related to the Internet, the Internet, special management. Although the current internet governance has been a lot of government departments involved, but the rapid growth of the internet will make more government departments involved. Therefore, governance requires a high degree of coordination of various departments, poor coordination, may prevent its development. This paper we focus on internet security management for the current domestic situation, proposed an optimized structure model for Internet security management organization, to better promote the internet's healthy develop.

2 The Existing Internet Security Management System

Recently, the Government is to standardize the focus of supervision of social and personal information behavior, and, the effective dissemination of information in solving internet problems; in the period ahead, the government regulation at the center is to train government, social organizations and the public's network Information literacy, and continuously build and adapt to innovation and social progress of network culture.

Government have to properly carry out the main responsibility of rule-making, monitoring and *implementation* of internet information management. The clearly rights, responsibilities and execution agencies is to ensure that an important prerequisite for its successful exercise. Figure 1 is a responsibility analysis chart involved in the agencies of Internet information management.

Industry associations and various organizations should play a key role in Internet regulation, industry. Industry Associations and business organizations should have industry standard with the government's *supervision*. This is the meso level of internet governance. The relationship between government and enterprises are compete with each other and complement each other during the proceed of internet governance.

At the meso level, the social institutions, business organizations and the impact of Internet people power will play the *important* roles. User autonomy, self-discipline, independent code of conduct, and the people to mobilize all social resources and wisdom to participate in Internet governance information's activities should be advocate. The following picture shows the firms' allocation of responsibility which are involved in Chinese Internet security.

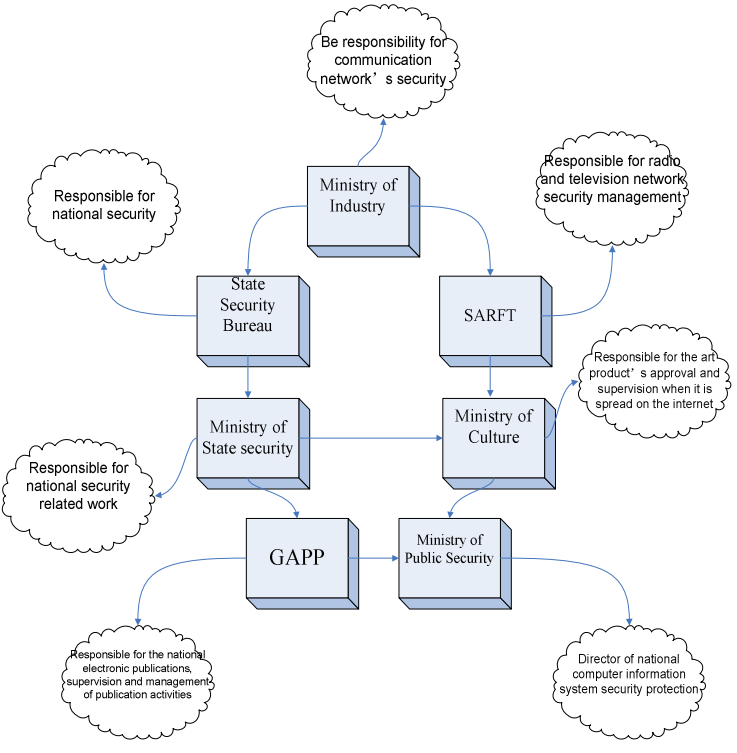


Fig. 1. Responsibility analysis chart involved in the agencies of Internet information management

3 Problems

Although the Internet's impact on society increasingly widespread,there are still a considerable number of government staff lack the knowledge to learn the consciousness of the Internet, do not have the internet experience, correct extraction and wrong extraction. They aren't familiar with the internet information transition rules, lack the network culture's construction and management. And Also know little about the governing country experience with the internet environment. Someone thought the internet is a virtual world, the problem on it couldn't have some important impact. This thought caused some local Internet culture construction and management of long-term are not strengthened; some hot face of online media do nothing to deal with slow, resulting in "small problem " into a "big Incident ", to cause harm to social stability and development.

Some network operators and Internet service providers lack of social responsibility, one-sided pursuit of economic interests, *ignoring* the spread of Internet information's Social benefits, shifting regulatory responsibility, resulting in management deficiencies, and allowing the spread of harmful information.

Some Internet users affected by the Internet's wrong concepts, such as "virtual space, free world and so on". Weak the rule of law and morality, is not responsible for statements made at random, spread rumors and false information, interference with

Internet information and communication order. Overall primarily in the following 8 questions:

1) Regulatory responsibilities allocation is unclear, there is a management blind spots.
2) Regulatory overlap management, \cross-management, and the distribution of responsibilities is unclear.
3) Parts of the management staff lack the enthusiasm for study. And their response always slow when the network events occurred. They don't know how to handle it.
4) The lack of internet culture and moral, a number of Internet users is not responsible for statements made at random, spread rumors and false information.
5) Some involved information security firms lack the responsibility, one-sided pursuit of economic interests, shifting regulatory responsibility, laissez-faire dissemination of harmful information.
6) Out of line management responsibilities and management tools. The departments are not responsible for the management of resources and management tools to master, leading to inefficient management
7) Part of the management "control center " duplication in serious condition, the technical level is not high, resulting in duplication and waste of resources.
8) *part* of the administration response to the lack of experience in multi-sectoral coordination to deal with the network to be multi-event coordination and inefficient

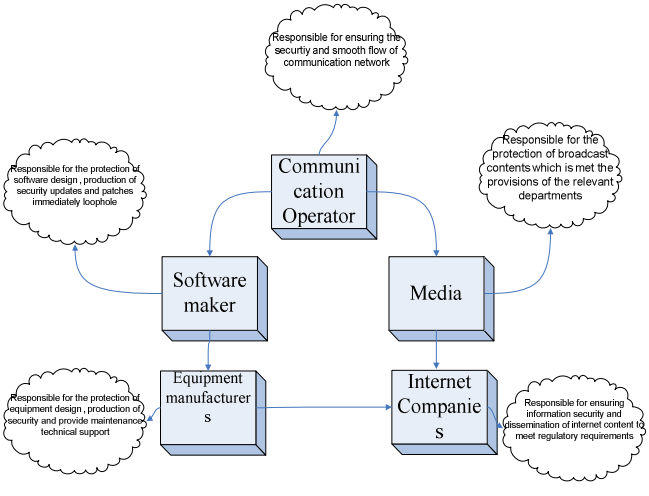


Fig. 2. The enterprise involved in internet security governance's allocation figure

4 An Optimized Organizational Structure Model of Internet Security Management

Internet content regulation related to the interests of all parties involved. It usually including the control subject, control objects and control aims.

Among them, the control of the main content on the Internet refers to the controlled unit. The control object is the content be controlled. And the control aim is the control subject's unit or persons.

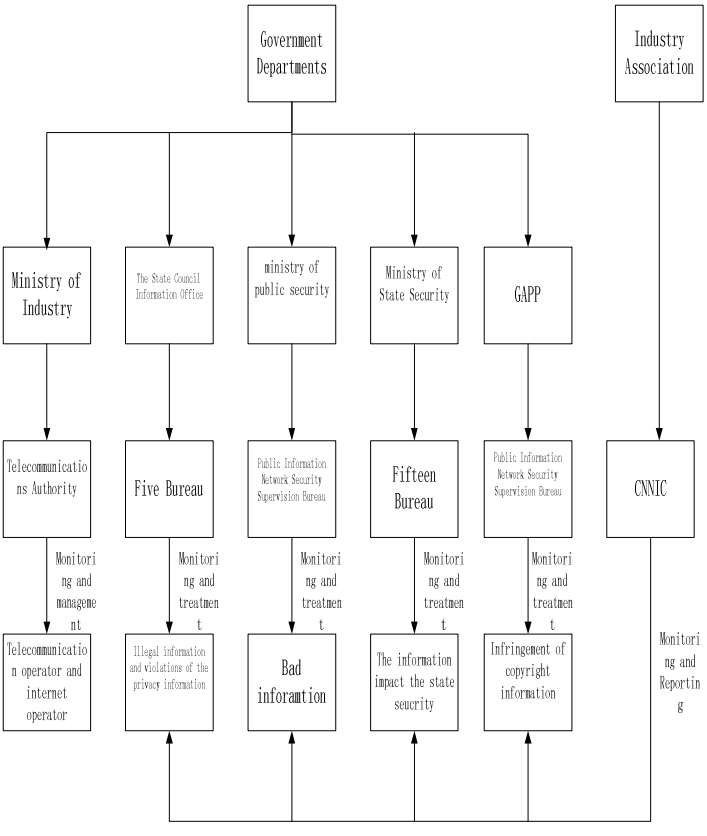


Fig. 3. The improved organizational structure for internet security management

From Chinese current situation, Chinese Internet content regulation subject, object and object include:

Control subject: Judicial branch of government, industry regulators and government authorized the establishment of trade associations.

Control object: The spread content on the internet which is harmful for individuals, businesses and state.

Control aim: Control by restricting Internet content and regulate the laws and the person or business subject, including the network of communication, the contents of the individual or business operations.

For the control subject have some harmful problem such as the main leadership of the existence of multi-sectoral, functional overlap and other defects.

So I think that should be based on the principle of division of functions will be further defined the functions of the control subjects, The control of the main draw of the Internet content regulation in different responsibility and authority to achieve the full range of Internet content ‘s effective control. in China. The figure below are the improved organizational structure.

5 Conclusions

In this paper, an improved organizational structure for internet security management is proposed. It improved the existing internet security management framework. The improved framework not only possesses a better internet management merchandise, but also has a new method for internet security management. It has a positive impact on national’s internet policy making.

References

[1] Rime, B., Philippot, P., Boca, S., et al.: Long-lasting cognitive and social consequences of emotion: Social sharing and rumination. In: Stroebe, W., Hewstone, M. (eds.) European Review of Social Psychology, pp. 225–258. Wiley, Chichester (1992)
[2] Lee, T.: Internet Control and Auto-regulation in Singapore (2005)
[3] Rodan, G.: The Internet and Political Control in Singapore (1998)
[4] Mitchell, D.: Internet Content Regulation in South-East Asia: Directions
[5] Drezner, D.W.: The Global Governance of the Internet: Bringing the State Back. In: Political Science Quarterly, vol. 119(3), p. 477. Academic Research Library (Fall 2004)
[6] Flint, D.: Regulation of Content on the Internet- An Australian Perspective
[7] Mifsud Bonnici, J.P., De Vey Mestdagh, C.N.J.: Right vision, wrong expectations: the European Union and self-regulation of harmful Internet content. Information & Communications Technology Law 14(142), 142 (2005)
[8] Communication to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions, Illegal and harmful content on the Internet, COM (1996) 487 final (October 16, 1996)

Study on the Business Model of Tourism Electronic Commerce in China

Jian Song

Shandong University of Finance and Economics, Jinan 250014, China
songj@sdfi.edu.cn

Abstract. Successful tourism electronic commerce must have successful business model. In this paper, the author first analyzes the status of electronic commerce models of tourism. Then it makes an analysis of current development of business model of tourism electronic commerce in China from four aspects, namely strengths, weaknesses, opportunities and threat. Besides, the author concluded the optimization principles of the tourism electronic commerce business models and the development trends of the tourism electronic commerce and gave some theoretical advices to our enterprises.

Keywords: tourism electronic commerce, business model, profit model, SWOT analysis.

1 Introduction

China, which has a historical culture and rich tourist resources, is already becoming not only a major tourist destination in Asia but also traveling powerful nation in the future world. Since up to the 1990s, the union of tourism and electronic commerce caused tourism electronic commerce. Its rapid fusion has created incredible value which we can not imagine. Its perfect union will certainly rapidly impel the development of electronic commerce and the traveling industry [1].

Success electronic commerce must have the successful business model. Business model of tourism electronic commerce is a new tourism operation way and the profit pattern based on information technology. Along with the vigorous development of tourism electronic commerce, the research on business model of tourism electronic commerce has vital practical and theoretical significance.

2 Tourism Electronic Commerce Status in Quo

As a new business model, tourism e-commerce can be traced back to the 1990s. It has developed rapidly because of its high efficiency and low cost. Many tourism e-commerce websites have became a considerable scale in North America, Japan and some European countries with the rapid development of international tourism recently.

2.1 Status in Quo of Tourism Electronic Commerce Abroad Structure

Online tourism electronic commerce in America has maintained a leading position in the world. Since 2007, its online travel transactions become than the core of the value

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 959–964.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

chain. Reports from IResearch shows that the scale of online travel sales grow 4.6% in America. And the value is 92.5 billion U.S. dollars. It is predicted by e-Market that online travel sales in 2011-2014 will continue to grow steadily in America. In 2014, it is expected to reach 119 billion U.S. dollars. Its online travel industry is not only a high degree of integration and the impact force, but also some of new technology and new business models.

Although tourism e-commerce in Europe is a few years later than the United States, it keeps up the pace of the U.S. tourism market. Danish Research Center for Regional and Tourism in Denmark found that the European online travel sales have reached 40.2 billion pounds in 2010, with an increase of 14%. It also predicts that total 2008 online sales will reach 44 billion pounds in 2011. The development of tourism e-commerce in Europe has showed the business combination trend of tourism electronic commerce, which is consistent to political and economic integration in Europe. Such as the largest travel information portal www.lastminute.com in United Kingdom and the largest hotel group Accor Group Travel Hotels in France become Europe's largest hotel distribution providers through the Global Distribution Systems.

The development of tourism electronic commerce in Japan is in a leading position in Asia. It put forward "tourism nation" to increase inbound tourism in recent years. They are actively exploring tourism electronic commerce from national to local. Now it has reached a certain scale in Japan. From the business model point of view, the tourism e-commerce has become a more mature B2B model (site of transport, accommodation, attractions and other business), B2C model (for tourists), C2B model (tourists "A la carte ", the enterprise bid disk access) and C2C model(visitors initiated the recruitment of their own group with the fan) [2].

2.2 Status in Quo of Tourism Electronic Commerce in China

Compared with foreign countries, tourist e-commerce started later in China as professional travel sites began to appear in 1996. At present, there are more than 5,000 travel information web sites in China, among which are more than 300 professional travel sites. Those web sites include regional sites, professional sites and portals on travel channel. According to the report "2007-2008 Development Report of Chinese online travel reservation" from IResearch shows that in 2007 Chinese online travel booking market reached 2.27 billion, compared with 47.5% increase over 2006. According to "2009-2010 Chinese online travel booking industry research report" from IResearch shows that in 2010 Chinese online travel booking market was 4.14 billion, compared with 3.74 billion yuan in 2009 an increase of 17.2%.

3 Hierarchical Business Models of Tourism Electronic Commerce

Tourism electronic commerce is for people who like to travel. These people include tourists, business people and all travel to a strange city to another. At present the main business model of tourism e-commerce is B2C and B2B2C [3].

3.1 B2C Model

This business model mainly is oriented to individual tourists. First individual tourists can query information through the network. Then they design their tourism calendar of

events, booking hotel rooms and joining a tour group. This model can be convenient for visitors to remote search and booking travel products. Also it can overcome the information asymmetry and incomplete information in tour market.

For example, the Spring and Autumn web site. Its profitable model is consisted of website of the tour, upstream of tourism enterprises, cooperation around the branch travel agency, airline ticketing agents, destination hotel and Internet market composition. Target market of the website is for tourism and holiday visitors. Profit of the Spring and autumn travel constituted mainly by the following aspects:

- Line booking agency fee, which is main source of profit of spring travel. It is formed in groups of the Spring International's and is got through the CST in the form of access.
- Hotel reservations agency fee. Customers can have two kinds of payment. One kind is prepaid. The other is the way around the payment. The website also has two profit channels. The former gets a profit by the return of Spring and autumn travel. And the other way is based on the return of profits realized in the form of destination hotels.
- Ticket booking fee. It is a way to get profit through price difference in the return of the post to achieve profitability.
- Spring International Development provided funding.

3.2 B2B2C Model

The model is also for the tourist consumers and travel businesses and tourism. Travel websites integrate consumer information of market information and visitors. Then they support suppliers and consumers connected to provide cash flow of and information flow of technical by electronic commerce. Websites has become the core of the whole industry chain nodes under this model, which is the most modern commercial spirit of the business model since the 20th century. The tourism electronic commerce is the largest travel agency who can response to customer service quickly and timely.

The most representative of this model is Ctrip. The profit model of Ctrip is mainly by its site, upstream tourism enterprises, destination hotel, airline ticket agents, travelling agents and netizen markets. Its target market is mainly for business travel clients.

As an intermediary e-commerce site, most revenue of Ctrip comes from intermediary business fees. The main source of its profits is:

- Hotel reservations agency fees. This fee, which is basically discount get from the destination hotel, is the main source of profit in Ctrip.
- Ticket booking fee. It is equal to the difference booking fee from airline ticket and from the customer.
- Travel fee in the hotel, air ticket booking fee and insurance fee. This income is used to return and post profits.
- Line booking agency fees. A number of travel groups are organized by Ctrip and other travel agencies.
- Advertising revenue. Ctrip also undertook the 21CN, Shanghai Hotline, CCTV, and other famous portals on the Travel Channel, which has considerable advertising revenue.
- Membership income. Members of Ctrip can purchase VIP membership card directly. Meanwhile they can get more price concessions.

In the above two kinds of patterns, the traveling website has played the irreplaceable role in the information procession. It can supply all kinds of tourist rich information and processing to commercial information.

4 Swot Analysis of Business Model of Tourism Electronic Commerce

Tourism electronic commerce has made great progress in China since it came into being. Although it can provide tourist information and travel experience for users, it also face great challenge. So we must sum up the current existence of tourism e-commerce industry and analyze advantages and disadvantages, opportunities and threats of tourism electronic commerce.

4.1 Advantages

- Convenience

Because tourism electronic commerce has overcome the tradition tourist service disappointment, it may provide the comprehensive destination information for the consumer and the destination preview for the decision-making. Moreover it may also solve the service prestige problem and provides all-weather cross region for the consumer the service not limited to time and space. Customers can pay their bill with the related bankcard [4].

- Euroky

Traveling website can transmit traveling information o global each corner with the aid of the Internet. Through the network consumers can access the net to obtain tourist business's material and the product information. Also tourist business can also develop marketing activity from all over the world through the network.

- Individuality

With the international traveling individual already arrived, the goal of individual traveler is not only going sightseeing, but also used to comfortable, free and regressive. The tourist market needed to transform the sightseeing tour to the personalized traveling. Tourism electronic commerce should also satisfy the entire populace to travel change through network.

- Preferential

While provides the convenience and quick service, tourism electronic commerce can bring the actual benefit to the consumer. These network hypothesized tourist service company not only causes consumer to complete the booking in any time to buy tickets, but also has the preferential benefit discount.

4.2 Disadvantage

Although tourism electronic commerce develops quickly in our country, it still had the very big disparity compared with the overseas traveling website.

- The website information is obsolete and the service items are unitary.

The website information is obsolete and the service items are unitary. Although the majority traveling websites all have the introduction to the scenic site and the hotel, the

content is not comprehensive and renews slowly, which has very difficulty to attract the customer's attention and interest.

- The website localization is inappropriate and the market subdivides not explicitly.

Now most traveling websites in our country take the website as a core, which leads to the website lack of communication. They can not make good use of the network resource to reduce cost of operation. Moreover, those traveling website can not meet personalization demand of customers [5].

- Website construction is redundant and network development is slow.

Our travel websites are often a replica of the foreign tourist site, which cause a pattern similar and redundant construction. Besides the whole travel industry also have no real network. In our country a broad sales and service to consumers have not formed.

- Electronic commerce software environment is poor.

4.3 Opportunity

Along with the Internet technology's rapidly expand, the tourism faced with the new opportunity. Many investors aim at the opportunity to enter the tourism electronic commerce market. The Ctrip traveling network takes the tourism electronic commerce through connects the upstream destination hotel, the airline and the downstream trade route customer. Though this can it provides each kind of tourism products and the traveling information to the downstream trade route customer. The experts pointed out that although tourism electronic commerce market faced with paroxysmal growth opportunity, but it still had many traps. Only when we are ready in the technology, the service, the brand as well as the channel can tourism electronic commerce hold its opportunity [6].

4.4 Threaten

Looking from the global scale, tourism electronic commerce in our country is still at elementary step. There are many problems in tourism electronic commerce regardless of from technology, service or business model. With the Chinese economy's fast development, the domestic tourist market also entered the high growth stage and the traditional traveling business model has been unable to satisfy the need of market. The tourism electronic commerce need reforms, only then can promote our country tourism electronic commerce.

5 Conclusion and Future Considerations

Through above SWOT analysis of business model to the tourism electronic commerce, we should follow the following several principles during construction of tourism electronic commerce in our country.

- Integrate tourism resources.

It is necessary to carry on the conformity and the strategic alliance among our traditional tourist business and the emerging traveling website in China. As for small tourist business they should be clear *about* market objectives, conformity technical superiority of traveling their agent and resources superiority of their own, which can be

benefit for on-line market. As for large-scale tourist business they should unify highly with Internet, establish electronic retailing system oriented to business agents and online sales system oriented to tourists for the sake of the foundation, consolidates and develops own brand. Only through this can they realize the collectivization and network management.

- Change management idea.

At present, our country traveling website should transform the idea of taking the transaction as the center to taking the service as the center. In order to reduce the distance with international tourist market, they should readjust the design proposal and the service mode continually.

- Subdivide the market.

The accurate localization market provides the individual consumption service and the enterprise service to the consumer. They also provide service differently according to the consumer's demand, which *includes* in the view of both the high-end expense community's service and the popular consumption community's service. Meanwhile the websites must provide the consumer route personalization tourist services and self-service design or local product purchase in order to meet the personalized need.

- Introduce new technology.

The comprehensive applications of mobile payment, short information service and global positioning system will bring a new revolution to the tourism and even the traveling electronic commerce. Regardless whenever and wherever the customer is, they can gains each kind of information through terminals mobile phone. The application of mobile electronic commerce technology will cause tourism electronic commerce service more perfect.

- Consummate macroenvironment.

Based on the traditional profession, we should establish Chinese traveling database, including traveling service information. We should provide information issue and inquiry for tourist, tourist business and traveling management organization. The government should consummate the software and hardware environment of tourism electronic commerce.

Tourism e-commerce is a general trend in the world. The combination of tourism and e-commerce is an original *creation* in e-commerce. Because of its less of physical transport it can be accepted easier by consumers. In short, a successful tourism site must have a successful business model. Tourism electronic commerce will provide differentiated value to customers by expanding and strengthening in our future.

References

[1] Wang, S.: Travel e-commerce: Two profit model travel sites. Managers 13, 121–123 (2008)
[2] Feng, M.: E-commerce development in China status quo of tourism. Modern Business 32, 50–51 (2009)
[3] Li, Z.: Development and key constraints of of Tourism e-commerce. Reform and Strategy 7, 130–133 (2009)
[4] Ting, L.: Our business model of mobile commerce. Journal of The Second Northwest University for Nationalities 3, 123–127 (2009)
[5] Chang, W.: Revisited travel e-commerce business model. Market Modernization 33, 140–141 (2007)
[6] iResearch. 2010Q1 Revenue of China's online travel booking market size of 11.9 billion yuan. [EB/OL] (May 24, 2010), http://news.iresearch.cn

Semantic-Based Wireless Video Surveillance System

Haijun Gu, Wei Liu, and Xiaohui Zhao

Department of Communication Engineering Jilin University Changchun, China
liuvivi512@163.com

Abstract. The technology of semantics is a hot research topic in recent years. Especially, it is extensively used for information retrieval and classification. For most of the traditional wireless surveillance is based on store-and-forward model, it is less active and inefficient in data retrieval. In this paper, we propose the wireless active video surveillance based on the semantic technology, and the basic model we used is shown. The wireless transfer model is based on pear-to-pear and the analysis model is object-oriented. In practice, we design the platform based on DSP and FPGA for image processing and storage. Then we recognize the objects by assembling difference image, and classify the objects according to their different features base on semantics ontology library we proposed. The user can retrieve the image by semantics ontology conveniently. In this paper, we describe the principle of ontology construction in detail. Finally, we have made experiment for proposed algorithm and get satisfied results. The rate of wireless transmission can reach 10-15 frame/sec under the condition of EDGE network, furthermore semantic technology greatly improve the efficiency, the accuracy rate of retrieval exceeds 96%.

Keywords: Semantic, Pear-To-Pear, Video, Retrieval, Wireless Video Transmission, Active Surveillance.

1 Introduction

Wireless video surveillance is a challenging research issue at all times. Generally, video analysis is mainly based-on low-level feature. However, with the development of semantics, it is growing interest in the semantic concepts understanding of video data [1].Therefore recent researches work on multimedia have shifted from low-level feature-based approaches to high-level semantic analysis[2]. Traditional video surveillance or DVR only record the interesting image and retrieve them by time, however, it generally cannot retrieve the image according to different features of objects. Semantic-based wireless video surveillance is an attempt of new approach.

In this paper, we designed a method to compute the object semantic in the interesting scene [3]. At first, we describe the basic model of system framework, including pear-to-pear wireless data transmission and object-oriented model analysis [4]; we emphatically introduce the analytical method based on semantics [5]. Then, we describe our system in detail. On one hand, we describe the hardware platform designed by us, composed of DSP and FPGA, which can realize real-time image process and storage [6, 7]. On the other hand, we detail the approach and algorithm, including the algorithm of object recognition [8], the method of object description [9], the approach

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 965–973.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

of object classification based on semantics [10]. Finally, the experiment results are given, the results show that we can achieve the wireless video transmission, and the retrieval efficiency is improved greatly. The semantic technology for information retrieval improves the activity of the surveillance.

2 The Basic Model of System Framework

2.1 Pear-to-Pear Wireless Data Transmission

Considering the mobile IP is dramatic and internal. Generally, wireless data transmission is based-on store-and-forward model. For the image data is stored by the servers, the users' privacy cannot be ensured. However, the monitoring system we proposed is based-on pear-to-pear model, the server is mainly used for address translation, and the security is enhanced.

2.2 Object-Oriented Modal Analysis

The real world can be recorded authentically by image data. Usually, we describe the real-world models in object-oriented form, so object-oriented model analysis, describing objects as properties, is very significant.

- Description of objects

Everything can be viewed as objects in the real world. Then the objects can be described as properties. The properties of objects are divided into static property and dynamic properties, as shown in figure1.

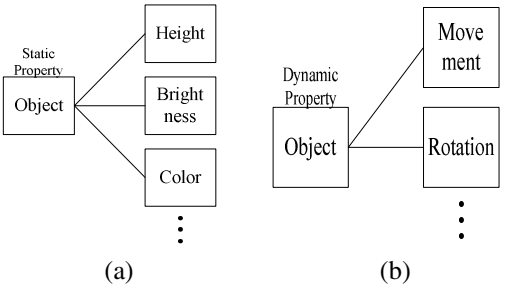


Fig. 1. Example of Description (a) static property (b) dynamic property

Here, the static properties are intrinsic, for example color, brightness, etc. The dynamic properties are related with the time domain, such as movement, rotation, etc.

- Object Recognition

In this paper, we design the method of object recognition, especially the moving objects. For moving objects, we propose the approach based on assemble difference image. The assemble difference image is acquired through comparing the real-time image with the template continuously. This approach can be used for moving objects

detection. In addition to shape feature and movement feature of the objects, we can distinguish human and automobile clearly.

- Construct Semantics Ontology

All objects are classified to be understood in human understanding; similarly, the objects in the real scene are classified as different classes. We can classify the complex scene clearly by adding semantic technology which helps the user in his/her image interpretation task.

In this paper, we classify the objects which have the same or similar properties as a group. Then we construct the semantics ontology. The principles of construction are similar to the object-oriented programming, such as inheritance, subsidiary. At first, we recognize the objects, distinguish human and automobile. Then classify the human and automobile as different subclass. For all the objects in video have inheritance, classes of all objects can be organized by the tree structure, as figure 2 shows.

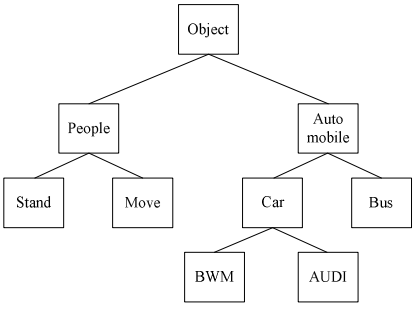


Fig. 2. The Diagram of Tree Structure

As the instance of figure 2, the people and the automobile are high level classes. For there has unique properties such as size, length, seats and so on, the car and bus are seen as the subclass of automobile class. So, the automobile class is called their base class. Similarly, the car class can also divide into several subclasses. As human understanding, we identify the things in a top-down process, from object to automobile, then to car to specific types. When a person watch a scene, he will find the object at first, secondly he will recognize that it is a people or an automobile, then will he recognize it as a bus or a tractor, even he is possible to recognize a certain brand of the bus.

As people are seen as a whole, it can be divided into some litter objects, such as gesture, movement status (move or stand) and so on. Movement status analysis can be seen as an independent analysis unit in some video analysis. The movement status object can be seen as the subsidiary of people object.

3 The Design of System

This system we designed is composed of hardware and software. The hardware is mainly used for the image process, and the software includes the classification based on semantics, data compression, image process and video display based on mobile phone. The structure of the system is shown as figure 3.

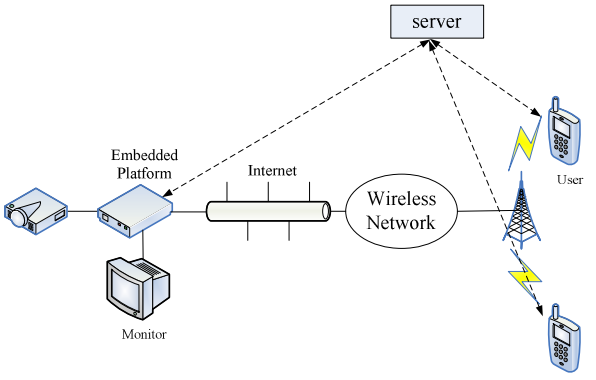


Fig. 3. The Structural Diagram of the system

3.1 The Design of Hardware Platform

The Image Process Unit is mainly used for the image storage, objects recognition and image retrieval, etc. We propose the FPGA and DSP scheme, FPGA is a better computing platform than PC to perform data-intensive video processing, especially the parallel processing capability; DSP is easier to perform the complicated algorithm and connect external unit and easy to develop the drivers.

We choose the Virtex-4 series high performance FPGA and C6000 series high-speed DSP of TI. They communicate to each other via the EMIF interface. The FPGA is used for image preprocess, including image filter, image binarization and high-speed image storage. The DSP is used for image recognition and retrieval. The basic structure of the platform is shown in figure 4.

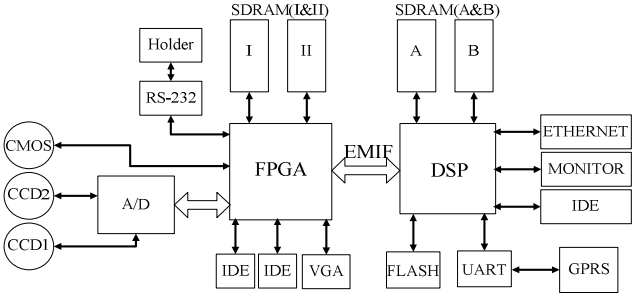


Fig. 4. The Structural Diagram of the platform

3.2 The Framework of Algorithm

- Object Recognition

Considering that people or automobile we interest in is mostly moving. According to the principle of assemble difference image, we need some difference images. In order to achieve better experiment results, we preprocess the image before recognition,

including image filter, image enhance, etc. In this paper, we use the shape feature, such as length, width and height, to distinguish people and automobile.

- Establish Ontology Library

At first, we summarize classification of the objects which are often seen in the image, such as people, automobiles, plants, and so on. In this paper, we are mainly interested in people and automobile. Secondly, considering that different objects can waken different interests on brain. Similarly, the ontology library is divided into different classes according to different types or feature. Finally, all the classes of the library are given a code, which is similar to the IP address. The code is called the unique identification code (UIC).Considering the capability of the platform, the description of UIC is similar to OWL, but it is much simpler.

Here, we use the UIC to distinguish the two classes. Sometimes, we are interested in one of the classes, and the interesting object can be set a high weight coefficient.

According to the description of section 2, the classes are hierarchical. Some subclasses are the inheritance of their base class, and the others are the subsidiary. The basic structure of the code definition is shown as figure 5.

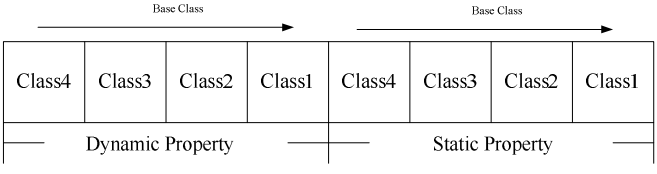


Fig. 5. The basic structural diagram of UIC

So the rule of the code definition is as follows:

For inheritance relationship, the code of automobile is NNNNNNN0 (there, N is the random data.), the code of car should be NNNNNN00; in the same way, the bus should be NNNNNN10. The last data '0' represents the automobile class.

For subsidiary relationship, the code of people is NNNNNNN1, the code of gesture class of people is NNN0NNN1, in the same way, and the code of movement status class of people is NNN1NNN1.

If there are some more interesting objects, we can define weight coefficient for each other. Thus we can recognize the object we interest. For example, if we are interested in the car class, we should define the weight coefficient of the car class as '2', the weight coefficients of others are defined as '1' or '0' according to the interest level.

- Object Classification

In this paper, the proposed ontology library can be used for object classification. We also propose the approach base-on learning mode. Specifically, in order to meet the real-time processing, we keep our algorithms as simple as we can. So we abstract the feature of the objects and then establish the feature library as the template, related to the class. Like this, the following object can be classified well and truly according to the feature library.

- Semantics Retrieval

Semantics retrieval is the key of video analysis in this paper. The object should be extracted through different ontology. The retrieval of image data should achieve the objective that finds the image according to specified object classes.

In this paper, we recognize the objects and sort them into some classes. Then we store the images in addition to the information of the semantic, make an index which provides a facility for retrieving. The basic structure of stored data is as figure 6 shows.

Fig. 6. The Basic Structural Diagram of Stored Data

- Data Compression

In wireless mobile environments, the bandwidth and client capability are limited. So the data transferred in wireless network must be compressed. In order to make the method of compression more suitable for real-time environment, we choose the M-JPEG decoding algorithm instead of MPEG-4.The comparison of the two algorithms is as table 1shows.

Table 1. The Table of Comparison

Algorithm	Complexity	Compressibility
MPEG-4	high	high
M-JPEG	low	Medium

3.3 Experiment Results

We test the proposed algorithm on the designed platform. The experiment mainly includes following several steps.

3.4 Objects Recognition and Classification

At first, we set the image mode as monochrome, and resolution as 320×240. Then we recognize the objects based on assemble difference image, and classify them based on semantics.

The scene we chose is the road, and the objects are mainly people and automobile. The results are as follows:

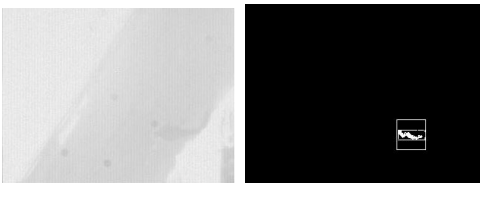


Fig. 7. The Image of People Recognition

As figure 7 shows, we recognize the people accurately. Since the object is moving, so we sort it into people class and moving subclass. Define the UIC as '00010001'; here we define the weight coefficient as '1'.

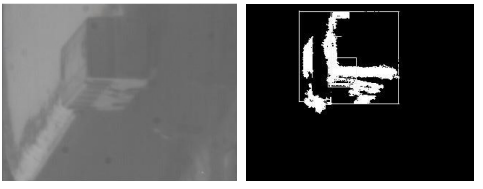


Fig. 8. The Image of Automobile Recognition

As figure 8 shows, we recognize the bus accurately, too. Since the object is a bus, so we sort it into automobile class and bus subclass. For it is moving, we define the UIC as '00010010', here we define its weight coefficient as '1'.

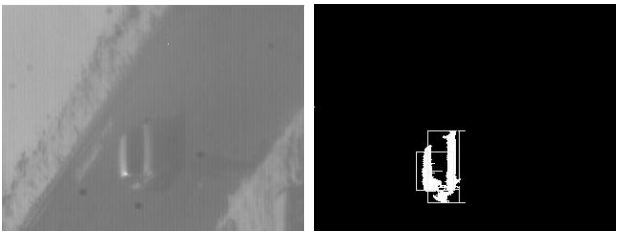


Fig. 9. The Image of Interesting Object Recognition

As figure 9 shows, there are two objects, a car and a people. Here we define the car's weight coefficient as '2', and define the people's weight coefficient as '1'. We recognize the car accurately, ignore the people. Since the interesting object is a car, so we sort it into automobile class and car subclass. For it is moving, we define the UIC as '00010000'.

As a whole, we recognize 1500 objects, the results are as table 2 shows:

Table 2. The Table of Results

	Total	Retrieved	Accuracy Rate
Object	1500	1446	96.40%
Automobile	896	885	98.77%
People	604	561	92.88%

3.5 Real-Time Image Storage

In this paper, we store the real-time image **into** the IDE disk. The IDE disk is driven by FPGA. The sequence chart is as figure 10 shows.

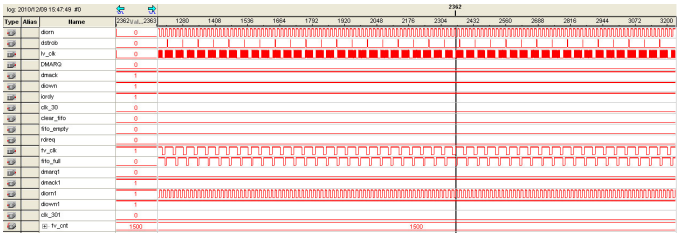


Fig. 10. The Sequence Chart of Image Storage

3.6 Wireless Video Surveillance

Considering the mobile network environment, we set the image model as color, and set the resolution as 176×144. We design the software with J2ME. We can display real-time video and choose the interested image through semantic index. The results are shown in figure 11.

Fig. 11. The Display of Mobile Telephone

Here, the left one is the real-time video. The rate of wireless transmission can reach 10-15 frame/sec under the condition of EDGE network. The right one is the retrieval mode. Here the UIC is '00010001', it is the people class and moving subclass.

4 Conclusion

In this paper, we propose some new approaches of video retrieval, and apply them to the wireless video surveillance system. Then we have made experiment for proposed algorithm and get satisfied results. The experimental results demonstrate that the system meets the real-time requirement and semantic technology greatly improves the efficiency, the accuracy rate of retrieval exceeds 96%.

References

[1] Li, B., Sezan, I.: Semantic sports video analysis: approaches and new applications. In: Proceedings of 2003 International Conference on Image Processing, Barcelona, Spain, pp. 17–20 (2003)

[2] Shen, L., Zhang, J.: Image Retrieval and Compressed Domain Processing. Lixs & Telecom Press, Ziaoguang, Beijing (2008)
[3] Aphade, M.R., et al.: Modeling semantic concepts to support query by keywords in video. In: Proc. IEEE Int. Conf. on Image Processing, ICIP 2002, NY, vol. 1, pp. 145–148 (September 2002)
[4] Corridoni, J.M., Del Bimbo, A.: Structured representation and automatic indexing of movie information content. Pattern Recognition 31(12), 2027–2045 (1998)
[5] Mezaris, V., et al.: Video object segmentation using Bayes-based temporal tracking and trajectory-based region merging. IEEE Trans. Circuits Syst. Video Technol. 14(6), 782–795 (2004)
[6] Bei, N., Wei, Y., Shi, G., Ye, J.X.: Developing a Smart Camera for Road Taffic Surveillance, pp. 826–831. IEEE (2008)
[7] Dias, F., Berry, F.: Hardware Design and Implementation on a FPGA-Based Smart Camera, pp. 20–25. IEEE (2007)
[8] Jiang, D.: The Research and Implementation of Target Trackong Algorithm Based on Mean-shift. Chongqing University, Chongqing (April 2009)
[9] Hanjalic, A., Xu, L.Q.: Affective video content representation and modeling. IEEE Transactions on Multimedia 7(l), 143–154 (2005)
[10] Miao, G., Zhu, G., Jiang, S., Huang, Q., Xu, C., Gao, W.: A Real-time Score Detection and Recognition Approach for Broadcast Basketball Video. In: IEEE International Conference on Multimedia and Expo., pp. 1691–1694 (2007)

Ultra-Low Power CMOS Charge-Sensitive Preamplifier

Zhou Yunbo[1,*], Yu Zongguang[2], and Yang Yu[2]

[1] School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China
[2] China Electronic Technology Group Corporation the 58th Research Institute, Wuxi 214035, China
ic_ybzhou@163.com

Abstract. An ultra-low power complementary metal-oxide-semiconductor (CMOS) charge sensitive preamplifier has been built using a CSMC 0.5μm DPDM process to achieve the ultra-low power dissipation requirement for portable digital radiation detector. The ENC noise of 363ē at 0pF with a noise slope of 23ē/pF can comply with the stringent low noise requirements. A 100mV/fC conversion gain at 20pF has been obtained. By operating the charge sensitive preamplifier works in the weak inversion region, the power dissipation is only 65.5μW (3.0V).

Keywords: Charge Sensitive Preamplifier, Weak Inversion Region, Random Telegraph Noise.

1 Introduction

When a Si semiconductor detector is used for the measurement of soft X-rays and low to high-energy gamma rays, the output signal is a weak charge pulse which pulse width is several tens of nanoseconds. To integrate the weak charge pulse and convert it into voltage pulse charge sensitive preamplifier with feedback capacitance is commonly used. [1~4][8]

Additionally, increasing demand for battery operated portable digital radiation detector has led to the stringent requirement for circuit designs to be more power aware. [5][6]

In this paper, a charge sensitive preamplifier (CSA) optimized for CMOS front-end readout ASIC is presented. Using wide channel devices for the input stage can easily make them bias into weak inversion region and get some extra advantages [5]. MOSFET working in weak inversion region can obtain higher voltage gain. Also because of low value of quiescent drain current, device has less energy consumption. Another advantage is that it has much better linearity than in strong inversion region. These features have been used in several early power-efficient designs. However, low frequency noise behavior for weak inversion operation is different from that for strange inversion operation [7]. So some design tradeoffs has to be used for optimizing the signal peaking time and equivalent noise charge (ENC) at low bias current. The design specifications are given in Table 1.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 975–982.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Table 1. Main design parameters of CSA

Parameter	value
Power supplies	3.0V±10%
Threshold range	160mV~200mV
Bias current	I_{bias}=1μA
Capacitance of detector	C_d=20pF
Feedback capacitance	C_f=150fF
Power consumption	P_{tot}= 70μW
ENC	400+30ē_/pF
Gain	100mv/fc
Linear dynamic range	2fc~200fc

2 Charge Sensitive Amplifier Design

2.1 DC and AC Analysis

The charge sensitive preamplifier, which acts as an interface between detector (input pulse signal) and the subsequent signal processing circuits, is shown in Fig. 1.

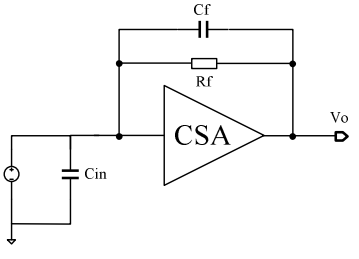


Fig. 1. Charge Sensitive Preamplifier

In this circuit, the feedback capacitance is used to collect the charge from the detector. Besides, the feedback resistor, known as a direct current feedback connected, is placed to discharge and stabilize the DC operating point of the amplifier. Consequently, if the DC gain A_0 is large enough, the output voltage can be expressed as

$$V_o = -\frac{Q}{C_f(1+\frac{1}{sR_fC_f})} \tag{1}$$

Since C_f is a constant, Vo is only related with the total charge Q. Therefore this circuit is implemented as a charge sensitive amplifier.

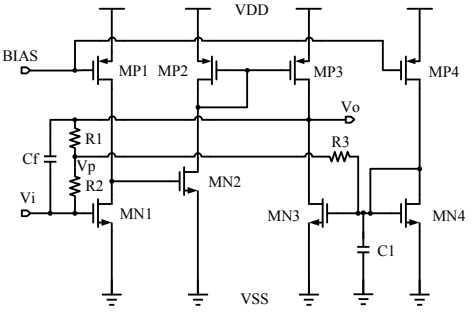


Fig. 2. Schematic diagram of CSA

The schematic diagram of CSA is shown in Fig.2. One can see in the schematic, a two stage cascaded single-ended structure is composed due to its high DC gain and low power. The gate of MN1 is the preamplifier input. The drain of MN4 is the preamplifier output. Feedback circuit consists of a Y-resistors network and a feedback capacitor Cf, which is able to provide two discharge channels.

Besides, the branch with R3 is used to set up gate bias voltage of input transistor MN1.

$$V_{G(MN1)} = \frac{R_1}{R_3 + R_1} V_p + \frac{R_3}{R_3 + R_1} V_o \tag{2}$$

When $R_3 << R_1$, Eq.(2) can be simplified as follows:

$$V_{G(MN1)} \approx V_p \tag{3}$$

As well known, MOSFET operating in weak inversion region generally has a high DC gain with an ultra low drain current [5][6]. Hence in our design, we choose input stage with W/L=300μm/1μm. The maximum conversion gain and minimum distortion bias point in this case is found from Fig.3 to be near 1μA.

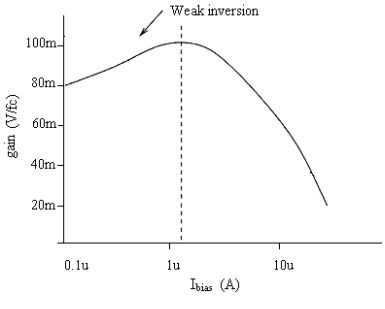


Fig. 3. Gain as a function of bias current for W/L ratio of 300μ/1μ

Depending on the analysis of small-signal equivalent circuit (Fig.2), the output changes to

$$V_o = -\frac{Q(1+s\tau)}{C_f(1+s\tau_1)(1+s\tau_2)} \tag{4}$$

Where the τ1 and τ2 are

$$\tau_1 = \frac{R_1^2}{R_3} C_f \tag{5}$$

$$\tau_2 = R_3 C_1 \tag{6}$$

This means the output voltage decays exponentially with two time constants.

The simplified small-signal equivalent circuit of the amplifier without feedback transistor and capacitor is shown in Figure.4.

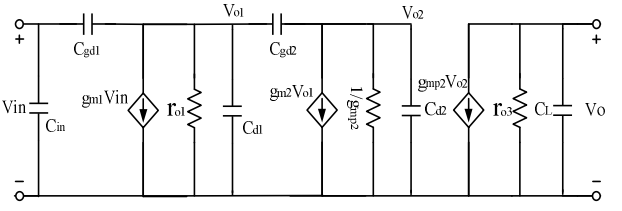


Fig. 4. Equivalent Circuit of the Amplifier

The open-loop gain $A_V = V_o / V_i$ is given by:

$$A = \frac{-g_{m1} r_{o1} g_{m2} r_{o3} \left(1 - s\frac{C_{gd1}}{g_{m1}}\right)\left(1 - s\frac{C_{gd2}}{g_{m2}}\right)}{\left[1 + s\left(C_{gd1} + C_{d1}\right) r_{o1}\right]\left[1 + \frac{s\left(C_{gd2} + C_{d2}\right)}{g_{mp2}}\right]\left(1 + sC_L r_{o3}\right)} \tag{7}$$

Where the DC gain A_0 is

$$A_0 = -g_{m1} g_{m2} r_{o1} r_{o3} \tag{8}$$

and the dominant pole P1 is

$$P_1 = \frac{1}{r_{o3} C_L} \tag{9}$$

Since the transconductance transistors are working in weak inversion region, the DC gain can be expressed as a constant which can be given by

$$A_0 = \frac{1}{\lambda_1 \lambda_3 (\zeta V_T)^2} \tag{10}$$

The channel-length modulation factor, λ, equals the reciprocal of the early voltage and is a strong function of channel length.

In addition, the smaller transconductance caused by a low bias current is result in that the right half plane zero will be probably at a relatively low frequency. However,

their frequencies are still greater than 10MHz without significant impact on the stability of this circuit.

2.2 Noise Analysis

The front-end readout ASIC' noise performance is usually expressed as the equivalent noise charge (ENC), which is defined as ratio of the total rms noise at the output of the pulse shaper to the signal amplitude due to one electron charge.

A significant effect that distinguishes weak inversion region operation from strong inversion region operation is low frequency noise behavior. In weak inversion region, low frequency noise is dominated by $1/f^2$ noise (here f is the frequency) [6][7].

The equivalent input $1/f^2$ noise source is given by

$$V_{eqi}^2 = \frac{K_{f^2}}{\left(\frac{C_{ox}WL}{g_m}\right)^2} \cdot \frac{1}{1+\left(\frac{f}{f_c}\right)^2} \tag{11}$$

Where K_{f^2} is the $1/f^2$ noise coefficient of the CMOS device working in the weak inversion region used [7]. g_m is the transconductance, and f_c the corner frequency. The $ENC_{f^{-2}}^2$ due to $1/f^2$ noise is obtained:

$$ENC_{f^{-2}}^2 = \frac{K_{f^2} f_c^2 C_t^2}{\left(\frac{C_{ox}WL}{g_m}\right)^2} \frac{\pi\tau_s^3}{q^2} \left| B\left(-\frac{1}{2}, n+\frac{3}{2}\right)\right| \frac{n! e^{2n}}{n^{2n+3}} \tag{12}$$

$C_t = C_d + C_f + C_{GS} + C_{GD}$, $C_{GS} = \frac{2}{3}C_{ox}WL$, $C_{GD} = C_{ox}WL_D$

And

$$g_m = \sqrt{2\mu C_{ox}\frac{W}{L}I_{DS}} \tag{13}$$

Where $B(x, y)$ is the beta-function, W, L and Cox represents the width, length and gate capacitance per unit area respectively of input transistor.

Optimum gate width exists for which the respective $1/f^2$ noise is minimal.

$$\frac{\partial ENC_{f^2}}{\partial W} = 0 \Rightarrow W = \frac{C_d + C_f}{(\frac{3}{2}L + L_D)C_{ox}} \tag{14}$$

3 Simulation and Test Results

Simulation has been performed with Synopsys hspice_vA-2008.03 simulator using the SPICE MODEL parameters of CSMC CMOS 0.5μm.

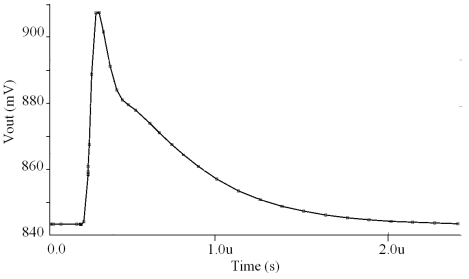


Fig. 5. Transient response at the output of the voltage amplifier

Fig.5 shows the transient response result of CSA for injected charges of about 6fc (C_d=20pF). A conversion gain of 100mv/fc is effectively reached.

Unlike semi-Gaussian pulse, the output waveform of CSA has two different attenuation time constants τ_1 and τ_2. The pulse come back to baseline rapidly because of the two different discharge path formed by R_1, R_2 and R_3.

The simulated transient response of CSA is given in Fig.6 for the input signals charging 2.67fc, 15.35fc, 28.03fc, 40.71fc and 106.8fc with applied to a 20pF capacitor at the preamplifier input.

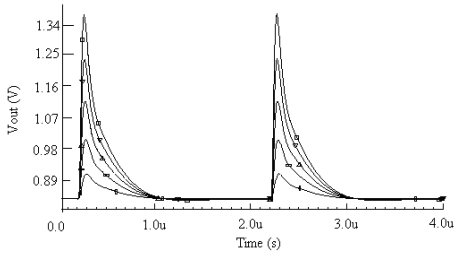


Fig. 6. Transient response at the output for five values of injected charge

The result shows that system has a great linearity within the scope of 2MeV. However, if we increase the input signal continuously which may cause a cut-off distortion in coping, system will go into a large-signal operating state. Fortunately within a certain range, the system still has a good linearity, and would not affect timing performance.

The optimized noise performance is presented in Fig.7 from which the ENC can be calculated. Considering a 100mv/fc conversion gain of the chip taken from Figure.4, the values of ENC is 363ē (C_d=0pF).

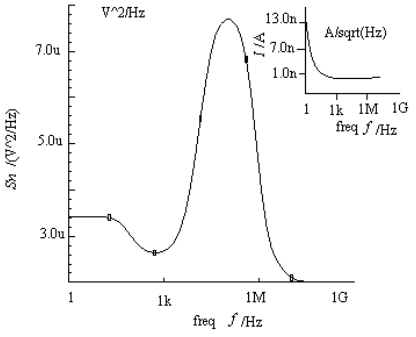


Fig. 7. Output noise spectral density

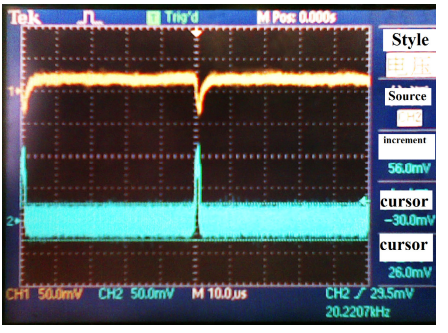


Fig. 8. Output waveform of the shaper and channel

The ultra-low lower CMOS front-end readout ASIC output signals are shown in Fig.8. The discrepancy between measured and theoretical results of the chip is always smaller than 4.4%.

As explain above, performance of charge sensitive preamplifier depends on model process and its structure, therefore, it is difficult to compare different circuits. Some results in recent literature are listed in Table 2 for reference.

Table 2. Performance comparison

	Process	Gain	Power/ channel	ENC
[3]	0.35µm	3.31mv/fc	1mW	382ē +21_/pf
[4]	0.35µm	2.81mv/fc	0.16mW	254ē +13.5_/pf
[8]	0.13µm	80mv/fc	0.6mW	351ē -4.8_/pf
This paper	).5µm	100mv/fc	0.1mW	363ē +23_/pf

4 Conclusions

An ultra-low power charge sensitive preamplifier, which is operating in weak inversion region, is described. Based on CSMC 0.5µm DPDM standard CMOS process, the whole system power consumption of 0.1mW with a gain of 100mv/fc is obtained. When input charge is 6fc and detector capacitance is equal 20pF, maximum count rate can reach 500Kcps, which is suitable for portable electronic dosimeter.

References

[1] Zhang, Y., Chen, Z., et al.: A Fully Integrated CMOS Readout Circuit for Particle Detectors. Chinese Journal of Semiconductor 28(2) (2007)

[2] Sansen, W.M.C., et al.: Limits of Low Noise Performance of Detector Readout Front Ends in CMOS Technology. IEEE Trans. Inform. Theory 37(11) (1990)

[3] Noulis, T., Siakos, S., et al.: Advanced Low-Noise X-Ray Readout ASIC for Radiation Sensor Interfaces. IEEE Trans. Inform. Theory 55(7) (2008)
[4] Noulis, T., Siskos, S., Sarrabayrouse, G.: Noise optimized charge-sensitive CMOS amplifier for capacitive radiation detectors. IET Circuits Devices Syst. 2(3) (2008)
[5] Comwe, D.J., et al.: Using the Weak Inversion Region to Optimize Input Stage Design of CMOS Op Amps. IEEE Trans. Inform. Theory 51(1) (2004)
[6] Gupta, S.K., Raychowdhury, A., et al.: Digital Computation in Subthreshold Region for Ultralow-Power operation: A Device-Circuit-Architecture Co-design Perspective. Proceedings of the IEEE 98(2) (2010)
[7] Leyris, C., Polorget, S., et al.: Random Telegraph Signal Noise Spice Modeling for Circuit Simulators. IEEE Trans. Inform. Theory (2007)
[8] Martoiu, S.: A Low Power Front-End Prototype for Silicon Pixel Detector with 100ps Time Resolution. IEEE Trans. Nucl. Sci. (November 14, 2008)

Reduction of Power Consumption in Wireless Sensor Networks for Railway Disaster Prevention and Safety Monitoring System

Tao Wu[1,*], Yusong Yan[2], and Xi Chen[1]

[1] Southwest Jiaotong University, No. 111, Erhuan Road, Chengdu and 610031, P.R. China
[2] Sichuan Normal University, No. 5, Jing'an Road, Chengdu and 610068, P.R. China

Abstract. Wireless sensor network has many advantages. One of the primary benefits is their independence from the wiring costs and constraints, which makes it is suitable for monitoring the environmental conditions around the high-speed railway. However, the energy of the sensor node is limited and railways are often located in remote area. If batteries have to be replaced frequently, the initial cost savings will be lost. Therefore, power efficiency is a critical requirement for wireless sensor networks. In this paper, two algorithms have been investigated for application of wireless sensor network in Railway Disaster Prevention and Safety Monitoring System, which are Advanced Direction-aware Algorithm and Cluster Head Selection Algorithm based on the Dynamic Energy Threshold respectively. Theoretically, both algorithms can effectively extend the lifetime of the wireless sensor network.

Keywords: Wireless Sensor Network, sensor node, energy consumption, high-speed railway.

1 Introduction

High-speed railway as a new mode of transportation in modern society has many distinct advantages, such as conducive to China's industrialization, beneficial for resource saving. China's High-speed railway is developing rapidly. At the same time, it is inevitably faced with many new problems. When a train's speed is up to 200 km/h or over, wind, rain, snow, earthquakes and other natural disasters, and public litter across the iron bridge, falling rocks, and other emergency foreign invasion may have a significant impact on the safe operation of the train [1]. For securing the safe running of the high-speed railway, these environmental parameters (including the strength and direction of wind, Rainfall, temperature, orbital vibration signals, and so on) need to be monitored to command trains correctly.

Unfortunately, most of China railway lines are without Railway Disaster Prevention and Safety Monitoring System (RDPSMS). Moreover, the existing RDPSMS is primarily done by using the traditional method to collecting regions' various information. For example, the anemometer is used to collect wind speed, and

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 983–991.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

rain gauge is used to collect rainfall values, etc. [2]. The lack of RDPSMS related to the railway track introduces high vulnerability to possible threats. Conventional wired sensor networks could possibly be used but the large length of existing track in China will introduce implementation problems. Comparing with these problems, the wireless sensor networks (WSN) is a very attractive and feasible solution for collecting, transmitting and processing the monitored data in this scenario.

Wireless sensor networks (WSN) contains thousands or even millions low-cost, battery-powered, small-size, and multi-functional sensor nodes, which cooperate on sensing different physical phenomenon. Tilak, et al [3] declare that sensor networks hold the promise of revolutionizing sensing in a wide range of application domains because of their reliability, accuracy, flexibility, cost-effectiveness, and ease of deployment.

Wireless sensor networks (WSN) used in RDPSMS has some advantages. It provides a good distributed intelligent monitoring network for watching the surrounding states of the railroad. The adoption of wireless sensor networks will greatly reduce the number of devices leads to enormously reduction of the wire complexity and the whole structural weight. Because wireless sensor network nodes have local signal processing function, a lot of data could be processed by the sensor nodes, which greatly reduces the amount of information required for transmission. Wireless sensor networks turns serial processing and centralized decision-making system, which is originally operated by the central processor, into a parallel distributed information processing system, greatly increase the speed of monitoring system and the reliability of decision-making [4]. The image of the RDPSMS based on wireless sensor network for the high-speed railway can be seen in Fig. 1.

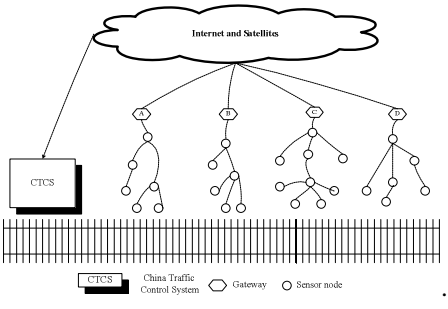


Fig. 1. An image of the RDPSMS based on wireless sensor network for the high-speed railway

This paper is organized as follows:

Section 2 discusses problem definition of the study. Section 3 presents some related work. The development of the model with different algorithms is discussed in Section 4. Results and future directions are concluded in Section 5.

2 Problem Definition

The independence and flexibility make the WSN able to be applied to many different scenarios, including which is discussed in the part I. However, almost all the nodes in

WSN are powered by batteries. The limited power supply is one of the most important issues of WSN. It is for this reason that many researchers are focusing on designing different energy efficient schemes for WSN. Among the numerous proposed methods, clustering is considered as one of the effective ways to save the energy for data transmission in a multi-hop environment [5].

In the RDPSMS based on WSN, railway maintenance workers have to verify remaining power of monitoring targets (sensors and gateways in the network) and replace their battery with a new one before the energy for each target is exhausted. Because railway is usually located at remote places, this work causes a considerable cost. Therefore, we should design the network so as to reduce the frequency of visiting the targets as much as possible. That means, we should design the energy consuming pattern of each sensor node in order to lengthen the visiting interval for replacing the battery as much as possible. In this paper, we will give an algorithm intend to maximize the minimum lifetime (the period until the energy is exhausted) of node[6], which will maximize the lifetime of the whole network.

In order to design the algorithm which has been discussed above accurately, we have to make some assumptions about the problem as follows:

- The location of each sensor is predetermined in the most suitable place to monitor the target.
- Sensor is operated by its own battery which must be replaced by a new one before its energy is exhausted.
- Each sensor collects data periodically.
- The number and locations of the gateways that can be deployed are also predetermined.
- Each gateway is equipped with wireless communication interface to the internet, sufficient memory and computing resources which stores and processes sensor data respectively.
- The data collected by the nodes is send to each gateway for a certain period, and then gateways transmit these massages to CTCS (China Traffic Control System) to control the speed of the train.

3 Related Works

Many studies for prolonging the lifetime of wireless sensor networks have been carried out till now. Minming has proposed an energy efficient transmission scheme in [7] which is based on cluster head election and power controlling (CHEPC) in heavy haul railway transportation monitoring system. In CHEPC, cluster head is periodically elected based on both energy and distance parameters. Akio et al [6] studied the resource constrained store-carry-forward wireless sensor networks to be utilized for the maintenance of railway structures, and proposed a method for designing these networks based on a metaheuristic algorithm. An energy-efficient data acquisition method for WSN applications using modern machine learning techniques has been proposed in [8]. This study was conducted using six different regression algorithms with five different datasets. A strategy called Distance

Estimation Broadcasting is designed in [5] to help a cluster member to estimate the distance between itself and its cluster head. Thus, the cluster members can use less energy to accomplish the data transmission. Moreover, a dynamic calculated Off-Duty Threshold is proposed to trigger the re-clustering operation when needed in this paper. Tang et al [9] proposed a method with consideration on the reliability of communication.

In contrast to these studies, we propose a method for prolonging the lifetime of wireless sensor network taking into consideration the characteristics of train.

4 Proposed Algorithm

Studies have shown that communication energy consumption accounted for the largest proportion of overall energy consumption of the sensor node. So, we should minimize unnecessary traffic to save energy in the wireless sensor network. There are two ways to reduce energy consumption as follows.

4.1 Advanced Direction-Aware Algorithm

Sensor node has four working conditions: transmission, reception, idle and sleep. The literature [10] points out that the seismic sensor node power consumption ratio in the table 1.

Table 1. Seismic sensor node power consumption

Working States	Column A (*t*)
Transmission	0.38-0.7W
Reception	0.36W
idle	0.34W
sleep	0.02W

From the above table, we can learn that maximizing sleep time of the node is the most effective way to save energy to extend survival time of node, thus improving the survival time of the whole network system. To solve this problem, we propose the Advanced Direction-aware Algorithm.

Assuming the direction of the train running is from A to C. We can learn that area A and area C are two adjacent regions of B from Fig. 2. If current serving area is B then C must be the target area. The nodes in region A should go to sleep; the sensor nodes in the region C should be wake up when the distance from train to the boundary of B is L.

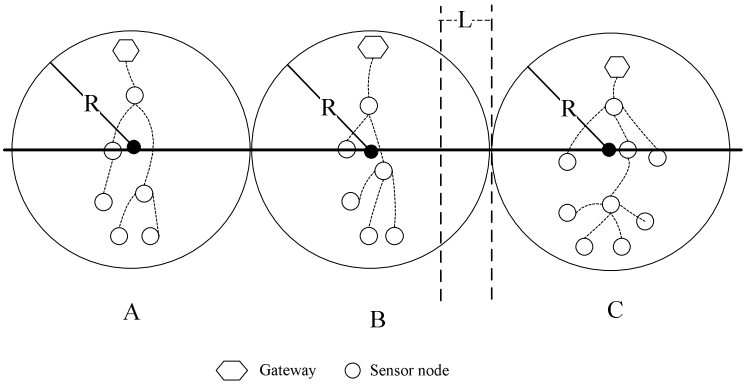


Fig. 2. An image of wireless sensor network coverage around the high-speed railway

Otherwise, if the direction of the train running is from C to A, current serving area is B then A must be the target area. The nodes in region C should go to sleep; the sensor nodes in the region B should be wake up when the distance from train to the boundary of B is L.

$$L = S + \mathrm{V}_0 T \tag{1}$$

Where, S is the maximum braking distance, V_0 is the maximum limit speed in region B, T is the lag time, which can be selected by experiences.

The proposed Advanced Direction-aware Algorithm in this paper encourages the sensor nodes in region B to be waken up, while inhibits nodes in region C to be awakened (when the direction of the train is from A to C, the train is in the region B). The image of Advanced Direction-aware Algorithm in region B can be seen in Fig. 3.

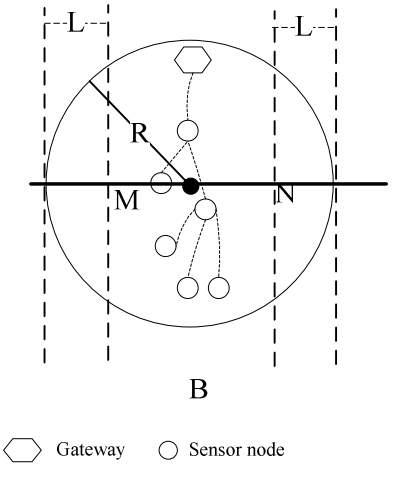


Fig. 3. An image of Advanced Direction-aware Algorithm in region B

The specific methods are as follows:

Step 1: After the train entering into MN interval, choose two time points randomly which are denoted by t_1, t_2, respectively. t_1 is not equal to t_2.

Step 2: Select N as a reference point, calculate the distance which is denoted by D_1 from the train to N when the time is t_1 and the distanc which is denoted by D_2 from the train to N when the time is t_2 separately.

Step 3: Compare D_1 and D_2. If $D_1 > D_2$, that means the direction of the train is from A to C， area C is the target region. If $D_1 < D_2$, that means the direction of the train is from C to A, area A is the target region. If $D_1 = D_2$, that means the train is in the static state, neither the nodes in region A nor in region C should be waken up.

$$WF_C = \begin{cases} 1, D_1 > D_2 \\ 0, D_1 \le D_2 \end{cases} \tag{2}$$

Where, WF_C is the wake-up factor of sensor nodes in region C. When WF_C equals 1, the sensor nodes of region C could be waken up, when WF_C equals 0, they should keep sleeping.

Step 4: When the train reached N point, wake-up sensor nodes in the region C.

In the step 2, we also could choose M as the reference point, then the equation of WF_C should be changed as follows:

$$WF_C = \begin{cases} 1, D_1 < D_2 \\ 0, D_1 \ge D_2 \end{cases} \tag{3}$$

4.2 Cluster Head Selection Algorithm Based on the Dynamic Energy Threshold

Many researchers proposed many different kinds of dynamic cluster head selection algorithm for the wireless sensor network. In these algorithms, cluster head will be picked up every round based on residual energy. The node which has the most energy will be the cluster head in the next round. These algorithms are reasonable, because the node that has more residual energy is more responsible to be a new cluster head. However, the frequent energy comparison and cluster head selection will cause every node energy loss in the network, and increase the system delay. In this part of article, we proposed a Cluster Head Selection Algorithm based on the Dynamic Energy Threshold. The algorithm give a reasonable decision for the cluster head should be re-selected in what condition.

We continually use the energy usage model described in [11] and [5]. This model is used to calculate the residual energy status, and determine the percent of live nodes

in the entire network. To transmit an l-bit data to a distance d, the radio consumes as follows.

$$E_{TX}(l,d)=\begin{cases} lE_{elec}+l_{efs}d^2, d<d_0 \\ lE_{elec}+l_{emp}d^4, d\geq d_0 \end{cases} \tag{4}$$

Where, l is the length of the message which is transmitted; E_{elec} is the energy used for sending out a single bit of data; l_{efs} is free space fading and is l_{emp} multi-path fading, these two factor should be chosen depending on the transmission distance d between the receiver and the sender; d_0 is the boundary value, it will be affected by the nodes transmission radius and $E_{TX}(l,d)$ is the energy consumption of l Byte length data transmitted d distances.

We can learn from the equations above, the energy dissipation of a node consists of two parts. The first one is consumed in running the radio electronic and the other is consumed is running the power amplifier.

Thus, the total energy consumption of the head node N before the round n should be as follows:

$$E_{con}=\sum_{i=0}^{i=n-1} E_{TX}(l_i,d_i) \tag{5}$$

And the residual energy of the node N is E_{Nr}.

$$E_{Nr}=E_0-E_{con} \tag{6}$$

Where, E_0 is the initial energy of the node N.

Suppose that the transmission data length is l_n and transmission distance is a_n in the n-round. Then the minimum energy required in the n-round is E_n.

$$E_n(l_n,d_n)=\begin{cases} l_nE_{elec}+l_{efs}d_n^2, d_n<d_0 \\ l_nE_{elec}+l_{emp}d_n^4, d_n\geq d_0 \end{cases} \tag{7}$$

$$W_{DTF}=\begin{cases} 1, E_{Nr}-E_n(l_n,d_n)>0 \\ 0, E_{Nr}-E_n(l_n,d_n)\leq 0 \end{cases} \tag{8}$$

Where, W_{DTF} is the Dynamic Threshold Factor. If W_{DTF} equals to 1, node N could be the head node in the n-round. If W_{DTF} equals to 0, node N could not be the head node any more and it should choose a new head node for the cluster by the cluster head election an power controlling (CHEPC) algorithm which is proposed in [5].

5 Conclusion

In this paper, we use the wireless sensor network in the Railway Disaster Prevention and Safety Monitoring System. Many advantages of wireless sensor network have been fully applied to monitoring environmental factors around the high-speed railway.

We proposed two strategies to extend the lifetime of the network. The first one is the Advanced Direction-aware Algorithm. This new scheme chooses the target area of wireless sensor network through judging the direction of the train, thus decreasing the energy consumption of the entire network by increasing the sleeping time of the wireless sensor nodes. And the second one is the Cluster Head Selection Algorithm based on the Dynamic Energy Threshold. This scheme re-selects the cluster head only when it is necessary. This algorithm avoids the frequent switching of cluster heads, not only reducing energy consumption but also reducing the system delay. In the future, we plan to consider more uncertain factors to improve the proposed schemes.

Acknowledgements. I would like to take this chance to express my sincere gratituede to my supervisor, Professor Yusong Yan, who has given me so much useful advices, and has tried his best to improve my paper. His willingness to give his time so generously has been very much appreciated.

I would like to offer my particular thanks to my friends and family, for their encouragement and support. Without their help, it would be much harder for me to finish my study and this paper.

References

[1] Zhang, W.: Railway Disaster Prevention and Safety Monitoring System for High Speed Railway. Railway Signalling & Communication, 80–81 (2010)
[2] Sha, Y.: Project of Disaster prevention and Safety Monitoring System in Railway for Passenger Special Line. Railway Signalling & Communication, 33–35 (2005)
[3] Tilak, S., Abu-Ghazaleh, N.B., Heinzelman, W.: A taxonomy of wireless micro-sensor network models. ACM SIGMOBILE Mobile Computing and Communications Review 6, 28–36 (2002)
[4] Shang, Y.: Research on Wireless Sensor Networks for Strain Monitoring, M.S Thesis, Nanjing University of Aeronautics and Astronautics (2008)
[5] Minming, N., Zhangdui, Z., Hao, W., Dongmei, Z.: An Energy Efficient Clustering Scheme for Mobile Ad Hoc Networks. In: 2010 IEEE 71st Vehicular Technology Conference (VTC 2010-Spring), pp. 1–5 (2010)
[6] Hada, A., Tsuchiya, R.: A metaheuristic algorithm for wireless sensor network design in railway structures. In: Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), pp. 231–236 (2009)
[7] Ni, M., Zhong, Z.: Energy Efficient Transmission in Wireless Sensor Network for Heavy Haul Railway Transportation. In: Wireless Communications Networking and Mobile Computing (WiCOM), pp. 1–4 (2010)
[8] Shafiullah, G.M., Thompson, A., Wolfs, P.J., Ali, S.: Reduction of power consumption in sensor network applications using machine learning techniques. In: IEEE Region 10 Conference, TENCON 2008, pp. 1–6 (2008)

[9] Tang, X., Xu, J.: Extending Network Lifetime for Precision-Constrained Data Aggregation in Wireless Sensor Networks. In: Proceedings of 25th IEEE International Conference on Computer Communications, INFOCOM 2006, pp. 1–12 (2006)
[10] Xu, Y.: Research on Energy-Efficiency Coverage Control Problems in Wireless Sensor Networks, M.S Thesis, Wuhan University of Technology (2010) (in Chinese)

The Integration Scheme of the Internet of Things and TD-SCDMA Network

Zhengyong Yu*

Huai'an College of Information Technology, Huai'an, Jiangsu Province, 223003, China
yonglly@sina.com

Abstract. In this paper, the realization of Mobile Internet of Things (M-IoT) based on TD-SCDMA network is given. The M-IoT service can extend the business application of TD-SCDMA system, enrich its traffic functions, and extend its application scenes. At the same time, M-IoT expands the application of IoT. The network framework of realizing M-IoT is described. The communication protocols are discussed in detail. Through OPNET modeling and simulation, the result shows that M-IoT can realize the mobile information interactivity for the fixed object and the mobile object. M-IoT based on TD-SCDMA network can not only expand the application of IoT, but also benefit the promotion of TD-SCDMA network applications.

Keywords: M-IoT, TD-SCDMA, OPNET, integration.

1 Introduction

Internet of things (IoT) [1] as a new generation of information technology, its application has infiltrated every aspect of daily life, and formed a certain scale of industry. IoT is now widely used in electric power, transportation, industrial control, retail, public services management, health, oil and other industries. It can achieve many functions, such as vehicle anti-theft, security monitoring, automatic vending, machine maintenance, public transportation management, and so on. It can improve production efficiency and has a positive of lowering production costs. In the 2G era, the insufficient bandwidth of the mobile communication network (such as GSM and CDMA) limited the diversity of business information carrying mode. So, the mobility development of IoT is also limited.

With the 3G TD-SCDMA mobile communication technology matures and the realization of commercial applications, a new era of the application for IoT is opened. TD-SCDMA network enhances the wireless access network bandwidth and provides the necessary access conditions for the mobility development of IoT. The integration of TD-SCDMA network and IoT can expand the application of IoT, such as the mobile video surveillance, telemedicine, smart transportation, smart logistics, and smart home applications, and so on. The network integration not only takes advantage of IoT development but also promotes TD-SCDMA application.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 993–1000.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Network Framework

2.1 Network Architecture

The base idea of developing M-IoT based on TD-SCDMA network is as follows. Firstly, it must full use the infrastructure and network elements in TD-SCDMA network. Then, based on TD-SCDMA mobile terminal the RFID reader [1]-[2] developed. So, the interconnection interface between IoT and TD-SCDMA network is provided. The management platform of IoT which is connected with TD-SCDMA core network is added. Finally, M-IoT based on TD-SCDMA network is realized at lower cost and is shown as Fig.1.

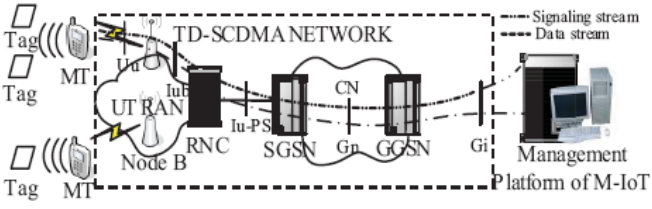


Fig. 1. Network architecture of M-IoT

From Fig.1, it can be seen that the changes of TD-SCDMA network can be summarized as the following five aspects. First, it is the network structure. In order to integrate IoT and TD-SCDMA network, the Management platform of M-IoT is added in TD-SCDMA core network. It provides some supports (such as the quality and the location of things, the information inquiry of things) for IoT in network side. Second, it is the interfaces between two networks. One is the air interface between the TD-SCDMA mobile terminal and things with the RFID tag. Another is the interface between the management platform of M-IoT and GGSN. Third, it is the communication protocol. SIP [3]-[4] is used as the signaling protocol of application layer. RTP is used as the traffic transport protocol of application layer. IP is used as transport protocol of network layer in client. At the same time, the network protocols of RNC and SGSN are extended. Fourth, it is the equipment function. TD-SCDMA mobile terminal has RFID reader function besides the communication functions. Additionally, it supports SIP, RTP and IP through extending its functions. SGSN supports SIP and RTP through extending their functions. Fifth, it is the wireless channel of TD-SCDMA network. The information of M-IoT is transmitted by using the data channels of the mobile terminal and TD-SCDMA network data links.

2.2 Network Element Function

1) Tag: It is a RFID tag. It stores information about things, such as the production data, the location and the usage. It accepts the information management and maintenance of the management platform of M-IoT.

2) Mobile terminal (MT): It is a dual-mode terminal which includes the TD-SCDMA mobile terminal function and the RFID reader function. It connects with

network and tags through air interface. It can operate the tag by accepting the order of the management platform of M-IoT. The tag information is read and processed by MT. Then, MT transmits the information to the management platform of M-IoT through TD-SCDMA network.

3) Radio network controller (RNC) and serving GPRS support node (SGSN): Their functions are extended. Two nodes, the Radio Network Controller (RNC) and Node B make the UMTS Terrestrial Radio Access Network (UTRAN). The Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN), as shown in Fig. 1, are connected to obtain mobility and session management information. The dashed lines in all figures in this paper represent signaling and the solid lines represent both data and signaling.

4) Management platform of M-IoT (MP-M-IoT): It provides the network support for M-IoT. It includes the server, the operation system and the database. The server realizes the network processing of the product information. The operation system provides the user interface of M-IoT and realizes the production information query and management. The database stores the information of things.

3 Communication Protocol

3.1 Communication Link

From Fig.1, it can be seen that the communication links include the signaling link and the data link, and each link is two-way. The signaling link is mainly used to establish and maintain the data communication between MT and MP-M-IoT. The signaling is the application layer signaling and adopts the SIP protocol. The communication link between MT and MP-M-IoT will be established by using SIP protocol when MP-M-IoT needs to send the information about things to the RFID tag of things, or MT needs to send the information which is read from the RFID tag to MP-M-IoT. In the communication process between MP-M-IoT and MT, SIP protocol is used to maintain the communication link. In order to reduce the occupancy of the TD-SCDMA network resources, the communication link will be released by using SIP protocol when there isn't the communication between MP-M-IoT and MT in a certain period of time.

The data link is the traffic information exchange link between the RFID tag and MP-M-IoT. The information of the RFID tag is encapsulated as IP format in MT. Then, the information is sent to MP-M-IoT by MT through TD-SCDMA network. MP-M-IoT will send the information of IP format to MT through TD-SCDMA network at first when it needs to send the information to the RFID tag. Then, MT converts the IP data packet to RFID communication message and sends it to the RFID tag.

3.2 Communication Protocol

The followings give the communication protocol of M-IoT.

3.2.1 Communication Protocol from RFID Tag to MP-M-IoT: As Fig.2 Shows

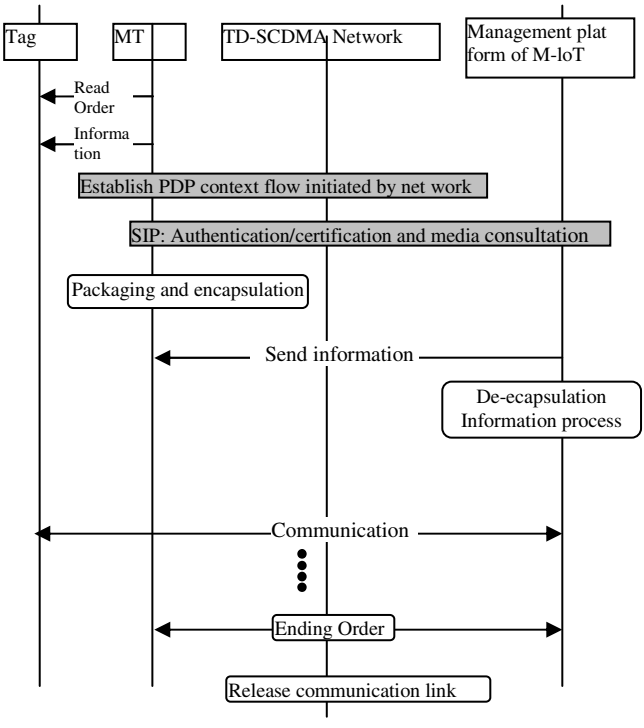


Fig. 2. Communication protocol from RFID tag to MP-M-IoT

Step 1: MT sends an order to read the RFID tag and reads the production information stored in the tag.

Step 2: The MT is triggered to initiate PDP context activation process in TD-SCDMA network when MT needs to send the information to MP-M-IoT.

Step 3: The session can be scheduled between MT and MP-M-IoT before MT sends information to MP-M-IoT, namely, it sends a request of establishing a communication link.

- MT sends a request of establishing a communication link to MP-M-IoT.
- Based on SIP protocol, the mutual authentication and certification is scheduled between MP-M-IoT and MT. At the same time, media consultation is completed.
- MP-M-IoT sends a response of establishing communication link to MT.

Step 4: The information of the RFID tag is encapsulated as IP format in MT. Then, the IP data are sent to MP- M-IoT by the communication link established by step 3.

Step 5: M-IoT server will process the received data and store them in the database. The operation systme of M-IoT can view the received message.

Step 6: Next, MT may continue to read the information of RFID tags and send it to MP-M-IoT, or MP-M-IoT sends information to RFID tags through TD-SCDMA network and MT.

Step 7: In order to reduce the occupancy of the TD-SCDMA network resources, TD-SCDMA network will send the ending order to MT and MP-M-IoT when there isn't the communication between MT and MP-M-IoT in a certain period of time. So, the wireless data channel and the packet data link are released.

3.2.2 Communication Protocol from MP-M-IoT to RFID Tag: As Fig.3 shows

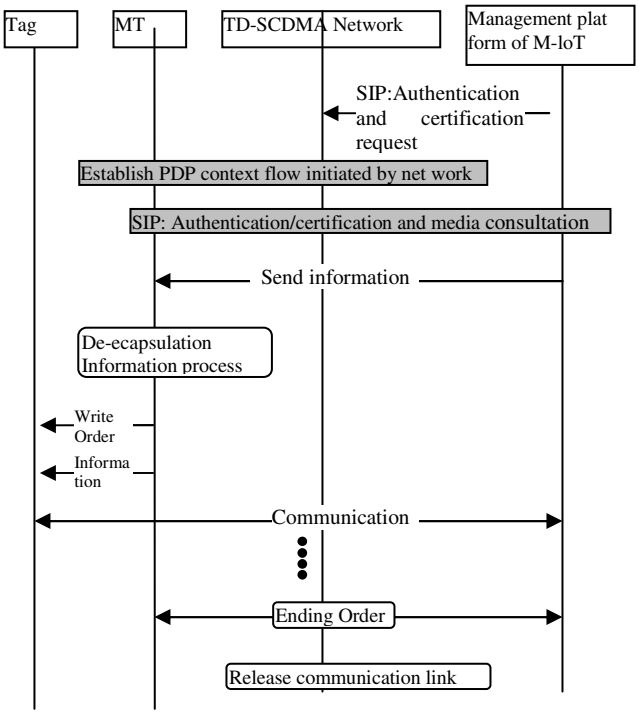


Fig. 3. Communication Protocol from MP-M-IoT to RFID tag

Step 1: In order to send informatin to MT, MP-M-IoT sends a request of authentication and certification to MT. So, PDP context flow is triggered by TD-SCDMA network-side.

Step 2: After establishing the PDP context paths between MT and GGSN, the mutual authentication and certification is scheduled between MP-M-IoT and MT based on SIP protocol. At the same time, media consultation is completed.

Step 3: MP-M-IoT retransmits information of tags to MT.

Step 4: MT de-encapsulates the receivd IP data packet and sends the written order and the information to RFID tag.

Step 5: Next, MP-M-IoT may continue to send information to RFID tags through TD-SCDMA network and MT, or MT reads the information of RFID tags and sends it to MP-M-IoT.

Step 6: So as to reduce the occupancy of the TD-SCDMA network resources, TD-SCDMA network will send the ending order to MT and MP-M-IoT when there isn't the communication between MP-M-IoT and MT in a certain period of time. Therefore, the wireless data channel and the packet data link are released.

4 Simulation and Result

4.1 Simulation Model

To verify the network performance of M-IoT, the network model is built by using OPNET network simulation software [5]. The network topology which is shown in Fig. 4 includes RFID tag (Tag_1~Tag_3), MT (MT_1~MT_3), Node B (Node B_1~ Node B_3), RNC, SGSN, GGSN, MP-M-IoT (M-IoT_server_and_database and Operation_system).

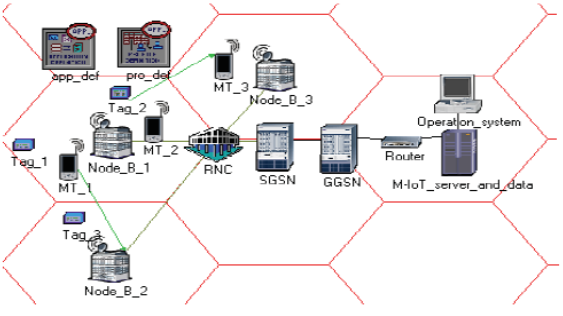


Fig. 4. Network model of M-IoT

4.2 Performance Analysis

Some assumptions are given as follows. The data communication of M-IoT is the only type of services in above network simulation. The data communication is generated between RFID tags and MP-M-IoT. Mobile terminal MT_1 and the RFID tag Tag_2 move along the path. The data communication is generated randomly. Finally, the simulation time is one hour. In Fig.5, Fig.6 and Fig.7, the network performances are got by running simulation mode.

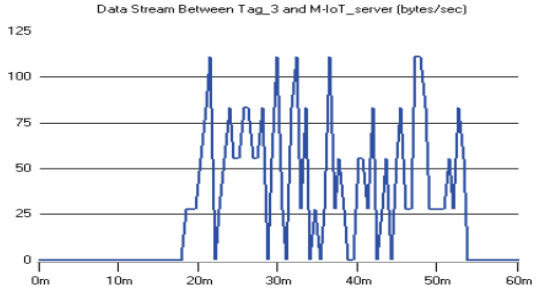


Fig. 5. Data stream between Tag_3 and the server of M-IoT

From Fig.5, it can be seen that Tag_3 and MP-M-IoT can communication with each other when the communication range of RFID reader of MT_1 covers Tag_3. So, the mobile information interactivity is realized between MP M-IoT and the RFID tags of the fixed objects.

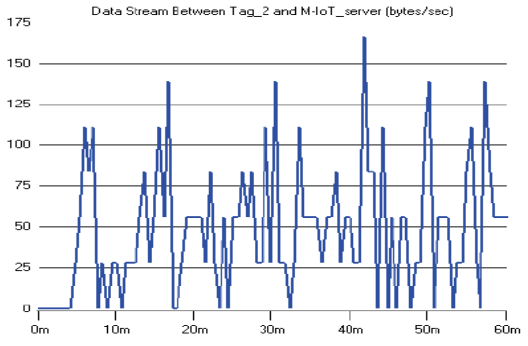


Fig. 6. Data stream between Tag_2 and the server of M-IoT

From Fig.6, it can be seen that Tag_2 mobile range is in the region the RFID reader network which is built by MT_2 and MT_3. The uninterrupted data communication is realized between Tag_2 and MP- M-IoT. So, the mobile information interactivity is realized between MP-M-IoT and the RFID tags of the mobile objects.

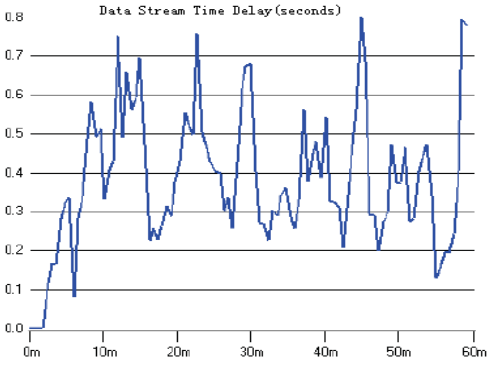


Fig. 7. Transmission time delay of M-IoT data stream

In Fig.7, the data stream time delay between the M-IoT server and the RFID tags is shown. The time delay includes the time of RFID reading and writing, the transmission time of TD-SCDMA wireless link, the transmission time of TD-SCDMA core network, the transmission time between GGSN and the M-IoT server, the data processing time of function entities. From Fig.7, it can be seen that maximum transmission delay is less than 800 milliseconds, and the time delay can meet the requirement of data communication.

5 Conclusions

In this paper, the framework of M-IoT based on TD-SCDMA is designed. And, the functions of main network elements are described. The communication protocols of M-IoT are given too. A system simulation mode is given by using OPNET software, and the information throughput and the communication time delay are analyzed.

References

[1] Yan, L., Ning, H.S., et al.: Internet of Things: from RFID to the Next-Geration Pervasion Networked System. Auerbach Publications, Taylor & Francis Group (2006)
[2] Ahson, S., Ilyas, M.: RFID handbook: Applications, Technology, Security, and Privacy. CRC Press, Boca Raton (2008)
[3] ETF Internet Protocol, RFC 3261, SIP: Session initiation protocol (June 2002)
[4] Kist, A.A., Harris, R.J.: SIP signaling delay in 3GPP. Fremantle WA. In: Sixth International Symposium on Communications Interworking of IFIP - Interworking 2002, pp. 211–222 (2002)
[5] Chen, M.: OPNET network simulation. China Tsinghua University Press, Beijing (2004) (in Chinese)
[6] International Telecommunication Union (ITU), ITUInternet Reports, The Internet of Things (November 2005)
[7] Seidler, C.: RFID Opportunities for mobiletelecommunication services. ITU-T Lighthouse Technical Paper (2005)
[8] Michael, M.P., Darianian, M.: Architectural Solutions for Mobile RFID Services on Internet of Things. Technical Report NRC-TR-2008-003, Nokia Research Center (April 2008)
[9] The EPCglobal Architecture Framework, EPCglobal Final Version (July 1, 2005)
[10] Fielding, R.: Architectural Styles and the Design of Network-based Software Architectures. Doctoral Dissertation, University of California, Irvine (2000)

Study on Product Data Structure and Realization of Model Based on PLM

Xianghui Zhan*, Xiaoda Li, and Xiangyang Lui

Department of Mechanical and Electrical Engineering,
Zhuhai College of Jilin University, Zhuhai, 519041, China
zhanxh_1215_163.com

Abstract. Product Lifecycle Management (PLM) is an integrated information-driven approach, and reduces each section of the product lifecycle from waste. Teamcenter is the leader in PLM solutions widely used by worldwide corporations. Teamcenter is valuable for product lifecycle, it aims to manage the product data more effectively, maintain the consistency of the data, shorten developing time, and make the process of design more flexible. To achieve this goal in design stage within Teamcenter and NX, the approach of implementing product data structure is presented, the way of how to create geometric model is analyzed. Furthermore, the conception and significance of PLM and PDM, the representation of product data structure, product design approach is clarified and discussed. In the end, one successful practice case is given.

Keywords: product lifecycle management, product data management, teamcenter, nx, product data.

1 Introduction

With the globalization process of manufacture and the intensifying corporations' competition, reduce costs, shorten the development cycle, product data integration, and improve product innovation have become the new requirements of enterprise survival and development. As the newest wave in productivity, PLM have become the strategic management's method of enterprise.

There are currently some mainstream PLM products in the market such as Siemens Teamcenter, PTC Windchill, Dassault Enovia, and SAP PLM etc. Among these, Siemens Teamcenter is one of the cutting-edge and high-level PLM software, with its powerful data and information management system enterprises can strengthen their competitive advantages in product R&D (Research and Development) field. Teamcenter has outstanding performances in document management and BOM management, it makes it easy to express the products' structure, and through the BOM it can manage the complexity and instability of the products. NX is another product of Siemens with a powerful suite of integrated CAD, CAE and CAM solutions, which bringing high definition PLM to product development.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1001–1008.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

In recent years, the study on Teamcenter mainly focuses on the system structure and its functional modules [1-4], lacking in the practical engineering product design process, especially in how to create product's geometry data with application of Teamcenter. The software application in this paper is all based on Teamcenter and NX, this paper mainly presents the realization of product data structure in Teamcenter, and the communication between the PDM system and the geometry model data through the NX Manager function. Afterwards the paper gives a practical enterprise case.

2 Introduction of PLM/PDM and Informatization

2.1 PLM Conception

At present the connotation of PLM lacks common understanding, so it is difficult to illustrate all the ingredients of PLM with a simple definition. The representative one is CIMdata, global leaders in PLM consulting, defines PLM as [5]:

- A strategic business approach that applies a consistent set of business solutions that support the collaborative creation, management, dissemination, and use of product definition information.
- Supporting the extended enterprise (customers, design and supply partners, etc.).
- Spanning from concept to end of life of a product or plant.
- Integrating people, processes, business systems, and information.

Fig. 1 shows the model of product lifecycle put forward by Siemens PLM software, including requirements, concept engineering, product engineering, sourcing, manufacturing engineering, product test, manufacturing production, sales/distribution, maintenance/repair, and disposal/recycling. The model's core is "knowledge" of each product or phase, covering the product of each function areas from requirements to disposal. Previously, the enterprises comparatively concerned the unit technologies within product life cycle, such as CAD, PDM, CRM, etc, and this leads to form "Information Island" between each unit technologies, solutions, and enterprise internal and external. PLM could seamlessly integrate each dispersed "Information Island", providing a unified data for each stage. Michael Grieves,a PLM expert, considered: PLM eliminates waste and efficiency across all aspects of a product's life--from design to deployment--not just in its manufacture [6].

With the development and promotion of Computer Aided Design (CAD)、Engineering Data Management (EDM)、Product Data Management (PDM)、Computer Integrated Management (CIM), the technology of PLM can come into formation. PDM is one significant part of it.

2.2 PDM and Product Informatization

With the development of computer technology, the physical object can be virtualized, and it can be expressed as 'information' in the system. In this virtual environment, PDM is an information warehouse, it is much more powerful than CAD or any other documents management functions, it is the main carrier of all information, and it can create, administer all these information.

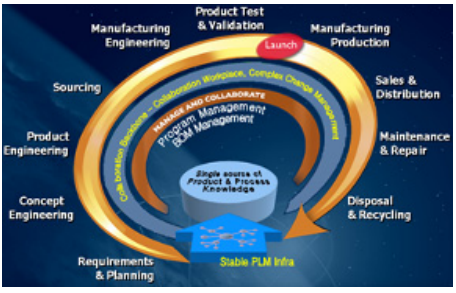


Fig. 1. The Model of product lifecycle

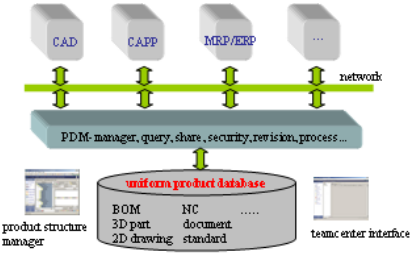


Fig. 2. PDM environments in teamcenter

Technically, PLM is one kind of technology that manages all product-related data in the whole product lifecycle; it is inextricably linked with PDM. PDM functionality is one subset of PLM, and it contains the entire contents of PDM. Product data management systems, manages product-related information like, engineering drawings, project plans, product specifications, CNC programs, analysis results, bills of material, engineering change orders, and much more such information [7].

As shown in Fig. 2, under the PDM environment in Teamcenter, product data structure is created and managed, through the user interface different users can do the operations, such as managing, inquiring, sharing, etc., to the data, like BOM, 3D part, etc., in uniform product database. In the entire product life cycle, the uniqueness of product data can be maintained.

3 Product Data Structure

3.1 Product Data Representation

Product data model is a concept model, which analyses information on the product and its relationship with other pieces of information by describing them formally and carefully [8].

As shown in Fig. 3, Product Structure Tree expresses the hierarchical relationships within product assembly, sub-assembly, and components, according to the product's formed structure it organized the related technical documentation and management documents orderly. Each tree node represents a component or a part, each node associates with the attribute information of the part (drawing number, material, weight, size, etc.) and relevant documentation (model data files, process files, etc.).

3.2 Realization of Product Structure in Teamcenter

3.2.1 The Basis of Document Management

Folder, Item/Revision, Dataset and etc. can be applied to realize the organizations of product data structure in Teamcenter. [9]

Folders are data objects that you use to organize product information. Folders can contain any other information object, including other folders. As well it can create a folder structure of the upper and lower levels to classify, organize all relevant data.

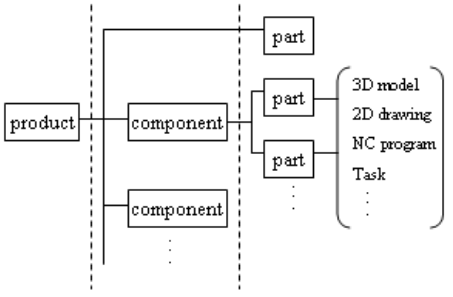


Fig. 3. Product structure tree

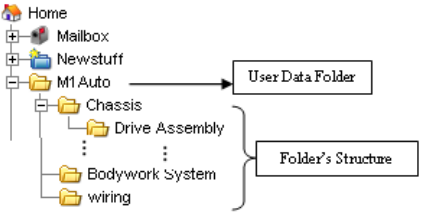


Fig. 4. Folder structure

Teamcenter has three folders Home, Mailbox, and NewStuff by default settings. As shown in Fig. 4.

Items and item revisions are the fundamental objects used to manage information in Teamcenter. Items are defined as representing a product, part, or component in the real world, such as reducer assembly, shaft etc. At least one Item revision would be created corresponding to each item. Item revision is used to manage changes (revisions) to your product information. Forms and master forms are data objects that display specific product information (properties) in a predefined template. Form associated with Item is called ItemMaster, they are created and destroyed in same time; Form associated with ItemRevision is called ItemRevision Master, they are created and destroyed in same time as well. As shown in Fig. 5.

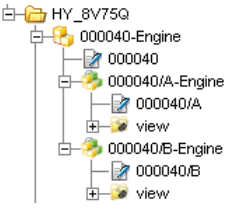


Fig. 5. Item structure

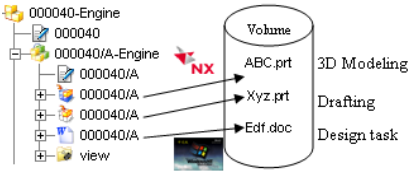


Fig. 6. Dataset managemment

Datasets are data objects that manage data files created by other software applications. Datasets describe the product objects on different aspects, such as part model, design requirements etc. Each dataset reference at least has two other objects: a dataset type that contains a list of tools that can operate on the dataset, and a tool that is used to create or modify the dataset. As shown in Fig. 6.

3.2.2 PSE

In Teamcenter Engineering, The Product Structure Editor (PSE) is the application for creating, viewing and modifying product structure and its associated occurrence data.

We can perform the functions with PSE as the followings:

- Build and edit product structure
- Browse and compare product structure
- Configuration of product structure

3.2.3 BOM

BOM is an organization relationship within the data, and it describes the product composition and all the material to be covered as data format. In different stages of product lifecycle, it will produce many different types of BOM, typically including the following three EBOM, PBOM, MBOM, therefore system requests accuracy, consistency, and other requirements to the product structure information transmitted by the BOM. Teamcenter BOM management solutions allow companies to effectively manage the BOM change history and the effectiveness of different stages to provide clear and accurate product BOM.

BOM views are used to manage product structure information related to items/assembly in Teamcenter, to save the product structure successfully, users should use View as the product version in the design stage, use MESetup in the processing stage, and use MEProcess in the manufacturing stage.

4 Model Creation

4.1 Integration of Multiple CAD Systems

For the customers located in different spots over the world with multiple CAD systems, they need to manage multi-CAD format data, which may come from their own system, or customers and suppliers. Thus, in a unified PLM environment to manage CAD assets of different series is crucial.

Parts from different CAD systems including NX, ProE, catia etc. can now be brought together into a common bill of material in Teamcenter. Teamcenter is able to automatically synchronize and manage CAD and visualized data, so that multi-site design team members can automatically browse complex assembly structure, do visualization, annotation, measurement, and analysis to the object. Multi-CAD system can make a fast real-time reflection on product design to realize the genuine collaborative engineering design project.

4.2 NX Manager

NX is the software introduced by Siemens PLM Software, which integrated CAD/CAE/CAM for three-dimensional parametric design. NX Manager is an application in Teamcenter Engineering environment, and it realizes the seamless integration between CAD and PDM data. It combines the power of NX in modeling and generating data from geometric shapes with the power of Teamcenter Engineering in storing and retrieving data in a controlled fashion.

With the communication and interaction between the two programs, users can create, access, and manage their data within a Teamcenter Engineering database; easily access to product design and structure; the structure tree in PSE corresponds to the assembly structure in NX; for the changes to the BOM table, Teamcenter and NX two solutions can keep pace with each other simultaneously.

For supporting concurrent engineering, separation of different types of data, and a significant reduction in data redundancy, the conception of master model is defined. Master model refers to the part models designed by the upper-stream personnel who is

responsible for designing, the downstream application personnel (drawing, processing and analysis personnel) only have the "read" powers not "write" powers to the models, which is "references." UGMASTER and UGPART is a type of dataset. UGMASTER is applied to save the three-dimensional master model, UGPART is applied to save the non-master model, which can be two-dimensional engineering drawings, analysis models, manufacturing models, etc. so the data derived and basic geometric shape data can be stored separately. In general, there is only one master model file, but non-master model files can be multiple. In Teamcenter, Specification and Manifestation these two data types are applied to define these relation, in essence they are a kind of assembly relationships, when the master model updates, the non-master model updates accordingly. As shown in Fig. 7.

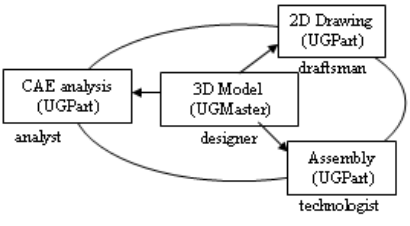


Fig. 7. Pattern of master model

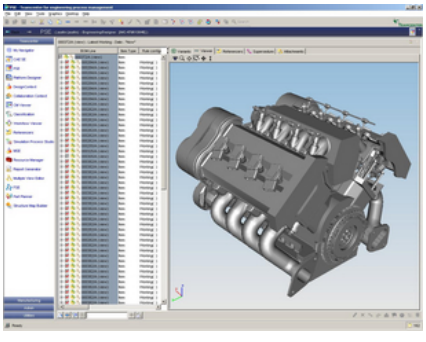


Fig. 8. Assembly visualization in PSE

4.3 Product Design Methodology

Bottom-up design methodology is to first create individual models isolatedly, then the models are combined into sub-assembly, parts of the assembly are generated finally. Bottom-up design methodology is the process to establish the assembly parts relationship, the assembly relations include Touch /Align, concentric, distance, and center and so on.

The top-down design methodology is a product design development process, according to the product's basic functions and design constraints, to plan on the overall product, determine the structure of product level, and then finalize the assembly relations and mutual constrains relations in the assembly components and parts. Top-down design method is useful to manage large assembly, effectively grasp the design intent, the product structural organization is clear and it is much better to transfer the design information in the team to achieve information sharing purposes. In Teamcenter, we can first build the product structure relations through PSE, and then with data communication between Teamcenter and NX manager the top-down method product design is achieved.

Undoubtedly, you are not tied up with one method to build up an assembly. For example, you can initially work in a top-down fashion, and then switch back and forth between bottom-up and top-down modeling.

5 Teamcenter Visualization

Teamcenter provides powerful 2D/3D visualization capabilities that can effectively transfer and express design intent, check the complete product data.

JT is the world's most widely used neutral data format, it can be created by using most of the major CAD software, it can comprehensively represent the relevant of model, and it can also be used in product visualization as well as the collaborative data sharing in all the PLM application software. Due to the independence of the CAD system, Teamcenter supports genuine multi-CAD assemblies in a neutral environment. As shown in Fig. 8.

Teamcenter is able to view all the two-dimensional formats, including PDF, BMP, JPG, CGM and DWG, etc.; JT format can be used in visualization to the variety of CAD assembly in major CAD systems.

6 Practice Cases

FuSheng Industrial Co. is the largest contract manufacturer of golf equipment in the world [10]. Its annual output is 15 million, and its global market share is 30%.

In order to achieve smooth cooperative mode of operation, in recent years, FuSheng sports equipment division one after another introduced NX product development solutions and Teamcenter product data management solutions, these two software is from the advanced Siemens PLM Software product lifecycle solution.

The reason why FuSheng chose NX is its excellent mold design capability, and the reason why FuSheng selected Teamcenter is the outstanding performance proved by the market and the excellent reputation in collaborative working technology field. The complete integration capability between the NX and Teamcenter accelerated FuSheng to make the decision.

After FuSheng used NX and Teamcenter, they electricalized all the data successfully, and shared all the data in real time with relevant departments. Siemens PLM guarantees the consistency of data update, in addition the clear project schedule reflects the time coordinating in the collaborative working thoroughly as well.

The Key to Success:

- Electricalized all the data, shared them in real time with relevant departments, and guarantees the consistency of data update.

Outcomes:

- Data electricalization overcame the problem of inconsistency of mold quality.
- Mold manufacturing time has been shortened to 7 days from 21 days dramatically.
- All in one glance project schedule reflects the time coordinating in the collaborative working.
- Unified management of data accelerates the application of knowledge on new projects.
- Data electricalization brings up more business opportunities.

7 Conclusions

The conception and significance of PLM and PDM is discussed in detail. Based on the representation of product data model, a methodology to construct product data structure is presented. NX combining with Teamcenter is a good way to create geometry model. Teamcenter visualization supports assembly from multiple CAD systems. At the end, the practical case shows leverage Teamcenter and NX as the enterprise's backbone in PLM.

Acknowledgements. It is highly appreciated that Siemens PLM software provided the NX software. This work is supported by a grant from the Teaching Quality Project (No. 10110208) and the teachers' training program 'Hundred Project' of Zhuhai College of Jilin University.

References

[1] Liu, Y., Zhang, Y., Xie, H., Qiu, J.: Research on Product Lifecycle Management System Based on Teamcenter. In: CCDC 2008, pp. 1683–1687 (2008)
[2] Zhang, Y., Bai, X., Xie, H., Liu, Y.: Research on Modular Innovation System for Lifecycle. In: CCDC 2008, pp. 1688–1691 (2008)
[3] Hou, S.: Research on Rapid Response Design System Based on Teamcenter Engineering Software. In: Intelligent Computation Technology and Automation International Conf., pp. 255–258. IEEE (2009)
[4] Ahmad, R., Yuqing, F., Hamidullah: Managing Product Data and Design flow Process. In: ISDA 2006, pp. 1–6 (2006)
[5] http://www.cimdata.com/plm/definition.html
[6] Grieves, M.: Product Lifecycle Management: Driving the Next Generation of Lean Thinking. McGraw-Hill (2005)
[7] Saak, A., Svuori, Immonen, A.: Product Lifecycle Management. Springer (2004)
[8] Brunetti, G., Golob, B.: Computer Aided Design 32(14), 877–887 (2000)
[9] Siemens PLM software, Teamcenter Engineering Product Data Management Student Guide (2007)
[10] http://www.plm.automation.siemens.com/zh_cn/about_us/success/case_study.cfm?ComponentTemplate=1481&Component=119492

The Research of DNS Safe Analysis and Defense Technology

Shao Ming-zhu*, Sun Bo, and Xie Rui-yun

Dept. of Computer Science,
Henan Mechanic and Electric Engineering College Xinxiang, China
shaomingzhu79@163.com

Abstract. DNS Service is the basic support of Internet, which security plays a vital role in the entire Internet. With the constant deepening of network application, DNS has been the hot spot goal of DDoS attack. In view of the characteristics of DNS service construction, this paper analyzes the security problem which DNS faces, especially the principle of DNS spoofing attack and DNS cache poisoning attack, and puts forward some corresponding methods and defense programs. In conclusion, to different kinds of users, they can use practical and effective defense measures according to their conditions and demand levels of information security.

Keywords*:* DDOS, NETWORK ATTACK, NETWORK SECURITY.

1 Introduction

DNS (Domain Name System, DNS) is a hierarchical and distributed database system which mapping the domain name to the IP address. It is the Internet's basic service, which security plays a vital role in the entire Internet. While there do many security risks exist in the DNS system itself especially in 2009, a variety of DDoS(Distributed Denial of Service)events happened frequently around the world In May 19th, hackers' attack led to the paralysis of a lot of Domain Name Resolution Servers, and many provinces' network interrupted, like Jiangsu, Anhui, Guangxi, and Hainan. In July 28, a new vulnerability appeared through BIND. This raises concerns of DNS service and the threat of security.

This paper states the working principle of DNS, analyses the main security problems faced by the DNS system, and raises the corresponding solutions and defense programs from the aspects of deploying DNS servers, preventing DDoS attack and preventing DNS spoofing.

2 DNS Service Working Principle

The working principle of DNS services is shown in Figure 1. Assume that the user needs to parse the domain name of www.test.com, while the target domain name is

*Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1009–1014.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

beyond the local DNS servers jurisdictions, at the same time, there is no record of the domain name in the local DNS servers' cache.

The resolution process of Domain Name System is described as follows:

1) Users appeal to the local DNS server to inquire DNS and request for resolving the IP address of www.test.com.
2) There are no corresponding records in the local DNS server, and then it turned to the root DNS server for help.
3) Root DNS server returns to the DNS server address of com domain.
4) The local DNS server appeals to the DNS server of com domain for resolution requests of domain name.
5) The DNS server of com domain returns to the DNS server address of test.com domain.
6) The local DNS server continues to make requires to the DNS server of test.com domain for domain name resolution.
7) The DNS server of test.com domain returns the IP address of www.test.com to the local DNS server.
8) The local DNS server returns the results of the domain name resolution to the Client, at the same time it also updates its cache records.

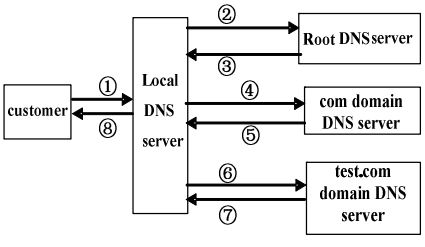


Fig. 1. DNS resolving schematic drawing

After receiving the reply packet, the client program will inquire whether the serial number and port number match the inquiries made by their own, if they match, the client program will accept the results; if not, it will discard the reply packet. This is the whole process of DNS service.

3 The Security Risk Analysis of DNS

DNS is a protocol which is not encrypted during the transmission process. Most DNS use BIND which code is on-limits, so there are security risks. Specifically, its security risks are as follows:

1) Lack of the legitimate verification, the client can not verify whether the received content of the response is legitimate or not, and the server can not verify whether the client's request is legitimate or not, either.
2) DNS mainly uses UDP (User Datagram Protocol) connectionless service, and each DNS access is independent, there is no process to establish a connection through three-way handshake. Though it speeds up the data transmission, it also led to poor defense.

3) Due to the on-limits feature of DNS, most of the DNS servers do not encrypt the data and control access, so customers can freely access each DNS server.
4) DNS uses Tree Topology. Although it is easy to query and manage, the SPOF (Single Point of Failure) which increase the security threats is obvious.

The vulnerability of DNS makes it become the weakest link in the Internet infrastructure; The Internet's dependence to DNS services also makes DNS become the hot target to a network attack.

4 The Analysis of DNS Attack and Defense Programs

4.1 DNS ID Spoofing and Defense Programs

DNS ID spoofing is a common DNS attack way in LAN. DNS message structure is shown in figure 2.

ID							
QR	OPCODE	AA	TC	RD	RA	Z	RCODE
QDCOUNT							
ANCOUNT							
NSCOUNT							
ARCOUNT							

Fig. 2. DNS Message Structure

The ID (identification) which lies in the head of the DNS Datagram is used to match the response and request datagram. During the process of resolution the domain name. Using a specific identification, the client first sends the domain name query datagram to DNS server. After the queries, it will use the same ID number sending the domain name response datagram to the client. Then the client will compare the received ID which comes from DNS response datagram with the query datagram ID which sent by itself. If they match, it means the received data is what they wait for, if not, they will be discarded. If camouflaging DNS server send a response datagram to the client ahead of the time, you can bring the client to the specified site so that to achieve DNS spoofing.

The Specific Process of DNS ID Spoofing is described as follows:

1) The cheaters proceed ARP spoofing to both the client host and the gateway at the same time in the LAN (and probably spoofing gateway and DNS server, or spoofing DNS server and client host.) On behalf of a client send ARP response datagram to the gateway host, changing the source MAC address as their own host's MAC address, at the same time, sending ARP response datagram to the client host in the name of the gateway, changing the source MAC address as the MAC address of their host. In this way, the data of gateway and the client must first pass the cheater's host.

2) We can monitor the DNS query message between gateway and the client and acquire the ID number, then send their own fake DNS response message to the client

in advance, after receiving the message, the client user will be oriented to the website which is appointed by the cheater, so the user's communication security is threatened.

3) The late response message sent by DNS server is discarded by the client and DNS ID spoofing becomes successful.

In view of this DNS attack, the following defense programs can be considered:

1) Prevent ARP spoofing. Because this DNS spoofing base on the ARP spoofing, if we can prevent ARP spoofing, we will prevent DNS spoofing. For example, we can bind the gateway router's IP address with MAC address in order to avoid deception.

2) In this DNS spoofing, the client will receive two DNS reply packets, one is legal, and the other is spoofing. In order to return to the client as soon as possible, comparing with the legitimate reply packet, the spoofing packet's message structure is simpler. It only has a response domain and without authorized domain and additional domains. According to certain rules or algorithms, we can monitor DNS response packet and distinguish the two packets, and then preventing DNS spoofing attacks.

3) For high-security-level individual services, we can access IP address directly and bypass the DNS service. Thus, all the attacks against the DNS can be avoided. In addition, we should make the secure configuration towards DNS server, install new software, restrict the range of IP address which is responded by DNS server and close the server's recursive function, etc.

4.2 Cache Poisoning and Defense Programs

In DNS service, in order to improve parsing efficiency, DNS will put the received mapping information related to the domain name and IP address in the high speed cache (Cache), for the late same request, DNS uses the information in the cache directly, if the information in the cache has been modified, when a request coming, DNS will return the wrong response.

In the Cache poisoning, the attacker will guess the message sequence number in DNS resolution process to forge the response packet of DNS authoritative server, the deceived server will receive the fake DNS response packet, and write records to their own Cache; therefore, the false domain name resolution information will be stored in DNS servers. In the cache's Period of validity, if this DNS server's client visits this domain name, it will be false pointed to the wrong analytical results, and even leads users' link request to the attacker's presupposed website.

Because DNS servers will query the domain name towards the upper DNS server after the expiration of cache records, so in the cache poisoning attack, the attacker will often first do DDoS attack to the authoritative DNS server, the server can not provide normal domain name resolution service, thus it can gain time for successfully attack the target serve.

In order to prevent the Cache poisoning attacks, we can take the following precautions:

1) Randomizing the UDP port. No longer use a fixed 53-port, randomly selected port in the range of UDP port, this can expand the combination space of ID number and port number more than 60000 times[1] ,significantly reduce the hit rate of cache poisoning.

2) Refreshing DNS timely. Rebuild the DNS cache, or according to the server's performance and network records, appropriately reduce the TTL (Time to Live) value of cache records, in order to prevent cache poisoning.

3) If necessary, restrict the updating of the dynamic DNS, and set the static DNS mapping table, so that some important sites can be protected from attacks.

4.3 Denial of Service Attack and Defense Programs

Denial of service attack is a low-tech but an effective attack method. The key position which DNS service lies in the Internet make it a hot target for cyber attacks. In 2007, 13 root name servers suffered denial of service attacks for nearly 3 hours; most of its flow is produced by other DNS servers' recursive queries. In the broken network event which happened on 2009 May 19th, it is also caused by the denial of service attacks.

Most DNS servers in the Internet are based on BIND software which has inherent security vulnerabilities[2]. An attacker could utilize the vulnerabilities in the BIND software to cause DNS server's paralysis or denial of service. We should update software timely and install the corresponding patches.

Another common denial of service attack is reflective attack. Attacker sent a large number of query request to DNS server, while the source IP address in the request packets is the attacked target's IP address. In this way, DNS server sent a large number of query results to the attacked target, and caused the network where the attacked target was in congested or can not provide the normal service.

Because the DNS query message's size is different from the response message's, with the expansion of Internet network, aiming at the particular query message, the length of the response message which sent back by the domain name server can exceed 512 bytes. When the server receiving the corresponding ENDS query message which including a space service request, the length of the response message which produced by it can increase correspondingly [3]. Therefore, the flow of DNS reflection attack has a strong amplifying effect and it is a method which is easy to implement and owns a strong effect. Baidu once suffered reflective denial of service so that it can not be accessed for up to 30 minutes.

Most importantly, the requests and connections sent by the reflective denial of service attack are real and effective. This caused the server under attack could not easily be detected when it suffered attack, in addition, the firewall and the safety inspection program operated in the attacked server are useless, because they can not prevent normal network requests and connections. At present, we can not fundamentally prevent such attacks, but can take some strategies to mitigate such attacks.

(1)Configure security policy to the border routers, take advantage of packet filtering technology to screen DNS message. Restrict the IP address of the query host, only allowing the host which belongs to this section address to query DNS servers. Restrict the flow of DNS request message, prohibit large-capacity DNS message in lowing into the network and prevent network from attack.

(2) Create backups DNS server and realize the load balancing[4]. Using the caching functions of multicast and host; allow users to access to DNS resolution services as soon as possible. It can improve the performance of DNS and reduce the load of DNS server. This can disperse the DNS server's single point fault risk to all

levels of the network, including other backup's server even host, to a certain extent, it can improve the whole network's security and reliability.

(3)Improve DNS server security configuration. Install the latest version of the DNS server software, upgrade and update software system timely. Using crosscheck, namely, after the server obtaining the host name corresponding to the IP address by the inverse query, then use the host name to query the DNS system's IP address which corresponding to it, if the results were consistent , then it is legitimate and it did response to the client. Through using ACL to restrict users' access to DNS server and restrict some areas' transmission, to prevent leakage of DNS database information.

(4)DNSSEC security architecture. IETF proposed the concept of DNS Security Extensions (Domain Name System Security, DNSSEC) in 1993, and fundamentally improved DNS's security[5]. The core idea is to sign the information of DNS through public key cryptography and provide authentication and integrity checks. In DNSSEC, the client not only checks whether the inquired records exists or not, but also deal with the signature, verification and other information. DNSSEC has strengthened its security in the DNS query and response, DNS dynamic update and DNS zone transmit, etc. Providing the integrity verification from port to port, it is a relatively sophisticated DNS security solution.

5 Conclusion

As a link of Internet, the security of Domain Name System directly determines the safety of the entire network. So the research of the attacking model and security prevention technology of DNS is particularly important. If DNS server has no protective measure completely, it will become network fraud and create the harm of money, and it will be unable to imagine.

Based on the current status of DNS service, this paper states the security risks in DNS protocol itself, aiming at the common DNS spoofing attacks, DNS cache poisoning and denial of service attack , it did a detailed analysis and defense design. It plays a positive effect in improving DNS's safety, reliability and attack-resistance. In the next generation Internet, DNS service will occupy an increasingly important position. How to realize DNS system's security in IPv6 will be our next focus.

Acknowledgements. The authors thank Sun Bo and Xie Rui-yun for the valuable discussion and recommendation.

References

[1] Kong, Z., Jiang, X.-Z.: DNS Spoofing Principle and Its Defense Scheme. Computer Engineering 36, 125–127 (2010)

[2] Yan, B.-R., Fang, B.-X., et al.: Detection and Defence of DNS Spoofing Attack. Computer Engineering 32, 130–135 (2006)

[3] Eastlake, D.: Domain name system security extensinas[S]. RFC 2535, IETF (1999-2003)

[4] Zhang, X.-M., Zhao, R.-C., et al.: Research on exploiting DNS for DoS attack and defence. Computer Engineering and Design 29, 21–23 (2008)

[5] Lioy, A., Maino, F., Marian, M.: DNS Security. In: Proc. of Terena Networking Conference (2000)

Study on the Dangerous Chemicals Transport Vehicles Monitoring System Based on RFID Technique

Wang Jianxin[1,*], Wang Tong[2], Ma Yunfei[2], and Tan Jialei[2]

[1] Department of Electronic and Information Engineering, Beijing Electronic Science and Technology Institute, Bingjing, 100070, China
wangjianxin@besti.edu.cn
[2] Beijing Municipal Institute of Labour Protection, Beijing Academy of Science and Technology, Beijing, 100054, China

Abstract. The project is based on RFID, wireless network, network, communication and database technique to develop management and technology standards of the dangerous chemicals transport vehicles of other province imported to Beijing, in order to establish the vehicles dynamic management system. The system adopts the fixed and movable RFID reader/writer methods to realize the supervision, management, trace, traffic control, emergency rescue, etc. This system can link with the GPS monitoring system of dangerous chemicals transport vehicles to improve the capacity of accident prevention and emergency response, and has a great significance to social benefit.

Keywords: wireless network, monitoring system.

1 Introduction

Dangerous chemicals transport is a high risk and accident-prone work. Dangerous chemicals transport vehicles are in hazard and can cause serious damage and loss of society, people and environment, during occurring dangerous chemicals fire, exploration, leakage accidents.

In recent years, dangerous chemicals transport accidents occur in Beijing. At about 22:50 of 30 October 2004, a vehicle of other province imported to Beijing which transports 200Kg phosphorus trichloride occurs leakage accident, and the leakage chemicals react with the rain to emit the thick toxic smoke; at about 23:00 of 31 March 2007, a vehicle of Tianjin transports 23.5 tons liquefied petroleum gas (LPG) collides the Yamenkou Bridge of Shijingshan District in Beijing, and leaks a mount of LPG. According to Beijing Municipal Environmental Protection Bureau statistics, Beijing takes place 22 dangerous chemicals highway transport accidents, involving gasoline, diesel, residual oil, LPG, calcium carbide, oil, coal tar, hydrochloric acid, n-Butyl acrylate, etc. from early 2006 to June 2007.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1015–1019.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 RFID System Analysis

2.1 RFID Advantages

There are two aspects of problems to manage and monitor other province vehicles loaded dangerous chemicals imported to Beijing. One is management mechanism problem: because the vehicles belong to the other enterprise outside of Beijing administrative area, and it is very difficult to realize the dynamic supervision; the other is monitoring technique problem: there are three methods in dangerous chemicals vehicles monitoring, involving GPS, Beidou and RFID (Radio Frequency Identification) system. But the vehicle-mounted GPS system is more expensive and not suitable to the Beijing real status; the Beidou system is less accuracy than GPS and depends on the ground networks; only the RFID system is cheaper than others, and the real-time characteristic is very high.

In USA, the Dow Chemical Company has used RFID system to monitor and control the inhaled toxic substances real-time during transport. In China, Shanghai adopts RFID system to carry out effective monitoring and management of millions of cans of dangerous chemicals.

2.2 RFID Characteristics

In order to solve the price and operation problems, this project adopt the RFID, wireless network transmission, on-line monitoring, communication and database technology to establish other province vehicles loaded dangerous chemicals imported to Beijing monitoring methods and system. In the vehicles management department of in and out of Beijing, there are fixed RFID monitoring system. And the project sets the movable RFID monitoring system in the key road and demonstration area. The other province vehicles can be dynamic monitored when passing the ports, which realizes the vehicles trace, traffic control, improves the capability of monitoring, emergency prevention.

The RFID that adopt the radio frequency to realize automatic target recognition by using non-contact method consists of two parts, the RFID reader/writer and the RFID electronic tag. In addition, the antenna is necessary to transmission and receiving data.

The other province vehicles loaded dangerous chemicals imported to Beijing will mount the RFID electronic tag which wrote the encryption driver, vehicle status, route and dangerous chemicals characteristic information. The authorized RFID reader can perform read and write operation to realize the vehicles automatic recognition and management.

3 RFID System Design

3.1 The RFID Monitoring System Structure

The RFID network is consist of many RFID nodes, which includes RFID reader/writer and information communication module to store the vehicles information, ensure the

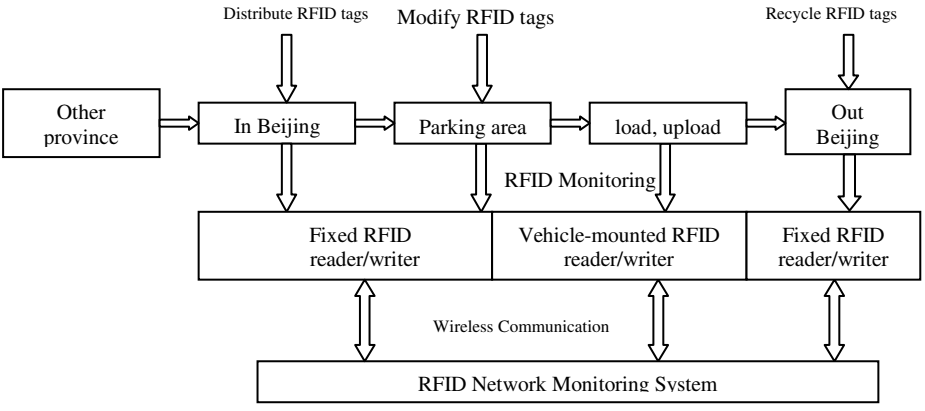


Fig. 1. The RFID Network Monitoring System Structure

information and communication synchronization of every nodes. The system software includes RFID tags distribution and logout, information management, authorization, alarm, GIS, statistical reports, system maintenance, equipment monitoring, etc. The RFID network monitoring system structure is as follow figure 1.

When the drivers get the RFID tags from the vehicle management department of in and out of Beijing, the information has transmitted to the distribution RFID tags server. If the vehicles pass the port, the RFID reader receives the RFID tags information, transmits to the center server, and then the data will be shown in GIS map or web browser. The system can print the reports via application management server if necessary.

3.2 Monitoring Position Settings

The project sets the fixed RFID reader/writer at the ports (in and out of Beijing) and special parking area to modify, read and write the RFID tags information.

The project sets the fixed RFID reader/writer in the key routes and intersections to realize the vehicles dynamic monitoring, such as prohibited routes, key roads and accidents areas.

The fixed RFID reader/writer with the wireless networks and public network communication module realizes the vehicles automatic positioning, data transmission, condition monitoring, remote control, remote reboot, remote alarm (power failure, emergency), intelligent trace, etc. If installing GPRS/3G module (4-band dual-network), the system can realize the reliable dual-mode data communication.

In order to realize the movable monitoring of the dangerous chemicals transport vehicles, some traffic management vehicles install vehicle-mounted RFID reader/writer and information communication module. The PDA is used to read and write the RFID tags information, the data transmit to the PDA server, and then transmits to the public network. This method plays an important role to realize emergency prevention when the important meeting held, important festivals, dangerous chemicals accidents emergency rescue.

3.3 RFID Tags Distribution and Recycle

Vehicle management department of in and out of Beijing distributes the RFID tags which write the above mentioned information. After the transport enterprises transact the vehicles passport of dangerous chemicals transport, the RFID tags distribute the drivers. When the vehicles leave Beijing the RFID tags recycle.

RFID tag information consist of two parts, one is tag identification which is unique ID and cannot be modified or deleted; the other can be modified, which is used to record the various information, such as name, sex, qualification, vehicle type, the total mass (tons), ratified load (tons), tank volume, etc.

3.4 Management Regulations and Technical Standards

The project establishes the relevant management regulations and technical standards in order to ensure the system performance smoothly. The standards and regulations includes the RFID tags data storage format, production method, distribution, recycle, supervision and well as department responsibilities, which can add to the current security and coordination system easily.

3.5 Demonstration Project

In the Beijing industrial layout readjustment, Fangshan District is as the petrochemical development area which depends on Yanshan Petrochemical Company. There are more than 30 chemical enterprises in Fangshan District and most enterprises are in Yanshan Petrochemical industrial area. There are 1700 vehicles comes from other province in and out this area every day, which is the largest distribution center of dangerous chemicals transport vehicles.

According to statistics, there are 73 production enterprises of dangerous chemicals in Yizhuang Development Area (Beijing Economic-Technological Development Area), and lots of vehicles come from other province are in and out this area.

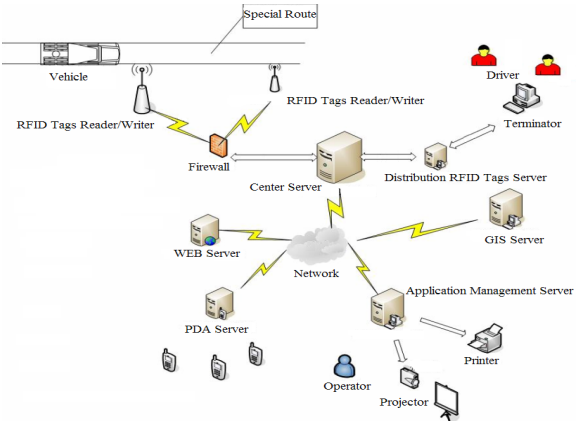


Fig. 2. The Demonstration System

The project establishes the RFID monitoring and management system in the Yanshan Petrochemical Company and Yizhuang Development Area which play a decision and supporting role to the GPS monitoring platform. The demonstration system is as follow figure 2.

4 Conclusions

The dangerous chemicals transport dynamic monitoring develops towards real-time, centralization, which integrates various technical methods into a unified platform and realizes the global decision from the whole city. The vehicles come form other province dynamic monitoring system improves the supervision efficiency, and has the advantage of low cost, wireless information process, high speed identification, tag reliable, can provide the technical support of the emergency rescue.

References

[1] Leonelli, P., Bonvicini, S., Spadoni, G.: New detailed numerical procedures for calculating risk measures in hazardous materials transportation. Journal of Loss Prevention in the Process Industries 12(6), 507–515 (1999)
[2] Fabiano, B., Currò, F., Palazzi, E., Pastorino, R.: A framework for risk assessment and decision-making strategies in dangerous good transportation. Journal of Hazardous Materials 93(1), 1–15 (2002)
[3] Pidd, M., de Silva, F.N., Eglese, R.W.: A simulation model for emergency evacuation. European Journal of Operational Research 90(3), 413–419 (1996)
[4] Rao, P.G., Raghavan, K.V.: Hazard and risk potential of chemical handling at ports. Journal of Loss Prevention in the Process Industries 9(3), 199–204 (1996)
[5] Maclean, A.D.: Chemical incidents handled by the United Kingdom public fire Services in 1980. Journal of Hazardous Materials 5(1-2), 3–40 (1981)
[6] Cumberland, R.F.: The control of hazardous chemical spills in the united kingdom. Journal of Hazardous Materials 6(3), 277–287 (1982)
[7] Bubbicoa, R., Maschiob, G., Mazzarotta, B., et al.: Risk management of road and rail transport of hazardous materials in Sicily. Journal of Loss Prevention in the Process Industries 19(1), 32–38 (2006)
[8] Davies, P.A., Lees, F.P.: The assessment of major hazards: The road transport environment for conveyance of hazardous materials in Great Britain. Journal of Hazardous Materials 32(1), 41–79 (1992)
[9] Winder, C., Zarei, A.: Incompatibilities of chemicals. Journal of Hazardous Materials 79(1-2), 19–30 (2000)

Implementation of Neural Network Backpropagation in CUDA

Jinfeng Liu* and Lei Guo

School of Automation, Northwestern Polytechnical University, Xian 710129, China
ljf@sjtu.org

Abstract. Training a multilayer Neural Network with Backpropagation algorithm is usually a very time consuming processing. In this paper, we propose an approach which uses CUDA programming model and exploits the computing power of Graphic Processing Units (GPUs) to accelerate the Backpropagation process. Experiments show that this method can achieve up to 7 times of speedup over the CPU counterpart.

Keywords: Neural Network, Backpropagation, GPU, CUDA.

1 Introduction

An Artificial Neural Network (ANN) is a mathematical model or computational model that tries to simulate the structure and/or functional aspects of biological neural networks. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks are non-linear statistical data modeling tools. They are usually used to model complex relationships between inputs and outputs or to find patterns in data.

Forward propagation is the process whereby each of the neurons calculates its output value, based on inputs provided by the output values of the neurons that feed it. Backpropagation is an iterative process that starts with the last layer and moves backwards through the layers until the first layer is reached. Assume that for each layer, we know the error in the output of the layer. If we know the error of the output, then it is not hard to calculate changes for the weights, so as to reduce that error. The problem is that we can only observe the error in the output of the very last layer.

Backpropagation gives us a way to determine the error in the output of a prior layer given the output of a current layer. The process is therefore iterative: start at the last layer and calculate the change in the weights for the last layer. Then calculate the error in the output of the prior layer.

Since a Neural Network requires a considerable number of vector and matrix operations to get results, it is very suitable to be implemented in a parallel programming

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1021–1027.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

model and run on Graphics Processing Units (GPUs). Our goal is to utilize and unleash the power of GPUs to boost the performance of a Neural Network.

2 A Brief Introduction to CUDA

The extraordinary GPU computing power is very attractive to general-purpose system development, which is referred to as general-purpose computing on GPUs (GPGPU). The first generation of GPGPU requires that non-graphics application must follow the general flow of the graphics pipeline (consisting of vertex, geometry and pixel processors), and memory access also has many restrictions, so it is very difficult to program.

The Compute Unified Device Architecture (CUDA) [1] from NVIDIA is a new technique for GPGPU. CUDA provides useful features for implementing generic data-parallel computing:

(1) Code is ANSI C extended with some keywords and data structures that eliminates the need of mapping the application to graphics API;
(2) Code can read from arbitrary addresses in memory;
(3) Full support for integer and bitwise operations.

In CUDA, the GPU is regarded as a coprocessor capable of executing a great number of threads in parallel. A single source program consists of the *host code* to be executed on the CPU and the *kernel code* to be executed on the GPU. The kernel code is usually computational intensive and data-parallel and is executed as many different *threads*. In CUDA, threads are organized into *thread blocks*, a thread block can have at most 512 threads, and threads belonging to the same thread block can share data through the shared memory and can perform barrier synchronization.

The hardware implementation of the device is a set of Streaming Multiprocessors (SMs). Each SM has a Single Instruction Multiple Data (SIMD) architecture; typically it has 8 processors that execute the same instruction on different data (threads). The blocks being processed in a multiprocessor are divided into SIMD groups of threads called *warps*.

The GPU's memories have different types in CUDA; they are specialized and have different access times and throughput limitations.

Global memory is a large, long-latency memory that exists physically as off-chip dynamic RAM (DRAM); it can be accessed by all threads in all blocks. *Shared memory* is a low latency, limited-capacity memory which can be only accessed by threads in a block. It is useful for data that can share and reuse to eliminate redundant accesses to the global memory. The *constant memory* uses a small cache of a few kilobytes optimized for high temporal locality and accesses by large numbers of threads across multiple thread blocks. The texture memory space uses the GPU's texture caching and filtering capabilities, and is best utilized with data access patterns exhibiting 2-D locality.

The properties of the different types of memories and the Suitability of CUDA have been summarized in [2][3].

3 Implementation of Backpropagation in CUDA

3.1 The Structure of ANN Used in This Paper

Fig.1 is an illustration of the neural network[4][5][6]. The input layer (Layer #1) is the grayscale image of 29x29 pixels. There are 29x29 = 841 neurons in the input layer. Layer #2 is a convolution layer with six feature maps. Each feature map is sized to 13x13 neurons. Each neuron in each feature map is a 5x5 convolution kernel of the input layer. Layer #3 is also a convolution layer with 50 feature maps. Each feature map is 5x5, and each unit in the feature maps is a 5x5 convolution kernel. Layer #4 is a fully-connected layer with 100 units. Each of the 100 neurons in the layer is connected to all 1250 neurons in the previous layer. Layer #5 is the final, output layer. This layer is a fully-connected layer with 10 units.

Altogether, there are a total of 3215 neurons in the neural network, 134066 weights and 184974 connections. The object is to train all 134066 weights so that, for an arbitrary input at the input layer, there is exactly one neuron at the output layer whose value is +1 whereas all other nine neurons at the output layer have a value of -1.

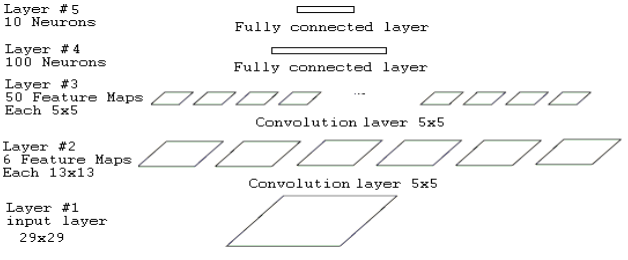


Fig. 1. The structure of ANN

3.2 General Description of Backpropagation Algorithm

Backpropagation algorithm consists of two phases: in the feed forward pass, an input vector is presented to the network and propagated forward to the output; in the backpropagation phase, the network output is compared to the desired output; network weights are then adjusted in accordance with an error-correction rule.

Steps of Backpropagation algorithm:

1) Initialization. Assign random numbers to synaptic weights, using a uniform random distribution.

2) Forward Propagation. Forward propagation is the process whereby each neuron calculates its output value based on inputs and the weight of connection indicated in equation (1)

$$x_n^i = F(y_n^i) = F\left(\sum_{j=0}^{C_{n-1}} \omega_n^{ij} \cdot x_{n-1}^j\right). \quad (1)$$

Where: x_n^i is the output of the i-th neuron in layer n, x_{n-1}^j is the output of the j-th neuron in layer n-1, ω_n^{ij} is the weight that the i-th neuron in layer n applies to the output of the j-th neuron from layer n-1. F()is the activation function.

3) Backpropagation computation:

The error due to a single pattern is calculated as follows:

$$E_n^p = \frac{1}{2} \cdot \sum (x_n^i - T_n^i)^2 \tag{2}$$

Where: E_n^p is the error due to a single pattern P at the last layer n; T_n^i is the desired output at the last layer; and x_n^i is the actual value of the output at the last layer.

Given equation (1), then taking the partial derivative yields:

$$\frac{\partial E_n^p}{\partial x_n^i} = x_n^i - T_n^i. \tag{3}$$

Equation (3) calculates numeric values for the derivative. We calculate the changes in the weights by applying the following two equations (4) and then (5):

$$\frac{\partial E_n^p}{\partial y_n^i} = G(x_n^i) \cdot \frac{\partial E_n^p}{\partial x_n^i} \tag{4}$$

Where: $G(x_n^i)$ is the derivative of the activation function.

$$\frac{\partial E_n^p}{\partial \omega_n^{ij}} = \frac{\partial E_n^p}{\partial y_n^i} \cdot \frac{\partial y_n^i}{\partial \omega_n^{ij}} = x_{n-1}^j \cdot \frac{\partial E_n^p}{\partial y_n^i} \tag{5}$$

We calculate the error for the previous layer, using the following equation:

$$\frac{\partial E_{n-1}^p}{\partial x_{n-1}^k} = \sum_i \omega_n^{ik} \cdot \frac{\partial E_n^p}{\partial y_n^i} \tag{6}$$

We take the numeric values obtained from equation (6), and use them in a repetition of equations (4), (5) and (6) for the immediately preceding layer. Meanwhile, the values from equation (5) tell us how much to change the weights in the current layer n, we update the value of each weight according to the formula:

$$(\omega_n^{ij})_{new} = (\omega_n^{ij})_{old} - \eta \cdot \frac{\partial E_n^p}{\partial \omega_n^{ij}} \tag{7}$$

Where: η is the "learning rate", typically a small number that is gradually decreased during training.

3.3 Implementation Details of Backpropagation in CUDA

3.3.1 Forward Propagation Computing

Because of the inherent structure limitation of ANN, parallelism can only be achieved within one level. For each network level, there is a CUDA function handling the

parallelism computation of neuron values of that level, the connections of the Neural Network are implicitly defined in CUDA functions.

Five arrays are defined to save the value of input layer and the weights between each layer, and transfer these arrays from host memory to device memory. The neuron values of layer 2 to layer 5 are allocated on device memory. The first kernel function calculates neuron values of layer 2. The second kernel function calculates neuron values of layer 3, and so on. So there are four kernel functions on Forward Propagation stage.

The pseudo kernel code of compute layer #2 is shown in Fig.2. Each block deal with one feature map and each thread compute one neuron value. This code is called by the host code by setting parameter of block with 6 and thread with 169.

Line 1 in Fig.2 gets the block ID and the thread ID, Line 3 defines a shared memory array s1, line 4 puts all elements of array gn1 to array s1, line 5 calculate the activation value by adding the weighted inputs together, it has two benefits to increase speed: first, loop unrolling eliminates the branch instruction and the loop counter update, second, array s1 is defined in shared memory and array gn1 is in global memory, using s1 instead of gn1 can effectively reduce access time. Line 6 applies an activation calculation. Line 7 puts result to array gn2, which contains the neuron value of layer #2.The other three kernel functions are basically similar with the first kernel function.

```
input：array gn1 contain neuron value of input layer, array gw1 contain
       weight between input layer and layer #2.
output：array gn2 is neuron value of layer #2.
```

```
1：bid=blockIdx.x;  tx=threadIdx.x;  ty=threadIdx.y;
2：wt=ty*2*29+tx*2;  result=0;
3：__shared__ float s1[29*29];
4：s1[] = gn1[]；//put all elements of array gn1 to array s1
5 ： result = s1[wt]*gw1[bid*26+1]+ s1[wt+1]*gw1[bid*26+2]+
s1[wt+2]*gw1[bid*26 +3]+…  // a total of 25 multiplication
6：  result=(1.7159*tanhf(0.66666667*result));
7：  gn2[13*13*bid+ty*13+tx]=result;
```

Fig. 2. Pseudo kernel code of calculating layer #2

3.3.2 Backpropagation Computation

1) Calculate output error of layer #5 using equations (4). Fig.3 is the pseudo kernel code which is called in the host code by setting block (10, 1) and thread (1, 1)

2) Adjust the weights in layer #4. Fig.4 is the pseudo kernel code which is called in the host code by setting block(10,1)and thread(100,1).

Line 1 in Fig.4 gets the block ID and the thread ID, line 3 and line 4 calculate the changes of weights in layer #4 according equations (5), line 5 adjust weights according equations (7), eta is the learning rate.

3) Calculate output error of #4 according equation (6). Fig. 5 is the pseudo kernel code which is called in the host code by setting block(100,1)and thread(10,1).

Line 1 in Fig.5 gets the block ID and the thread ID, line 2 defines a shared memory array st, in line 3 each thread in one block calculate one part of result and store in

array st, line 4 lets these threads synchronized. In line 5 means if the thread ID is 0, line 6 add ten parts of result together.

4) Adjust the weights in layer #3.Fig. 6 is the pseudo kernel code which is called in the host code by setting block(1251,1) and thread(100,1).

```
input：gn5 is the actual output value of layer #5,tn5 is the desired output of layer #5
output：dx5 is the output error of layer #5
1:  bid=blockIdx.x ;
2:  dx5[bid]= DSIGMOID(gn5[bid])*(gn5[bid]-tn5[bid]);
```

Fig. 3. Pseudo kernel code of calculate output error of layer #5

```
input：dx5 is the output error of layer #5,gn4 is the neuron value of layer #4,gw4 is the weights of layer #4
output：newdw4 is adjusted weights of layer #4
1: bid=blockIdx.x;   tx=threadIdx.x;
2: wbegin= bid *101;
3: dw4[wbegin]+=dx5[bid];
4: dw4[wbegin+tx+1]+=dx5[tx]*gn4[tx];
5: newgw4[wbegin+tx+1]=gw4[wbegin+tx+1] –eta* dw4[wbegin+tx+1]
```

Fig. 4. Pseudo kernel code of adjust weights in layer #4

```
input：gw4 is the weights of layer #4,gn4 is the neuron value of layer #4,dx5 is the output error of layer #5
output：dx4 is the output error of layer #4
1: bid=blockIdx.x;   tx=threadIdx.x;
2: __shared__ float st[10];
3: st[tx]=dx5[tx]*gw4[tx*101+1+ bid];
4: __syncthreads();
5: if(tx==0){
6:     dx4[bid]= st[0]+st[1]+…st[9];
7:     dx4[bid]= DSIGMOID(gn4[bid])*dx4[bid];   }
```

Fig. 5. Pseudo kernel code of calculate output error of layer #4

```
input：dx4 is the output error of layer #4,gn3 is the neuron value of layer #3.gw3 is the weights of layer #3.
output：newgw3 is adjusted weights in layer #3
1: bid=blockIdx.x ;
2: tx=threadIdx.x；
3: dw3[1251*tx+bid]=dx4[tx]*gn3[bid-1];
4: newgw3[1251*tx+bid]= gw3[1251*tx+bid] –eta* dw3[1251*tx+bid];
```

Fig. 6. Pseudo kernel code of adjust weights in layer #3

5) The rest calculations

Calculate output error of layer #3; the kernel code is similar to Fig.5. Adjust the weights in layer #2; the kernel code is similar to Fig.6. Calculate output error of layer #2; the kernel code is similar to Fig. 5. Adjust the weights in layer #1; the kernel code is similar to Fig. 6.

4 Experimental Results

We have implemented the Neural Network Backpropagation described above on NVIDIA GeForce GTX9800+.This GPU has 16 SMs, 128 1.836GHz processing cores, with 896MB onboard memory. The GPU is installed on a desktop computer equipped with a 2.8GHz Core 2 Duo CPU E7400.

One pass of Backpropagation computation spend CPU 7.457ms, and spend GPU 1.091ms. The GPU gain about 7 times speedup than CPU.

5 Conclusions

Calculation the weight of multilayer complex Neural Network using backpropagation is a very time consuming processing, it may take several hours or more than one day if uses common algorithm executed on CPU. In this paper, we mainly describe an approach of calculating backpropagation using GPU in CUDA. Experiments show that this approach can greatly reduce computation time. This paper both has practical use in backpropagation calculation and reference value for general GPU computation.

References

[1] NVIDIA Corporation. NVIDIA CUDA programming Guide Version 2.2 (April 2009), http://developer.nvidia.com/cuda

[2] Ryoo, S., Rodrigues, C.I., Baghsorkhi, S.S., Stone, S.S., Kirk, D.B., Hwu, W.-M.W.: Optimization principles and application performance evaluation of a multithreaded GPU using CUDA. In: Proceedings of ACM PPOPP 2008, pp. 73–82 (2008)

[3] Hwu, W.-M.W., Rodrigues, C., Ryoo, S., Stratton, J.: Compute Unified Device Architecture Application Suitability. Computing in Science and Engineering 11(3), 16–26 (2009)

[4] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE 86(11), 2278–2324 (1998)

[5] Simard, P.Y., Steinkraus, D., Platt, J.: Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis. In: International Conference on Document Analysis and Recognition, pp. 958–962. IEEE Computer Society (2003)

[6] O'Neill, M.: Neural Network for Recognition of Handwritten Digits (2006), http://www.codeproject.com/KB/library/NeuralNetRecognition.aspx

Research and Application of Serial Communication Modules Based on VB

Zhang Tao

Education Technology Center, Heilongjiang University, Harbin, 150080
mfdzzz1981@163.com

Abstract. With the development of the computer technology, human activities have changed a lot. The use of computer control is also a tendency. The objective of utilizing the computer language to control the device is a part of realizing automation. The paper, aims at achieving serial data communication through VB programming. In addition, we explain the knowledge of the serial data communication as well as serial data reception and transmission in detail. Besides, the brief introduction has been made in the end of the paper about the application of the serial data communications.

Keywords: VB, Serial technology, Serial Communication.

With the development of computer technology, serial data communications technology in society has been widely used, Such as: RS232/RS485/RS422 and other serial, Especially in the industrial, medical and other industries in the application of, Serial communication is a very widely used means of communication, its speed as fast although not parallel, but the small number of the wiring, easy implementation, etc., has long been very popular, Here we look at what is the first serial data communication.

1 The Serial Data Communication and Knowledge of the Serial Port

With computer technology, network technology, the rapid development of communication technology, Data communications is becoming increasingly important. Both communication between computer and peripheral equipment, including computer and information exchange between computers. This said, the computer communication is the exchange of information with the outside world. Since serial communication is a transmission line in a transmission of a message transmission line less used and can make use of existing telephone network for information transmission, therefore, particularly suitable for long-distance transmission. External device for serial storage, Such as terminals, printers, disks, etc, Exchange data using the serial mode is also very common. In real-time control and management, Classification using multiple computers distributed control system, communication between the CPU are generally serial[1].

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1029–1033.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

Serial communication is used between the peripherals and a computer data signal line(Ground and control lines may be required), Data in a data signal transmission line bit by bit, Each data occupy a fixed length of time(Figure 1). This line of communication data used less in the long-distance communication can save communication costs and, of course, the transmission speed slower than parallel transmission.

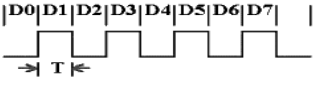


Fig. 1. Serial Communication Chart

Serial port is the CPU and serial devices encoding converter. When data is sent from the CPU through the serial port,the byte data into a serial bit. Receiving data, the serial bits are converted to byte data. In the Windows system, Serial port is part of system resources. Applications to use serial communication must be made before using the operating system resources to application requirements, Open the serial port, Communications resources must be released after, Close the serial port[2].

2 The Serial Communication Receive and Transmit Data

Serial communication using VB language writing process is actually very simple, just comes through the VB in the MSComm control can be achieved. Mainly through the following steps to complete[3]:

a, arranged in the form of a MSComm control, As a serial communication channel, Before the project began to design, select [Project] \ [Part], Specify MSComm32.OCX, in the toolbox in Figure 2 shows the components of successful communication reference;

Fig. 2.

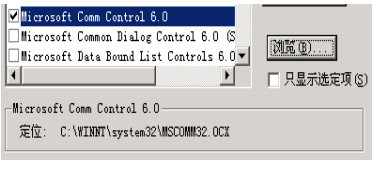


Fig. 3.

Description: If you can not in accordance with the above method in [Part] and successfully invoked the component, you can also use the program's browsing(Figure 3), to the Windows\system directory to find MSComm32.OCX, open the file.

b, The design MSComm control properties, in the form onload event, add the following code sets the control properties.

MSComm1.CommPort = 1

```
'set the port number, port number can not conflict with the port on the computer
MSComm1.Settings = "2400,N,8,1"
'Set the baud rate, 2400 for the connection speed, this must be set to the same baud rate of communications equipment
MSComm1.InputLen = 0
'Specify the length of the string read into the serial port. The default value is 0, only the time to read all the data in the input buffer.
MSComm1.InputMode = comInputModeText
'Data through the INPUT attribute to retrieve the text form
MSComm1.PortOpen = True
'before using the serial port must be open
MSComm1.RThreshold = 1      'receive data need to use
'To set or return the number of characters to receive events triggered when the receiver registers to set the number of characters, it will lead to OnComm incident receive the event.
MSComm1.SThreshold = 1      'Send the data needed to
'To set or return event caused the number of characters sent, set when the transmit register to the number of characters, it will send a trigger event OnComm incident.
```

c、Add OnComm events

```
Private Sub MSComm1_OnComm()
Select Case MSComm1.CommEvent
Case comEvReceive
'If you have received information, the following procedure is triggered
Dim ST As String
  ST = MSComm1.Input        'receive information
Case comEvSend
'If you send information to trigger the following procedure
  Dim ST As String
  ST = MSComm1.Output       'send information
End Select
End Sub
```

d、Settings button to send the data through the serial port Command2

```
Private Sub Command2_Click()
MSComm1.Output = “This is a test example!”
‘send data
End Sub
```

e、Set a text box text1 receive data through the serial port

```
text1.text= MSComm1.input
‘receive data
```

Through the above step 5 you can use VB to implement serial communication, At the end of the process, don’t forget to use MSComm1.PortOpen = false close the serial port has been opened, In the above code has been tested under win2003 + VB6.0, The code is taken from the Heilongjiang University multimedia Centralized control system parts of the central source.

3 The Serial Communication Technology

Serial communication technology based on these advantages, it has been widely used, The main application areas: Industrial Automation、Medical Laboratory、Remote control of gates 、Centralized control 、Tank Monitoring, etc. Heilongjiang University of multimedia language lab, is to use RS-232 serial communication technology developed centralized control system for multimedia(Figure 4). System uses a computer with the RS-232 serial communication controller hardware system to achieve control of the classroom use of computer software, Electric curtains, Projector, VCD, DVD, VCR, Showcase, Air conditioning and other equipment[4].

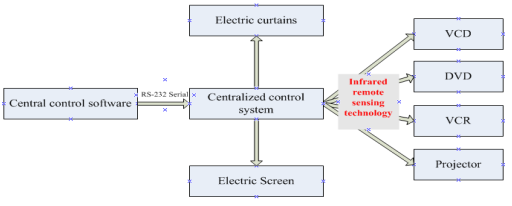


Fig. 4. Flow chart of centralized control system

Heilongjiang University in the Multimedia Language Laboratory use of computer software to control devices within the control of the classroom(Figure 5). It is mainly through the use of serial communication technology to send commands corresponding function button, In the control software can be used on the classroom computer serial port to control many devices. Many of the current control system on the market, After learning the knowledge of the above, the reader will understand the principles in the control system, serial communication with VB is not difficult to achieve[5].

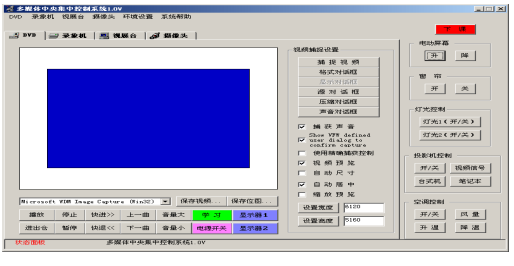


Fig. 5. In the control of multimedia language laboratory software system

4 Summary

This article explains a single serial device receiving and sending data, In fact, the process of serial applications often use multiple serial applications, Application of multi-and single serial port on the same principle applied, Only need to create multiple MSComm control, on the basis of each MSComm control program, the

computer can be achieved with multiple serial devices to communicate. VB serial communication method used to achieve similar, Readers can make the necessary changes to the practical application of, Hope that through this introduction to readers on the serial data communication technology has a simple awareness and understanding of how to use VB to communicate with serial devices.

References

[1] Wang, H., Yue, L., Yue, Z.: The application of MSComm in VC++6.0 port communication. Journal of Changchun Institute of Technology (Natural Sciences Edition) (01) (2009)
[2] Zhang, C.: The real time data collection by serial communication port based on VB. Machine Development (10) (2005)
[3] Pan, T., Chen, R.: Realization of Distributed Supervisor System for Air Analyzer Based on VB Serial Communication. Industrial Control Computer (12) (2006)
[4] Dong, H., Wang, Z., Shi, X.: The method of realizing serial communication based on MSComm. Microcomputer Information (27) (2007)
[5] Yang, J.: Based on Visual Basic RS232 serial communication with the temperature monitoring system. Electronic Test (12) (2007)

Design and Application of Export Model

Lei Xiong

Dept of Computer Science, Jiangxi Blue Sky University, 330098 Nanchang, Jiangxi, China
xionglei@163.com

Abstract. One of the common functions of pages such as report and search is to export data according to various patterns. As for implementing the export function of Word and Excel, the paper researches a variety of current and common Response output methods successively. By analyzing its advantages and disadvantages, it summarizes a reuse export model which can export various styles and patterns. The model adopts strategic design pattern to extract all the common characteristics and encapsulate them into abstraction layer in the process of export and finally make them adapt to the design principle---- only add, no modify when they are being maintained and modified.

Keywords: Strategic Design Pattern, Change Point Encapsulation, Extract Html, Export Style.

1 Introduction

It is certain that this export scheme is aimed to ASP.NET, which isn't in the range of WinForm. To implement generality, this scheme must support exporting Word and Excel simultaneously, even any other pattern. First, it needs to probe various methods to export Word and Excel, including the most common Response output method at present, so as to implement its functions entirely. Then, adopt strategic design pattern. Define general interfaces on abstraction layer on the basis of all concrete implementation to be inherited by all specific export class.

In this way, in the process engineers can make a declaration through abstract interfaces and implement functions through concrete export class so as to achieve the effect of maintainability and expandability. For example, when there is need to switch exporting Word into exporting Excel, the declaration needn't be modified but it only need change the types of class. Take another example. If a new export pattern requires to be expanded, it just needs to customize a export class to inherit general interfaces. At this point, it is consistent with the design principle---- only add, no modify.

2 Response Output Method

This method emphasizes writing the output stream of files into browser directly and that then when Response outputs, \t is used to separate data (equal to parse) and \n to change lines.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1035–1038.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2.1 JavaScript Output Method

JavaScript is the best choice in ASP.NET, because JavaScript is a pure script language which is only implemented on clients with no effect on the performance of servers.Even so, there are two flaws. First, Excel must be installed on clients, or it cannot be exported. Second, it is closely linked to whether browser permits ActiveX. In the case of default, browser won't permit loading ActiveX, which means that it is not suitable to clients.

2.2 DataGrid Output Method

This method must be based on DataGrid control to implement export. Its key codes are mainly divided into three parts.The first part is configuration, which is primarily configuration types and encoding. Here is the code.

```
HttpContext.Current.Response.AppendHeader("Content-
Disposition","attachment;filename=Excel.xls");
HttpContext.Current.Response.Charset ="UTF-8";
HttpContext.Current.Response.ContentEncoding =System.Text.Encoding.Default;
HttpContext.Current.Response.ContentType ="application/ms-excel";
```

The second part is to extract Html code, which is mainly to get HTML code of the table fromDataGrid control. Here is the code.

```
ctl.Page.EnableViewState =false;
System.IO.StringWriter  tw = new System.IO.StringWriter() ;
System.Web.UI.HtmlTextWriter hw = new System.Web.UI.HtmlTextWriter (tw);
ctl.RenderControl(hw);
```

The third part is export. Response exports files according to the HTML of tables. Here is the code.

```
HttpContext.Current.Response.Write(tw.ToString());
HttpContext.Current.Response.End();
```

Even so, there are two flaws. First, it must rely on DataGrid control. Second, apply paging in DataGrid. In this way, it cannot export all the information of select statements but that of current pages.

2.3 DataSet Output Method

The concept of this method is consistent to DataGrid, which is to output the exported information on clients. Differently, DataGrid output method emphasizes generating HTML code of table through DataGrid, while DataSet output method stresses generating it through DataTable.

Even so, there is still one flaw. DataSet is the primary DS object of DotNet1.0. But currently the latest version DotNet4.0 advocates generics. If DataGrid is applied, it appears out of fashion and too fussy, which is not beneficial to later update and maintenance.

2.4 HTML Output Method

Its concept is in agreement with that of DataGrid. What is different is that DataGrid output method emphasizes generating HTML code of table through DataGrid while

Html output method emphasizes customizing a server control which generally applies Span element or Table element. Add “runat=sever” after its definition. And then get the HTML code of table through the control directly. Owing to that, the disadvantage of this method is equal to that of DataGrid output method.

3 Generic Export Solutions

By analyzing and comparing various Response output methods, there are generally three steps in the process of export, which are to configurate Response, to extract Html code and to output Html code. The mode of extracting Html code is different in form in the four Response output methods above. For example, JavaScript output method proposes getting Html code of table directly, whose code is window.clipboardData.setData("Text",document.all('table1').outerHTML). DataGrid output method suggests getting HTML code of table from DataGrid control, whose code is ctl.RenderControl(hw). DataSet output method suggests from DataTable. Html method suggests from server control. All in all, as long as related HTML code is obtained, export will be implemented. However, all the methods mentioned above require external force to extract Html code. If we change the concept, we can customize and construct HTML code totally. In this case, first, we needn’t depend on specific control, and second we can customize the effect of export based on requirement. For example, adding extra information such as title, footer, page number, verifier in DataGrid which can only export table needs to add plenty of meaningless efforts. If the concept is changed, customizing and constructing HTML can implement the same effect and even patterns and styles can vary with the requirements so as to avoid increasing extra work load.

Whether it is to export Excel or Word or other patterns, there lie three steps to go through in the process: configurate Response, construct Html code and output Html code.

Among them, configurating Response is relevant to the type of export files. If Excel files need exporting, the export type must be appointed as Excel in configuration. If Word files are exported, the type has to be Word. However, constructing Html code is related to specific pattern which must be customized based on requirements. In nature, the process of customizing is equal to writing Html code. Differently, here Html code is written as character string. Outputting Html code can be solidified completely, because this step doesn’t distinguish export types and patterns, which belongs to generic steps.

The top layer is abstract class EKExport. Export method is used to implement export which is the only common method. This method cascades the following three methods including InitResponse, BuildHtml and OutHtml in the sequence of implementation. But other methods are private. On one hand, they can be invoked by Export method; on the other hand, they will be reloaded by the classes on the second and third layers. For example, InitResponse() method is used to take charge of initializing export object to decide whether it is to export Excel or Word directly. It is reloaded by subclass on the second layer here. BuildHtml() method takes charge of the specific pattern for export to decide the effect directly. It is reloaded by subclass on the third layer here. Different from InitResponse(), this method is virtual, which is

exported in the form of table in default. OutHtml() method can be invoked by any types and patterns, which belongs to typical common. So it is defined to be private.

Such a structure can be adapt to the design principle ---- only add, no modify. If a new type needs exporting, we can add a subclass on the second layer. If it needs to be exported in specific patterns, we can add a subclass on the third layer. In this way, the whole component possesses better maintainability and reusability.

4 Conclusion

In conclusion, the paper not only can implement the export function of Word and Excel but also adopts strategic design pattern to make export a generic export component which is easy to maintain, expand and reuse. First, configurate Response, and then construct Html code and in the end output Html code. It encapsulates two change points---- configuration and construction to the largest extent and implements the export effect of various patterns at lowest cost.

References

[1] Shi, X.-M.: Export JSP Page Data that is Based on JavaScript to Excel. Modern Computer (March 2010)
[2] Jin, C., Chen, D.: Design of Dynamic Data Exporting Module in MIS. Computer Engineering (January 2002)
[3] Liu, Z.-S.: Method and Implementation of Massive Data Export in Excel Format. Computer and Modernization (February 2010)
[4] Dong, Z.-L., Liu, Y.-X.: Research and Implementation of a Universal Data Export Tool. Fujian Computer (February 2008)
[5] Yang, D., Chang, M., Ren, X.: PDF and Its Application in the Field of Electronic Publishing. Computer Application (January 1999)

A Processing Pipeline for Virtual 3D Urban-Based Web3D Author Tools

Zhang Wei

Computer Academy of Jilin Normal University Jilin China
Wzhang@sina.cn

Abstract. Over the last years, a high demand for the Virtual 3D city and campus visited on the internet has evolved. At the same time, there are a lot of different ways to bring 3D on the internet: VRML, SecondLife, Vivaty, PDF 3D… Many 2D processes can be enhanced with 3D data, and the Web3D Consortium spearheaded the ISO certified X3D. X3D is a royalty-free open standards file format, and run-time architecture to represent and communicate 3D scenes and objects using XML on the Web, between applications and across distributed networks and web services. In this paper, we present an solution for the establishing of a Virtual 3D campus using the 3ds max 2009 and the Web3D Author tools. This approach consists of a three-step process, in which 3D urban model is made, exported from the 3ds max 2009 and then edited in the Web3D Author tool-Vivaty, transformed into various scene graph formats such as X3D.The paper concludes with the experiences gathered from implementing and using this approach and provides an outlook as to how the lessons learned can be used in application and standardization.

Keywords: X3D, Web3D, 3D urban model, LOD, Virtual 3D Urban.

1 Introduction

Despite the growing number of Web3D applications, but to create realistic virtual environment, also needs a lot of tedious modeling work, the Web3D Author tools such as X3d-Edit 32 can only makes some simple model, it's far from the actual needs. We must use the existing three-dimensional graphics development tools such as 3ds max 2009 to establish the 3D scene, the next section of this paper introduces the detail production process of the modeling, although there is nothing about the new techniques, it contains the experiences in the actual development project. Section 3 proposes how to modify the 3D scene using the Web3D Author tool-Vivaty, also includes about the way to add the interactive and set up the LOD. It's so convenient than using the X3D Edit-32. In addition, The Vivaty Studio 1.0 provides the exporting tool about the compression format of the X3D files, it can greatly reduce the size of the X3D files. Section 4 presents how to publish the X3D files into the web sites. Section 5 presents the discussion and future direction.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1039–1043.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 3D Urban Modeling

The modeling has long been one of the largest workload in the construction of virtual scene, here's a real project to use 3D modeling solutions, the specific production process is as follows:

The data collection of the city plane map. The use of existing urban planning maps or aerial photos plane, collected into city streets, the location and the area around the building and other data. 3D geospatial data and satellite images are becoming more widely used thanks to the continued improvement of geobrowser applications [5]. Google Earth, Microsoft Virtual Earth, NASA World Wind and ESRI ArcGIS Explorer are well-known geobrowsers. Additionally there are many 3D virtual earth applications providing satellite imagery display including Geosoft Dapple, Lunar Software EarthBrowser, Skyline TerraExplorer, GeoFusion GeoPlayer, Poly9 FreeEarth, etc. The Google map 11 is excellent, we can use it to download the map from their server, and export the high resolution images. But the cost must be high, if it is difficult to get these type data, we can reference about the step 3.

Importing the plane map data into three-dimensional modeling software-3ds max 2009. In this way, we can reference the plane map to fix the location and area of the building model as shown as in figure 1, and in this step, we must be attention to choose the "Lock the view with the map" in the dialog about "setup the view" in the environment of the 3ds max 2009.

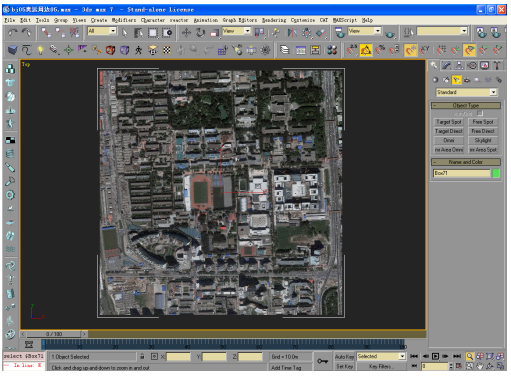


Fig. 1. Import the data of the plane map

Collect the data about the individual building. Using the digital camera to collect the corresponding data, and get the images of the each side of the building according to the building's basic shape and its length, width and height.

Establishing the model of the building

According to the data, make the model of the individual building using the software-3ds max 2009 as shown as in figure 2. Modeling process needs to pay attention to the following:

1)Uniform the use of "meter" as the scene unit.

2)Import the map data into the top view, and reference the photographs to generate models.

3)Require an accurate grasp for the proportion.

4)Make great effort for the more concise if it does not affect the accuracy of the models.

5)Using the approach of polygon modeling, Prohibited include the Re-face, rotten face, dead pixel.

6)Remove the bottom model, the polygon that its normal direction for the perspective, the user view, and the camera view can not be seen.

Establishing the virtual scene. On the basis of the map plane, immerge all the individual building into the virtual scene, and place them in the right place.

According to the maps and photos, processing the material texture mapping of the models

This process needs to grasp the overall color of the building body, that can not be without contrast and be too strong, the use of texture mapping files need to be stored in the same directory with the max scene files. It will be better to store them in the directory named map.

The model for the above, consideration should be given for the settings of the details of the models, so you can under the current model, simplify the model for the Web3D Author tool-Vivaty 1.0 in action. In order to make it possible for different levels of the natural transition between LOD models, we should ensure that the same number of LOD models of buildings on the precise alignment in space, different levels should also ensure their appearance.

After the above steps, we can get the virtual 3D urban scene, and then we need to export it saved the file as the type of wrl. Because of the 3ds max 2009 can only support the format of the VRML, the file can not be transferred into the format of the X3D. We also should be attention to the settings in the exporting, we can delete "./" in item of the "use prefix", retain the prefix folder- maps (suppose the texture map files be stored in here) as shown as in figure 3, using this approach can avoid the faults.

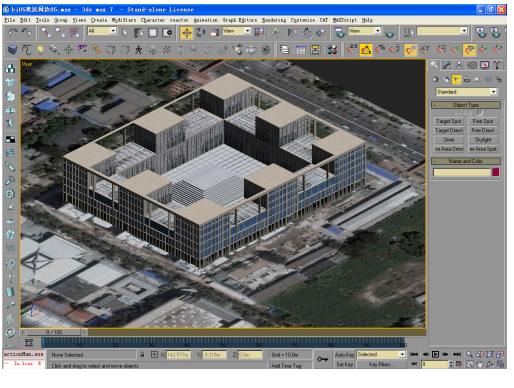


Fig. 2. Make the model of the building

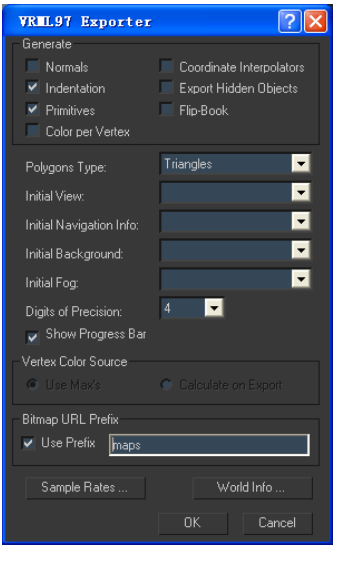


Fig. 3. Exporting Dialogue

3 Editing in the Web3D Author Tool-Vivaty Studio1.0

Although the company of the vivaty stopped the services of their website because of some financial problems, the Web3D Author tool-Vivaty Studio 1.0 is really good for use. Firstly, we should use the command “Import X3D or VRML…” under the menu of the file, to import the scene established in the 3ds max 2009 as shown as in figure 4.

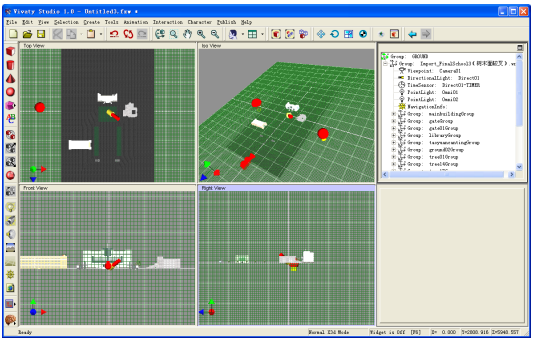


Fig. 4. Exporting the wrl file

The navigation of the models in the Vivaty Studio 1.0 on the right, displays the names of the models, we can expand it and choose the detail item, and use the page of its properties to modify its parameters such as the materials, locations and size. There are some advanced tools about “Create LOD”, “Create a Proto Instance Node” and “Create Route”, etc. In order to add the interaction, we also can use the tool of the “Create Script”. All these tools can be found in the tool bar on the left of the Vivaty Studio 1.0. Taking into account of the final works in the network transferring problems, the trees in the scene and other models can be made with the crossed patches mapping with the texture mapping, we can modify them in the 3ds max 2009 using the tool of the “Look at Constrain” with the camera. We established the virtual scene of our school using the Vivaty Studio 1.0 as shown as in figure 5.

4 Publish the Scene File

After the modified work about the scene, we can use the tool of the exporting in the Vivaty Studio 1.0, and in the exporting dialog, choose the item of the “Compressed (GZIP)”, it can reduce the X3D file size greatly. Finally, we can embed the X3D file in the web pages and it can be browsed using the Browser such as Internet Explorer 8.0. Surely it needs to install the The appropriate plug-in such as Vivaty player or Bs Contact.

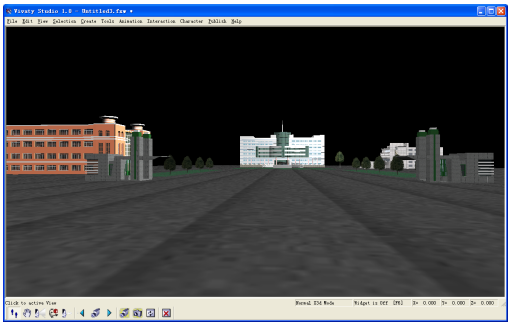


Fig. 5. The virtual scene of our school

5 Discussion and Future Directions

This paper describes the development process of the Virtual 3D Urban using the application of 3D Modeling Tool-3ds max 2009 and Web3D Author Tool-Vivaty studio 1.0, this work was based on goals set forth in the original Virtual Technical Requirements workshop. Experimental results demonstrate that the technical details of a standards-based X3D approach provides excellent infrastructure for real-time 3D simulation using high-resolution using the explorer such as the Internet Explorer. 3D information is pervasively used in Web environments. This remarkable growth of Web3D services is enabled by a rich and free infrastructure along with easy-to-use by the Web3D Author tools. It allows people who utilize the Web3D to become Web3D Author tools users without special programming training or knowledge. In the same manner, we believe that the royalty-free infrastructure for building an interoperable, real-time 3D simulation environment of information will influence the growth of scientific simulation, education, Commerce and analysis.

References

[1] Autodesk, Autodesk 3ds Max Services & Support
[2] Brutzman, D., Daly, L.: X3D: Extensible 3D Graphics for Web Authors. Morgan Kaufmann Publishers (2007)
[3] Brutzman, D.: X3D-Edit Authoring Tool (2009)
[4] Brutzman, D.: X3D Schematron Quality Assurance, QA (2009)
[5] Butler: Virtual globes: The web-wide world. Nature 439(7078), 776–778 (2006)
[6] Chon, J., Lee, Y.-W., Shibasaki, R.: Web-based interactive Virtualization of 3D video mosaics using X3D standard. Journal of Zhejiang University Science A (2006)
[7] Daly, L., Brutzman, D.: "X3D: Extensible 3D Graphics Standard," Standards in a Nutshell column. IEEE Signal Processing Magazine 24(6), 130–135 (2007)

Categorization of Medical Documents Using Hybrid Competitive Neural Network with String Vector, a Novel Approach

Seyyed Mohammad Reza Farshchi and Mahdi Yaghoobi

Department of Artificial Intelligence, Mashhad Branch,
Islamic Azad University, Mashhad, Iran
bshiveex@gmail.com

Abstract. Text categorization is one of the well studied problems in data mining and information retrieval. Even if the research on text categorization has been progressed very much, traditional approaches to text categorization require encoding documents into numerical vectors which leads to the two main problems: huge dimensionality and sparse distribution in each numerical vector. Although many various feature selection methods are developed to address the first problem, the reduced dimension remains still large. If the dimension is reduced excessively by a feature selection method, robustness of text categorization is degraded. The idea of this research as the solution to the problems is to encode medical documents into string vectors and apply it to the novel competitive neural network as a string vector. The quantitative experiment results demonstrate that this method can significantly improve the performance of medical document classification.

Keywords: Text classification, Medical documents, Neural network.

1 Introduction

In the last 10 years content-based document management tasks (collectively known as *information retrieval*—IR) have gained a prominent status in the information systems field, due to the increased availability of documents in digital form and the ensuing need to access them in flexible ways [1]. Text categorization (TC a.k.a. text classification, or topic spotting), the activity of labeling natural language texts with thematic categories from a predefined set, is one such task. TC dates back to the early 60's, but only in the early 90's did it become a major subfield of the information systems discipline, thanks to increased applicative interest and to the availability of more powerful hardware. TC is now being applied in many contexts, ranging from document indexing based on a controlled vocabulary [2], to document filtering [5], automated metadata generation [3], word sense disambiguation [4], population of hierarchical catalogues of Web resources, and in general any application requiring document organization or selective and adaptive document dispatching.

In many contexts trained professionals are employed to categorize new items. This process is very time-consuming and costly, thus limiting its applicability. Consequently

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1045–1054.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

there is an increasing interest in developing technologies for automatic text categorization [11].

A number of statistical classification and machine learning techniques has been applied to text categorization, including regression models [2], nearest neighbor classifiers [3], decision trees [4], Bayesian classifiers [7], Support Vector Machines [5], rule learning algorithms [6], relevance feedback [9], voted classification [10], and neural networks [8].

The research on text categorization has been made very much progress in context of machine learning and data mining. It requires encoding documents into numerical vectors for using one of traditional algorithms for text categorization [4].

A corpus which is a collection of documents is mapped into a list of words as the feature candidates. Among the candidates, only some are selected as the features. For each document, a numerical value is assigned to each of the selected features, depending on the importance and presence of each feature. However, encoding medical documents so causes the two main problems: huge dimensionality and sparse distribution [12].

In order to solve the two main problems, this research use the novel method that medical documents should be encoded into string vectors. A string vector refers to a finite set of strings which are words in context of a natural language. In numerical vectors representing medical documents, words are given as features, while in string vectors, words are given as feature values. Features of string vectors are defined very variously as properties of words with respect to their posting, lexical category, and statistical properties, but in this research, the highest frequent word, the second highest frequent one, and so on are defined as features of string vectors for easy implementation. By encoding medical documents into string vectors, we can avoid completely the two main problems: huge dimensionality and sparse distribution.

We proposed the competitive neural network, as the approach to medical text categorization. Before creating it, traditional neural networks, such as MLP (Multi Layers Perceptron) with back propagation receives numerical vectors as its input data. Differently from the traditional neural networks, the proposed neural network receives string vectors. It has the two layers as its architecture: the input layer and the competitive layer. It is expected for the proposed model to improve the performance of medical text categorization by solving the two main problems.

The rest of this paper is organized as follows. The principle of TC and previous works is given in Sections 2. The structure of string vector of proposed model is given in Sections 3. Section 4 describes the novel neural network model. Section 5 we will mention the simulation result and significance of this research. Conclusions are presented in Section 6. All manuscripts must be in English. Please keep a second copy of your manuscript in your office (just in case anything gets lost in the mail). When receiving the manuscript, we assume that the corresponding authors grant us the copyright to use the manuscript for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

2 Previous Works

This section is concerned with previous works relevant to this research. Even if many approaches to text categorization already proposed, we will mention the four representative and popular approaches: KNN (K Nearest Neighbor), NB (Naive Bayes), SVM (Support Vector Machine), and Neural Networks.

It requires encoding medical documents into numerical vectors for using one of them for text categorization; the two main problems are caused. String kernel was proposed in using the SVM for text categorization as the solution to the two main problems, but it failed to improve the performance. In this section, we will explore the previous works on previous approaches to text categorization and previous solution to the two main problems.

The KNN may be considered as a typical and popular approach to text categorization [1]. The KNN was initially created by Cover and Hart in 1967 as a genetic classification algorithm [2]. It was initially applied to text categorization by Massand et al in 1992 [3]. KNN was recommended by Yang in 1999 [4] and by Sebastiani in 2002 [1] as a practical approach to text categorization. Therefore, the KNN has been aimed as the base approach in other literatures as the base approach.

The Naive Bayes may be considered as another approach to text categorization. It was initially created by Kononenko in 1989, based on Bayes Rule [5]. Its application to text categorization was mentioned in the textbook by Mitchell in 1997 [7]. Assuming that the Naive Bayes is the popular approach, in 1999, Mladenic and Grobelink proposed and evaluated feature selection methods [7]. The Naive Bayes has been compared with other subsequent approaches in text categorization [10].

Recently, the SVM was recommended as the practical approach to text categorization [10]. It was initially introduced in her magazine article by Hearst in 1998 [8]. In the same year, it was applied to text categorization by Joachims [9]. It was adopted as the approach to spam mail filtering as a practical instance of text categorization in 1999 by Druker et al [10]. Furthermore, the SVM is popularly used not only for text categorization tasks but also for any other pattern classification tasks [10].

Neural Networks may be considered as an approach to text categorization, and among them, the MLP (Multiple Layers Perceptron) with back propagation is the most popular model. The neural network model was initially created in 1986 by Mcelland and Rumelhart, and it was intended to for performing tasks of pattern classification and nonlinear regressions as a supervised learning algorithm [9]. It was initially applied to text categorization in 1995 by Wiener [8]. Its performance was validated by comparing it with KNN in his master thesis on the test bed, Reuter21578. [6]. Even if the neural network classifies documents more accurately, it takes very much time for learning training documents.

The string kernel was proposed as the solution to the two main problems which is inherent in encoding documents into numerical vectors. It was initially proposed by Lodhi et al in 2002 as the kernel function of SVM [15]. The string kernel receives two raw texts as its inputs and computes their syntactical similarity between them. Since documents don't need to be encoded into numerical vectors, the two main problems are naturally avoided. However, it costed very time for computing the similarity and failed to improve the performance of text categorization.

This research has three advantages as mentioned in section. The first advantage of this research is to avoid the two main problems by encoding medical documents into alternative structured data to numerical vectors. The second advantage is that string vectors are more transparent than numerical vectors with respect to the content of its full text; it is easier to guess he content of document by seeing its string vector than by its numerical vector specially when we want to classify medical data (medical data are more complex). The third advantage as one derived from the second advantage is that it is potentially easier to trace why each document is classified. Therefore, this research proposes the novel neural network which received string vectors of medical data as its input data because of the three advantages.

3 Strategies of Encoding Medical Documents

Since medical documents are unstructured data by themselves they can not be processed directly by computers. They need to be encoded into structured data for processing them for text categorization. This section will describe the two strategies of encoding documents with the two subsections: the traditional strategy and the proposed strategy. The first subsection describes the former and points out its demerits, and the second subsection describes the latter and mentions its merits.

4 Formal Description of TC

Text categorization is the task of assigning a Boolean value to each pair $< d_j, c_i > \in D \times C$ where *D* is a domain of documents and $C = \{c_1, c_2, ..., c_{|c|}\}$ is a set of predefined categories. A value of T assigned to $< d_j, c_i >$ indicates a decision to file d_j under c_i while a value of F indicates a decision not to file d_j under c_i. More formally the task is to approximate the unknown target function $\Phi := D \times C \rightarrow \{T, F\}$ (that describes how documents ought to be classified) by means of a function $\Phi := D \times C \rightarrow \{T, F\}$, called the classifier.

5 Numerical Vector

A traditional strategy of encoding documents for tasks of text mining, such as text categorization is to represent them into numerical vectors. Since input vectors and weight vectors of traditional neural networks such as back propagation and RBF (Radial Basis Function) are given as numerical vectors, each medical document should be transformed into a numerical vector for using them for text categorization. Therefore, this subsection will describe the process of encoding medical documents into numerical vectors and what are their attributes and values.

Fig. 1 illustrates the process of extracting feature candidates for numerical vectors from medical documents. If more than two documents are given as the input, all strings

of documents are concatenated into a long string. The first step of this process is tokenization where the string is segmented into tokens by white space and punctuations. In the second step, each token is stemmed into its root form; for example, a verb in its past is transformed into its root form, and a noun in its plural form is transformed into its singular form. Words which function only grammatically with regardless of a content are called stop words [5], and they correspond to articles, conjunctions, or pronouns. In the third step, stop words are removed for processing medical documents more efficiently and reliably for text categorization.

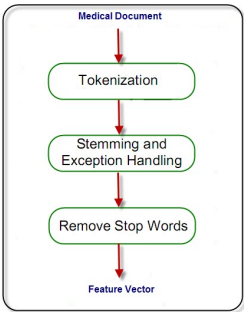


Fig. 1. Flowchart of feature extraction of medical documents

6 String Vector

An alternative strategy of encoding medical documents for text categorization is to represent them into string vectors. In this subsection, we describe this strategy and its advantage in detail. However, this strategy is applicable to only proposed neural network, while the previous one is applicable to any traditional machine learning algorithm.

A string vector is defined as a finite ordered set of words. In other words, a string vector is a vector whose elements are words, instead of numerical values. Note that a string vector is different from a bag of words, although both of them are similar as each other in their appearance. A bag of words is an infinite unordered set of words; the number of words is variable and they are independent of their positions. In string vectors, words are dependent on their positions as elements, since words correspond to their features.

Features of string vectors are defined as properties of words to the given medical document. The features are classified into the three types: linguistic features, statistical features, and positional features. Linguistic features are features defined based on linguistic knowledge about words in the given document: the first or last noun, verb, and adjective, in a paragraph, title, or full text. Statistical features are features defined based statistical properties of words in the given documents; the highest frequent word and the highest weighted word using following equation.

$$weight(w_k) = tf_i(wk)(\log_2 D - Log_2 df(w_k) + 1) \quad (1)$$

Where $tf_i(wk)$ is the frequency of medical word, w_k , D is the total number of medical documents in corpus.

Positional features are features defined based on positions of words in a paragraph or the full text: a random word in the first or last sentence or paragraph, or the full text.

We can define features of string vectors by combining some of the three types, such as the first noun in the first sentence, the highest frequent noun in the first paragraph, and so on.

A formal description of string vector is defined as a set of words which is ordered and has its fixed size. It is denoted by $[s_1, s_2, ..., s_d]$ where s_i denotes a string, and there are d elements. When representing documents into string vectors, their sizes are fixed with d, and it is called the dimension of string vectors. Since the elements are ordered in each string vector, two string vectors with their identical elements but different orders are treated as different ones. The reason is that each position of an element has its own different feature.

Table 1 illustrates differences between string vectors and numerical vectors. The first difference is that numerical values are given as elements in numerical vectors, while strings are given as elements in string vectors. The second difference is that the similarity measure between two numerical vectors is the cosine similarity or the Euclidean distance, while that between two string vectors is the semantic average similarity. The third difference between the two types of structured data is that features for encoding documents into numerical vectors are words, while those for encoding them into string vectors are statistical linguistic and posting properties of words. Therefore, a string vector is the vector where numerical values are replaced by strings in a numerical vector.

There are three advantages in representing medical documents into string vectors. The first advantage is to avoid completely the two main problems: the huge dimensionality and the sparse distribution.

Table 1. The Comparison of Numerical and string vectors

	Numerical Vector	***String Vector***
Element	Numerical value	String
Similarity measure	Inner products, Euclidean distance	Semantics similarity
Attributes	Words	Property of words

The second advantage is that string vectors are characterized as more transparent representations of medical documents than numerical vectors; it is easier to guess the content of medical documents only from their representations. The third advantage is that there is the potential possibility of tracing more easily why medical documents are classified so.

7 Structure of Proposed Neural Network

This section describes the proposed competitive neural network, in detail, with respect to its architecture, training, classification, and properties. The proposed neural network follows self organizing map (SOM) in that synaptic weights are connected directly between the input layer and the competitive layer, and the weights are updated only when each training example is misclassified. However, note that the proposed neural network is different from SOM in context of its detail process of learning and classification, since it uses string vectors as its input vectors, instead of numerical vectors. The competitive layer given as an additional layer to the input layer is different from the hidden layer of back propagation with respect to its role. The learning layer determines synaptic weights between the input and the competitive layer by referring to the tables owned by learning nodes. The learning of proposed neural network refers to the process of competition weights stored in the tables.

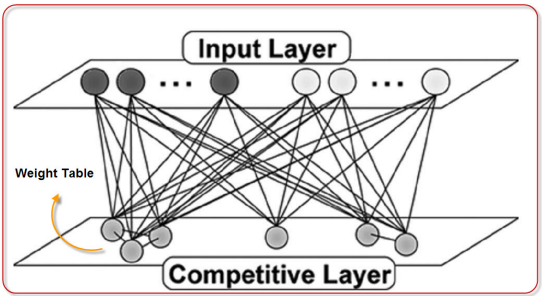


Fig. 2. The Architecture of proposed neural network

Each training example is classified by summing the initial weights and selecting the category corresponding to the maximal sum. If the training example is classified correctly, the weights are not updated. Otherwise, the weights are incremented toward the target category and those are decremented toward the classified category. The winner weights are generated as the output of this process.

In the competitive neural network, each example is classified by summing the winner optimized weights, whether it is a training or unseen example. Weights connected to itself from the input nodes as its categorical score. The weights are decided by referring the table which is owned by its corresponding learning node. The category corresponding to the output node which generate its maximum categorical score (winner category) is decided as the category of the given example. Therefore, the output of this process is one of the pre-defined categories, assuming that the competitive neural network is applied to text categorization without the decomposition.

8 Simulation Result

This section is concerned with the empirical validation of the performance of proposed competitive neural network. We use the collection of medical news, called medicalnews.com, as the test bed. This set of experiments involves the five approaches:

KNN, NB, SVM, NNBP, and our proposed neural network. The F1 measure is used for evaluating the performance of each approach to text categorization. In this section, the test bed and configurations of the approaches involved in the set of experiments are described, and the results of the set of experiments are presented and discussed.

The partition of the test bed, Medicalnews.com into the training and test set is illustrated in table 2. The test bed is given as a small collection of medical articles for entering the first evaluation, and its source is the web site, www.medicalnews.com. The collection was built by copying and pasting medical news individually as the plain text files. In the test bed, the five categories and the 1,200 medical articles are available. The collection of news articles is partitioned into the training and test set by the ratio 7:3, as shown in table 2.

The configurations of the involved approaches are illustrated in table 3. The parameters of the SVM and the KNN, he capacity and the number of nearest neighbors, are set as our and three, respectively but the NB has no parameter. The parameters of the NNBP such as the number of hidden nodes And the learning rate are arbitrary set as shown in table 3.

Medical articles are encoded into 560 dimensional numerical vectors and 60 dimensional string vectors. The results of comparing the involved approach with each other are presented in Fig. 3. Among the five method, the black bar indicates the micro-averaged F1 measure of each method. The gray bar indicates the macro-averaged F1 measure of each method, respectively. Our proposed approach shows its best performance to the NNBP, but the performance of our proposed approach is comparable to that of NNBP.

Table 2. Training and Test Set in Medicalnews.com

Category Name	*Training Set*	*Test Set*	*Total*
Digestive	400	175	575
Odontology	360	155	515
Cardiography	170	75	245
Cardiac	430	185	615
Physiology	250	110	360
Total	1610	700	2310

Table 3. Parameter Settings

Algorithms	*Parameter Settings*
SVM	Capacity = 5.0
KNN	# Nearest Value= 6
Naïve Bayes	N/A
NN With Back Propagation (BP)	# Hidden Layera=15 Learning Rate=0.2 #Training Epoche=500
Proposed Method	Learning Rate=0.2 #Training Epoche=150

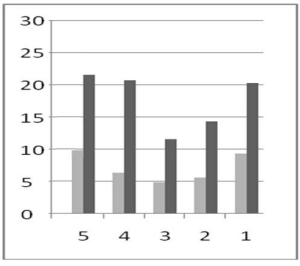


Fig. 3. The micro and macro-average F1 result for 1-SVM, 2-NB, 3-KNN, 4-proposed method, 5-NNBP

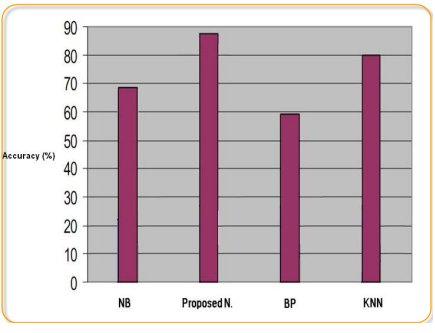


Fig. 4. The accuracy rate of three traditional approach and proposed method

Let's discuss the results from the set of experiments which were illustrated in Fig. 3. Even if the proposed neural network is not better than NNBP in the task, both are comparable to each other with respect to the performance of text categorization. Note that it requires very much time for training NNBP as the payment for the best performance. The NNBP is not practical in dynamic environments where NNBP must be trained again, very frequently. Hence, the proposed method is more recommendable than NNBP with respect to both the learning time and the performance.

Figure 3 show the accuracy of five methods on above test medical documents. The result from this picture show that the proposed neural network has more reliable than other traditional method. The accuracy rate of the proposed neural network on test bed is more than 86% but the best traditional approach have 80% accuracy rate.

9 Conclusion

The four contributions are considered as the significance of this research. For first, this research proposes the practical approach for medical document categorization, according to the results of the set of experiments. For second, it solved the two main problems, the huge dimensionality and the sparse distribution which are inherent in encoding medical documents into numerical vectors. For third, it created a new neural network, called competitive neural network, which receives string vectors differently from the previous neural networks. For last, it provides the potential easiness for tracing why each document is classified so.

References

[1] Sebastiani, F.: Machine learning in automated text categorization. ACM Computing Survey 34, 1–47 (2002)
[2] Jo, T.: Neural text categorizer for exclusive text categorization. Journal of Information Processing System 4, 10–40 (2008)
[3] Chen, J.: Feature selection for text classification with Naïve Bayes. Journal of Expert System with Application 36, 46–65 (2008)
[4] Guyon, I., Elisseef, E.: A special issue on variable and feature selection. Journal of Machine Learning Research 4, 14–34 (2008)
[5] Manning, C., Schutze, H.: Foundations of Statistical Natural Language Processing. MIT Press (1999)
[6] Mouradi, M., Bouyakhf: Integrating WordNet knowledge to supplement training data in semi-supervised agglomerative hierarchical clustering for text categorization. International Journal of Intelligent Systems 16, 929–947 (2001)
[7] Yang, Y.: An evaluation of statistical approaches to text categorization. Information Retrieval 1, 142–146 (1999)
[8] Chen, C., Sun: A self-adaptive personal view agent. Journal of Intelligent Information Systems 18, 173–194 (2002)
[9] Mitchell, T.M.: Machine Learning. McGraw-Hill (1997)
[10] Soda, P., Vullo: Text categorization for multi-page documents a hybrid naive Bayes HMM approach. Journal of Intelligent Information Systems 18, 195–217 (2002)
[11] Goldberg, J.: An approach to learning in text categorization. International Journal on Artificial Intelligence Tools 5, 229–253 (1996)
[12] Cristianini, N., Taylor, J.: Support Vector Machines and other Kernel-based learning methods, vol. 43, pp. 130–145. Cambridge University Press (2000)
[13] McClelland, J., Rumelhart, D.: Parallel Distributed Processing. MIT Press (1986)
[14] Haykin, S.: Neural Networks: A Comprehensive Foundation. Macmillan College Publishing Comp., New York (1994)

Study of Computer Room Network Monitoring Management System

Dianjun Xu

Library, Bohai University, Jinzhou, China
dianjunxu@qq.com

Abstract. In this dissertation, from the college computer room reality, based on: existed resources to develop a network monitoring management system. Major research of the dissertation is how to make use of the existed conditions on the implementation of effective management of computer room problems. The major functions of the system are: how to authorized student who can use the computer, monitoring the running processes of computer, monitoring the remote computer desktop, remote control computer, send a remote message, log monitoring, and database management. System features: Integrated authentication, Extended the process to identify ways, to achieve the RLE image coding algorithms, Remote desktop automatically poll.

Keywords: Network Monitoring, Remote Control, Computer Room Management.

1 Introduction

On one hand, computer network monitoring protects the network computer hardware and software security, on the other hand is for monitor computer usage. The monitoring computer uses, network size and rapid growth in the number of users, only on a simple increase in manpower and resources, it will lead to a sharp rise in management costs and rely on a single management style also makes artificial management Xiaoshuai not rapidly improve, the above methods can no longer meet the requirements of the development room [1-3], so we have to change the computer's management, the introduction of network monitoring and management system, which formed mainly of computer software to monitor supplementing by artificial management of comprehensive management system to effectively address this problem. At present the computer room in an open style management, there are some defects, mainly reflected Zai user management and use of monitoring two aspects [4-7]: In the user management side, it's not on the machine authorization student Jin Xing Ran Zhen, Jean Cha on the plane of the Gee Faxing, the students are within the specified time, whether using the computer or not. Computer is used in the regulation, whether in class time for students to play games, it opens to do other illegal procedures or nothing to do with the class relies heavily on teachers and administrators to students on the machine and manually manage the situation, there is no specific monitoring software to supervise students using the computers case due to the large number of students, location is relatively dispersed, and it is the hidden behavior, resulting in inefficient labor management. For the current

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1055–1061.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

regulatory system, it does not make full use of teaching resources, resulting in waste of resources, as well as the computer room to improve management efficiency; we need a network monitoring system to meet the needs of [8-10]. Therefore, the network computer room monitoring system needs to include the following features: the competence of the students is on the machine authorization certification; only by authorized certification students can use the computer. Students use monitor the situation on the plane, which the students during school hours to supervise the use of computers and remote control when necessary. Send a remote message to the students. The establishment of the log system, the violation of the student record is for inspection. To facilitate the use of teachers and students, data management module timetabling teachers, students make an appointment on the machine and program inquiries, and other functions can be used remotely via the Internet.

2 The Overall Function of the System Analysis and Design

The system functions are to achieve the certification authority and network monitoring two parts. System designs two modules: data management module and network monitoring modules to achieve system functionality. Data management module is the basis for realization of the entire system, which is mainly to provide data for the network monitoring module support, set the monitoring system parameters and for database maintenance. Network monitoring module is based on the data management module that is responsible for licensing certification and network monitoring concrete implementation, network monitoring approach including process control, screen monitoring, log monitoring, sending long-range news. Data management module includes teaching management, personnel management, equipment management, database maintenance, system that set up five sub-module; network monitoring module including authorization certification, process control, screen monitoring, log monitoring, remote message 5 module. Overall system function modules as are shown in Figure 1. The link between the two modules of the database, it indicating that the relationship between modules is shown in Figure 2.

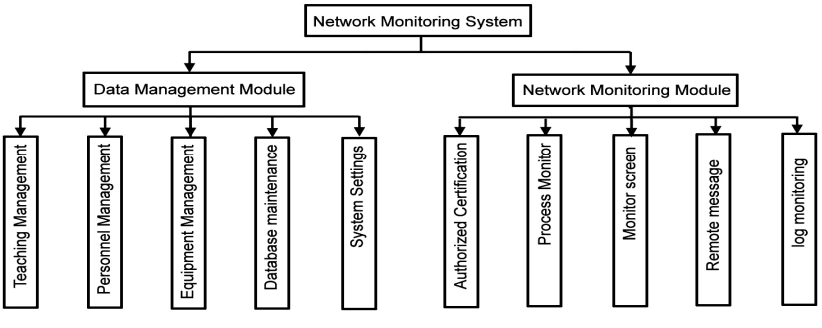


Fig. 1. System functional modules

Database stores the monitoring management-related data including the courses, classes, curriculum, personnel, computer equipment and process information. The focus of system design is for network monitoring, focusing on the process, both control and video surveillance. Authorized certification is characteristic of the system; process control system is of the rich features of this design. Structure is in the system, the system will take C / S mode to achieve, and students were monitored and the master of the administrator side terminal.

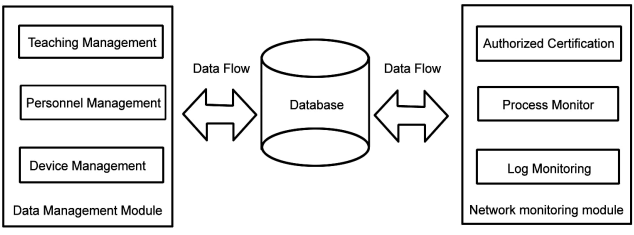


Fig. 2. Relationship between data management and network monitoring

3 The Functions of Data Management Module

Data management module is for the certification authority to provide students with identification, curriculum information by setting the work schedule and start date, as Wang Lou Jiankongmokuai Jinhangshenfen time the legitimacy of legality verification and validation O'clock Shiyong provide Shenkao; data management module is for database backup and to recovery operations to maintain the database. The function modules are shown in Figure 3.

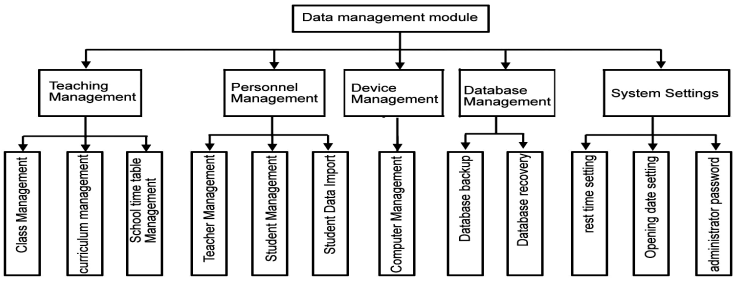


Fig. 3. The Functional Block Diagram Data Management

Data management module achieves the database-related functions through the operation to provide data support for the monitoring module. Teaching management sub-module is used to implement authorization authentication feature, which includes classes in management, curriculum management, curriculum management of the three modules. Main class management class is setting a different name to distinguish different classes. Course management courses are used to add modifies delete operation. Curriculum management shall arrange more a teacher or administrator teaching programs for students

that is based on the machine timetable for teaching content, and modify the delete operation. Curriculum is to validate the time when the Authorized is mainly on the basis of legitimacy. Personnel management sub-module is also used to implement authorization authentication feature, which includes the teacher management, student management, and science data into three modules. Teacher management information are for teachers to add modify delete operation. Student management is for the management with the basic information on students, which is mainly to add, modifies basic information deletion, password and log state changes. Student data is for the bulk import feature to import student data. In this system, student information is updated more frequently and more quantity, if this way is to manually input, then the efficiency is very low, so taking a standard format for the data source (such as the EXCEL table) to bulk import the student data, individual student information to the student management to operate, which is greatly improved the programs friendly, ergonomic design embodies. Device management sub-module includes computer management. Computer Management major is a computer code, IP address, status and other information that can be used to manage. Computer Management screen is mainly used for monitoring arrangements and the curriculum. Course scheduling is required to determine whether there are sufficient computers for students to use, only when sufficient number of computers is the arrangement to be successful. Distinguish between different computer monitor screen is a sign that IP address, IP address that comes from this. Database maintenance sub-module includes a database backup and restores two functions. Database backup recovery yes data management system is integral to the function, it is mainly in order to enhance the security system designed once the system crashes, it Mango to ensure the system of data security; re-deployment of the system is offered the chance to make losses reduce to a minimum. System Settings sub-modules are including timetables set, school time settings and administrator password that are set of three modules: the main work and rest time is set for the summer and winter schedules are different designs, it can adjust the time for each lesson. Opening time setting is the arrangement for curriculum development in accordance with the school calendar, and the system to determine by reading the current opening hours for the first few weeks. Rest time settings and set the opening time are the time to judge the legitimacy of the foundation. Password settings are on the administrator password to manage.

4 Network Monitoring Module Function

Network monitoring module includes five functions: authorization certification, process control, screen monitoring, remote message and log monitoring, which were deployed in the student side and the administrator side, depending on the data management module that provides data to support. Network monitoring module of the data stream is shown in Figure 4.

Student-side functions are as follows: they complete the work: students are on the machine authorization certification to ensure the students fax by using the computer; real-time monitoring computer to run the process and prevent students from running illegal programs; Boohoo local computer is for real-time screen image and Dui coding; receiving different types of administrator sends a message, and making the appropriate treatment; the monitoring system is agreed to write the log information table for later reference; startup, self-protection lock the computer's functions. Student-side monitoring system includes the main sub-module license authentication, the process of monitoring, screen capture, messaging, and logging. Authorized certification modules: functions are at startup, self-protection, lock computer, when the legitimacy of

authentication and certification: monitoring system at startup is a necessary function without this feature, monitoring capability would be greatly reduced. Self-protection function is to prevent users from terminating control procedures and to avoid regulation. Computer lock feature is the student who can't pass authentication, lock the computer screen, mouse and hot key part of the system to prevent unauthorized use of computer students. Authentication is the function of the student login to check the legality of student status and to prevent the legitimate students twice a login, authentication is passed, and the login successful message log record is written. Time He Faxing certification consists of two phases, the first stage is after the success of the student login to check the students; time is in the curriculum arrangements (i.e. the legal time) on the machine within; second stage is the regular inspection of the legitimate use of the current, students are in the remaining time, when the legitimate use of time is less than certain limit warning to students to remind of the preservation of information, the students are over the legitimate use of time immediately remove the students authorization, lock computer. Process control modules are as follows: the process of monitoring sub-module of the process are to achieve the function to have access to real-time process control lists and classification: the process of classification to obtain the list of functions that students are certified through the authorization process from the server to obtain classified list, which is saved locally. Real-time process monitoring function is to regularly inspect the machine running the process, the process of classification and compare the list to determine the legitimacy of the process, the process are for different types of treatment accordingly in order to achieve the process of monitoring. Capture screen image sub-module are as follows: function capture the local computer screen images and image compression, image capture function of the local computer screen through the operating system, screen capture function on the local computer and desktop image capture and save it. Time the legitimacy of certification consists of two phases, the first is the success of the student login to check the student curriculum arrangements in the original image file to capture a bitmap format, when the resolution is 1024 X 768, which saves for the 32-bit bitmap, the entire screen image file size is 1024 X 768 X 4 = 3M, if there is no compression transmission directly to the Internet that will be time-consuming, taking up a lot of bandwidth and affecting the image transmission efficiency, so the image is compression. Image coding functions are as follows: the original image file in accordance with a certain image coding algorithm, compressed into a smaller image file. Message processing modules: a message processing function is the administrator that sends the message for processing. Guan Libyan sends the message that is divided into three categories, firstly control information such as a mouse keyboard message; the second command message, such as shutdown, restart, lock computer, updating process are of the classification list of orders; thirdly, news tips, such as prompt is with after-school homework on students. Students receive on the local mouse and keyboard control the message according to message content, mouse and keyboard events in the local simulation to conduct the appropriate action. Receives shutdown, restart; lock the command message to call the local function to handle it. When prompted with the message received the message box is in the form of local to show to students. Logging modules: Module of the logging function is to monitor the system that is agreed to write the database information in the log table. Agreed information system includes: student personal information, time, log type and the operation information (such as students log on, run the illegal program, unedited operating procedures) and so on.

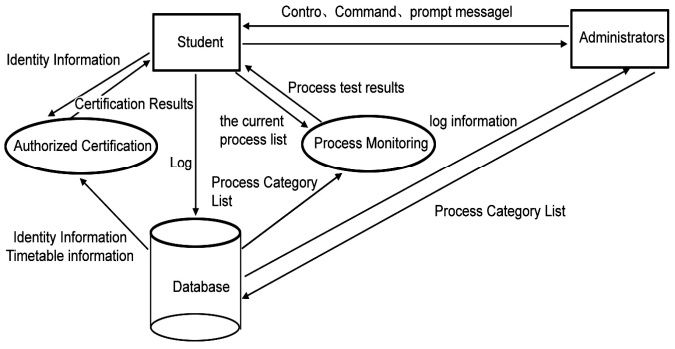


Fig. 4. Monitoring module data flow diagram

Administrator client features: Monitoring system administrator side completes the work: the process of classification is of the set list, so that students end the process of monitoring carried out according to this list. Monitoring remote computer screen, and if it is necessary, it controls the mouse and keyboard operation. It can send command messages to the remote computer or a prompt message. Real-time monitoring log information according to the system administrator to set the alarm, or access to log records, operational records to find students. Monitoring system administrator side includes the process monitoring, remote computer screen monitoring, remote message, log monitoring of four sub-modules. Process control modules: process control function is the process of setting, real-time updated list of students to end the process of classification and treatment unknown process; process of setting the corresponding process is to add modify delete operation table for the students to end the process of monitoring carried out to provide the basis to judge. The process of classification list includes the following main properties: the process name, and the process path flag. Process name and process the path are used to distinguish different processes to ensure its uniqueness and avoid a miscarriage of justice in the process of monitoring time. Flag is used to distinguish the black list, white list and application development platform are for three types of processes. Students end the process of classification in real-time updates the list to ensure that the process of the database table with the list of students classified information side of the consistency of the process to improve the resilience of the process of monitoring. Dealing with the unknown process is found during the previous machine to process the unknown process. Remote computer screen monitoring modules: screen monitor features include: screen monitor: client computer by monitoring the students to observe real-time screen on the machine case of students; remote control: if it is necessary, control student client computers, keyboard and mouse. Screen monitor is the first time by way of poll to see all the students end computer screen image, if it is necessary, it can be a single computer for remote control. Remote message sub-module: Remote message function of student client computer messages or prompts an order message. Log monitoring modules: features include real-time alarm, search logs, access logs and log maintenance. Real-time alarm is timed according to the setting of alarm conditions to search for the corresponding log information and to remind the administrator to the appropriate treatment to improve the adaptability of the monitoring program. Search function is in accordance with certain

conditions, the log records are in search of qualified log. Search log function is to view detailed information on a particular log. Log maintenance is to back up the log as a text file and delete the system log expired.

5 Conclusion

This paper introduces the basic conditions of the engine room, and it is in accordance with the system structure to provide an overall planning and design; then it introduces the system of two modules: data management module and network monitoring module function, and their links between two major modules are respectively corresponding to the various sub-module functions described.

References

[1] Mobasher, B., Cooley, R., Jaideep, S., et al.: The Design of Automatization Management System of Computer Rooms. IEEE Press, New York (1999)
[2] Shahabi, C., Zarkesh, A.M.: Design of Distributed Management System for Computer Laboratories in Universities. IEEE Press, Birmingham (2001)
[3] Yan, T., Jacobesn, M.: Development of Distributed Management System to Computer Laboratory Groups in University Campus Network. WAM Press, Paris (1999)
[4] Xu, L.: The The Implementation and Application of Distributed Management System to Computer Laboratory Groups in University Campus Networks. Anchorage Press (2001)
[5] Bezdek, J.C.: Development of a WEB-based administration system on multi-campus computer laboratories network. Ph. D. D, Mathematics, Cornell University, Ithaca, New York (1993)
[6] Kamel, M., Selim, S.Z.: Development and Application of A Hospital Management System Based on Campus Network 27(3), 421–428 (1994)
[7] Zadeh, L.A.: Developing of a Book Management System Operating on Network 8, 338–353 (2001)
[8] Kamel, M., Selim, S.Z.: Design of the General Integrated Environment for Multimedia Teaching System Based on Campus Network 61, 177–188 (1994)
[9] Al-Sultan, K.S., Selim, S.Z.: The Integration Between Information Technology and Physics Course Teaching And Its Teaching Application, 26(9): 1357–1361 (1993)
[10] Miyamoto, S., Nakayama, K.I.: The automatic system of computer-management on LAN 16(3), 479–482 (2001)
[11] Bezdek, J.C., Hathaway, R., Sabin, M., Tucker, W.: Realization of the Computer Laboratory Management System Based on the Network Database C/S Mode. Cybernet 17(5), 873–877 (2001)
[12] Kameci, M., Selina, S.Z.: Establishment of Networkedized Management Pattern for Modern Computer Room 27(3), 421–428 (1994)
[13] Jamel, M., Salim, S.Z.: Composition Module of Open Management System and Application of Management Mode in Computer Laboratory 61, 177–188 (2005)
[14] Al-Susan, K.S., Selim, S.Z.: Track Basic Equipment Detecting Information Management System based on Web 26(9), 1357–1361 (2003)
[15] Meyanmoto, S., Nakayama, K.I.: Design and Implementation of Computer Lab Management System Based on Web Service (2002)

The Application of Online Traffic System on the Wireless Internet

Zhiqiang Yao[1], Chiyuan Li[2], Ke Xu, and Han Xia[3]

[1] Computer Science Dept. North China Institute of Aerospace Engineering, Langfang, China
[2] Fire Engineering Dept. The Chinese People's Armed Police Force Academy, Langfang, China
[3] Department of Information Engineering, Qinhuangdao Institute of Technology, Qinhuangdao, China
[4] Finance Dept. North China Institute of Aerospace Engineering, Langfang, China
yaonciae@126.com, chinafire_lcy@126.com, qhdxuke@126.com, xiahannciae@126.com

Abstract. Wireless internet is gaining wide spread usage in traffic applications. This paper explores how simple busline query applications can be developed for mobile devices and introduces general framework for designing such online traffic applications. Using the framework, based on Microsoft database software SQL Server 2000, T-SQL clause is employed to program one stored procedure for busline query, which applies the depth searching algorithm. The framework utilizes mobile phones simulator software, Nokia WAP Toolkit, which interacts with a server-side ASP (Active Server Pages) in conjunction with the associated SQL Server database. The design and implementation of the management information system are discussed in the paper, and test results are presented.

Keywords: Traffic Applications, Stored Procedure, Depth Searching Algorithm, WAP Toolkit, ASP, SQL Server.

1 Introduction

Management information systems[1] are becoming more and more popular in the information age[1] and the wireless Internet[2] is a new revolution upon us, one that will affect the world on a scale similar to that of the wired Internet. This paper demonstrates the general framework for designing online traffic management information system on the wireless internet. In the urban traffic system[3][4], busline query is very popular; for the purpose of demonstration of the framework, busline query application on the wireless network is developed. This framework utilizes Microsoft Active Server Pages (ASP)[5] in conjunction with the associated SQL Server[6] database to create a truly dynamic application using WML (Wireless Markup Language)[7] to develop application on the mobile phones.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1063–1070.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

2 Database Design and Implementation

2.1 Design the Database

SQL Server 2000[6] is used as the datase to provide the solution for online traffic query. At first, we design one database: Traffic and in this database, we design the related tables to save traffic data: Bus_Line and Bus_Stop.The design of these two tables are shown in TABLE 1 and TABLE 2.

Table 1. Design of the data table : Bus_line

ColumnName	Data Type	Remarks
BusID	int	Automatic Increase by 1
BuslineName	Varchar(30)	BusLine Name
isSameRoute	bit	If the back route is the same as the going route, isSameRoute equals 1, or else, it equals 0.

Table 2. Design of the data table : Bus_station

ColumnName	Data Type	Remarks
BusID	int	Reference the BusID column of table BusLine
Station	Varchar(30)	BusStation Name
Orders	int	The order of bus station in its busline

2.2 Program the Stored Procedure

Every busline has a fixed going line, every going line has fixed stations and the order of every station in its line is fixed; if different lines pass by one same station, we can exchange bus lines at this station. If we set out from one named station, the data for going to the destination can be seen like a tree, as to one station in several different bus lines, at any one busline, the before station and the next station can be seen as the child nodes of this station. So we can program the stored procedure[6][8] with the depth searching algorithm[9] which is commonly used for tree-like data. When we search until the destination station, we query out the search resut.

In addition, the loop limitation in the flow chart of the stored procedure (Fig. 1) is based on such conditons:

- If bus lines are different and bus stations are are not different, exchange is done (Symbol ‘ex’ is used in the query result).
- If bus lines are not different, bus statioins are different and the back route is the same as the going route, both the before node and the after node are the child nodes of the bus station.
- If bus lines are not different, bus statioins are different and the back route is different from the going route, only the after node is the child node of the bus station.

- The deepest level is where bus station equals @end_Station.
- The string length of the query result in the table #temp is less than 1000; that is to say, during the search, if there is no solution to reach the end station within the max length 1000, we assume there is no query result.

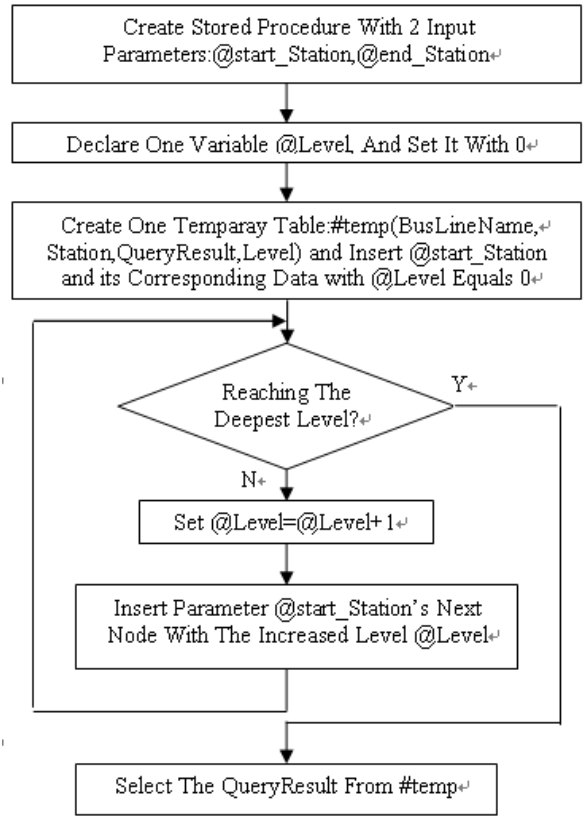


Fig. 1. Flow Chart of the Stored Procedure Based on the Depth Searching Algorithm

In respect to the two tables:Bus_Line and Bus_Station, we design one view[6] VW_Bus for the convenience of query.

Create View VW_Bus As

Select a.BusLineName,b.Station,a.isSameRoute,b.Orders From Bus_Line a, Bus_Station b Where a.BusID=b.BusID

According to the flow chart shown in Fig. 1 and the above illusation, on the platform of SQL Server 2000, we use the processing tool-----Query Analyzer[6] to program the stored procedure SP_BusLineQuery with T-SQL[6] and the source codes are the followings:

```
CREATE PROC SP_BusLineQuery
@start_Station varchar(10),@end_Station varchar(10) AS
BEGIN DECLARE @level int
SET @level=0
SELECT BusLineName,Station,QueryResult=
CAST(BusLineName+': '+Station as varchar(1000)),
Orders=Orders,[Level]=@level INTO #temp FROM VW_BUS WHERE
Station=@start_Station
WHILE @@ROWCOUNT>0 AND NOT EXISTS(
SELECT * FROM # WHERE Station=@end_Station)
BEGIN SET @level=@level+1
INSERT#temp( QueryResult,BusLineName,Station,Orders,[Level]) SELECT
QueryResult=a. QueryResult +CASE
WHEN a.BusLineName=b.BusLineName THEN '->'+b.Station ELSE ' ex
'+b.BusLineName+': '+b.Station END,b.BusLineName,b.Station,b.Orders,@level
FROM #temp a,VW_BUS b WHERE a.[Level]=@level-1 AND (
a.Station=b.Station AND a.BusLineName<>b.BusLineName
OR a.BusLineName=b.BusLineName AND
(b.isSameRoute=1 AND (a.Orders=b.Orders+1 OR a.Orders=b.Orders-1) OR
b.isSameRoute=0 AND a.Orders=b.Orders-1)) AND LEN(a.QueryResult)<1000
AND
 PATINDEX('%[ >]'+b.Station+'[-)]%',a.QueryResult)=0
END SELECT 'solution'=Line FROM #temp WHERE
 [Level]=@level AND Station=@end_Station END
```

2.3 Testing the Stored Procedure

We insert these sample data into the database and test the stored procedure on Query Analyzer[6].

In the sample data, we assume there are 3 bus lines:

BusNo.1 which passes by A,B,C,D,E,D,C and A with the isSameRoute equals 0.
BusNo.4 which passes by A,B,C,D,E and F with the isSameRoute equals 1.
BusNo.7, which passes by B,C,G,H and I with the isSameRoute equals 1.

The testing result is shown in Fig.2

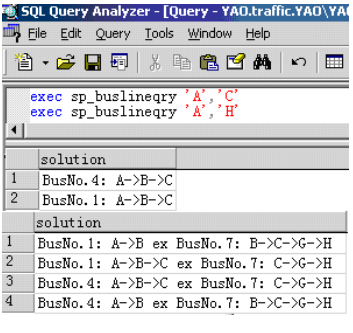


Fig. 2. Test of the Stored Procedure sp_buslineqry

3 Software Architecture for Busline Query Applicaion

Dynamic applications are applications that build content “on the fly” in response to request made. They can reflect the latest traffic information from the database. Microsoft Active Server Pages (ASP)[5] in conjunction with the associated SQL Server[6] database to create a truly dynamic traffic application. The conceptual software architecture is shown in Fig.3 and the architecture contains the following components:

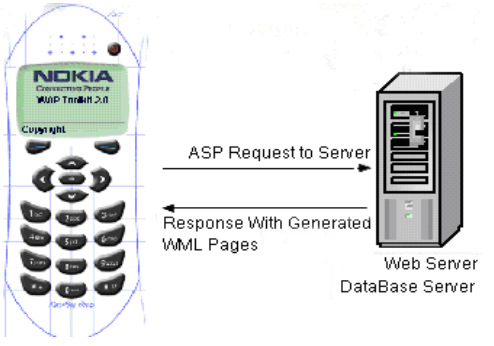


Fig. 3. Architecture of the Dynamic Application

3.1 WEB Server

IIS (Internet Information Service)[10] is configured to create one virtual directory(Fig.4.) containing one file: query.asp, which connects the database server, gets the query results from the database server and responds with generated WML pages.

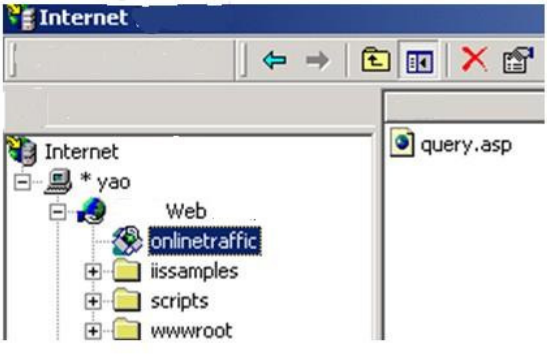


Fig. 4. Configuration of IIS to Create Virtual Directory

4 Application Design

The client software was developed using Nokia Wap Toolkit[2] 2.0. This kit enables developers to create applications on the mobile phone based on the WAP (Wireless Application Protocol)[2]. Wireless Markup Language(WML) is used to implement the Wireless Application Protocol (WAP) specification and creates pages that can be displayed in a WAP browser. WMLScript[7], a client-side scripting language, is used for developing tasks such as user input validation, generation of error message and other Dialog boxes etc. on the mobile phone. The Toolkit contains tools for editing both WML and WMLScript, and it can display the resulting code on a mobile service simulator (Fig.5). Fig.6, Fig.7 and Fig.8 illustrate the desired outcome of the busline query application.

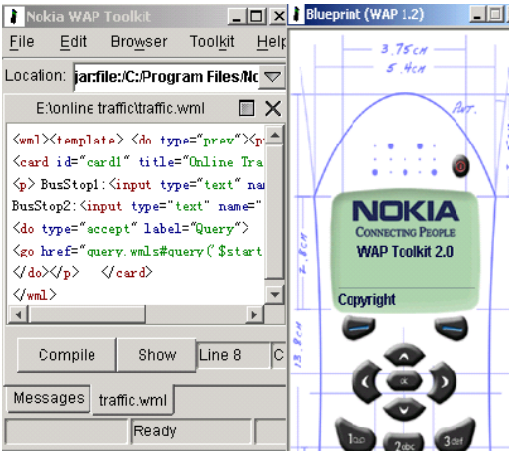


Fig. 5. Nokia Wap Toolkit for Editing WML

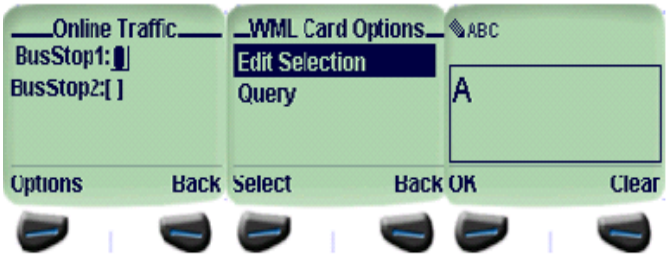


Fig. 6. Keying the BusStop1 and BusStop2 to Query

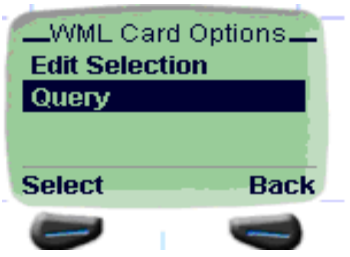


Fig. 7. Pressing the Query Label to Query

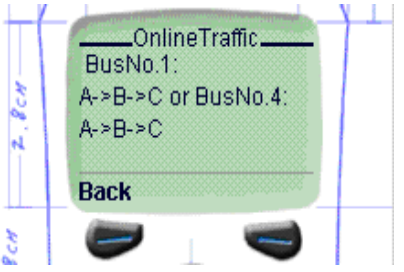


Fig. 8. Viewing the Result of the Busline Query or WMLScript and Displaying the Resulting Code on a Mobile Service Simulator

The application consists of the following source files:

traffic.wml: this file is stored on the client mobile phone and is responsible for displaying the user interface and calls the query function defined in query.wmls script to display the busline query result.

query.wmls: this file is stored on the client mobile phone and is invoked by traffic.wml when the client wants to query the busline from one stop to another. It sends ASP[9] request to the query.asp file stored in Web server.

query.asp: this file is stored in the Web server[11] and is automatically invoked when query.wmls script is executed in the client mobile phone. It connects to the database which is stored in Database server, gets the query results from the database server and responds with generated WML pages[7].

5 Testing of the Programs

In order to test programs, Nokia Wap Toolkit[2] is installed to the Web Server whose system software is Window 2000 Server[10]. If we want to go from station A to station H, Fig. 9 shows the resulting screens on the WAP Toolkit, we can exchange busline at station B or station C.

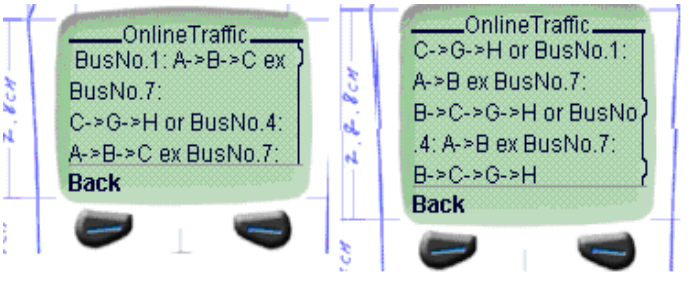


Fig. 9. Results for Online Traffic From BusStop A to BusStop H

6 Summary

With the help of the depth search algorithm, the stored procedure of busline query is created and tested successfully on the platform of the SQL Server 2000 database; With the application of WAP Toolkit, the WML and WMLScript web files are designed to send the users' request to the Web server, the corresponding dynamic ASP programs are developed to process the request and acquire the query result from the database server and finally the busline query result is shown on the mobile phones with the dynamically generated wml files and the testing result is satisfactory.

This conceptual design framework can easily be used for any traffic applications, such as subway traffic, airline traffic and railway traffic applications. As more and more online traffic applications are conducted using mobile phones, this design framework is flexible enough to be used on the wireless internet.

References

[1] Oz, E.: Management Information System Six Editon. Course Technology, 208–254 (January 2008)
[2] Fife, R., Lee, W.M.: Webmaster's Guide to the Wireless Internet. Syngress, 114–160 (December 2001)
[3] Ma, L., Liu, X., Liao, D.: The Shortest Path of Public Transport Net and Related Optimal Taking Bus Problem. Mathematics in Practice and Theory 34(6) (2004) (in Chinese)
[4] Zhu, Y., Wang, J., Lu, H., Zhang, J.: Database Technique Applied to the Urban Traffic Information Service System. Systems Engineering-Theory & Practice 25(12) (2005) (in Chinese)
[5] Mitchell, S.: Designing Active Server Pages, pp. 182–203. O'Reilly (September 2000)
[6] Thernstrom, T., Weber, A., Hotek, M.: MCTS Self-Paced Training Kit (Exam 70-433): Microsoft® SQL Server® 2008—Database Development, pp. 306–453. Microsoft Press (February 2009)
[7] Frost, M.: Learning WML & WMLScript, pp. 54–134. O'Reilly (October 2000)
[8] Cao, Y.: Application with SQL Server Procedure in System Development. Modern Electronics Technique 28(20) (2005)
[9] Ouzzani, M., Bouguettaya, A.: Query Processing and Optimization on the Web. Distributed and Parallel Databases 15, 187–218 (2004)
[10] Liu, L., Yang, G.: Windows 2000/XP NetWork Construction Technology Application, pp. 241–276. Hope Electronics Press, Beijing (2002) (in Chinese)
[11] He, L., Chen, Y., Ding, R., Zhang, S.: The Informationalization of Small and Medium-size Business Based on the ASP Platform. In: 2008-The Seventh Wuhan International E-Business Conference (2008)

A New Trusted Monitoring Algorithm to IA Based Distributed System

Shao Liang and Wu Fang

Department of Humanity and Information, Zhejiang College of Construction
Hangzhou Zhejiang 311231, China
shaoliang.hz@gmail.com, wf552554@163.com

Abstract. In order to increase the control capability of network system, it is essential to have a network monitoring which can greatly improve its reliability and stability. Considering to how to improve the reliability of network monitoring, whether the monitoring result is trusted or not? This paper proposed a trusted monitoring algorithm in IA (Intelligent Agent) distributed system with trusted computing technology. After extracting and analyzing the monitoring delay and mutation parameter of node states. It is able to trustily predict the node states and realize trusted monitoring. The experimental result is shown that, the results can keep steady when node states are changed and it validates the good robust of algorithm. It provides a trusted nodes states map for system communication and execution.

Keywords: IA, trusted computing, trusted monitoring, distributed system.

1 Introduction

Owing to the expansion of network scale and the increase of network services and applications Network has in-depth people's daily life. However, there are also some network security problems along with it, and network security situation is very grim. Network monitoring, as a key part of the network management, can timely discover problems before the network performance reduce to its deadline, by collecting data、monitoring network quality of service、and robust performance analysis and display capabilities. So it can improve the reliability and stability of network [1].

What's more, in order to solve the unsafe problem of computer and network structure, the basic idea of trusted computing [2] is created. Trusted means an entity whose behavior in the realization of the given target is always the same as expected [3], and emphasizes the results of the behavior is predicable and controllable. Trusted computing means the components、operations and processes involved in calculating are expected in any case, and able to protect from the physical virus. It is essentially measure reliability for computing environment with trust metric technology, realize the basic framework by roots of trust and chain of trust [4]. But there are some problems here: (1) theoretical research lags behind practical research, (2) some key technological yet to be overcome, (3) Lack of supporting and trust software system, (4) lack of the combination of security mechanism and fault-tolerant mechanism [5].

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1071–1079.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

For IA system, the reliability of the executing ability and executing result, by each IA entity, has a directly effect on system or its reliability. This is a problem deserving of study that how to build a trusted system model and how to realize a trusted network monitor model based on IA[6]. In Paper [7] a computation trust model is proposed in which the confidence information based in direct prior interactions with the target agent and the reputation information from trust network are used. Sarvapli D. Ramchurn, Nicholas R. Jennings, Carles Sierra and Luis Godo [8] develop such a trust model, based on confidence and reputation, and show how it can be concretely applied to guide agents in evaluating past interactions and in establishing new contracts with one another.

This paper, based on the study of IA technology and trusted computing technology, proposed a trusted monitoring algorithm in IA distributed system. After extracting and analyzing the monitoring delay and mutation parameter of node states. Combined with Confidence、 Reputation and basal factor. It is able to trustily predict the node states and realize trusted monitoring. The experimental result is shown that, the results can keep steady when node states are changed and it validates the good robust of algorithm. It provides a trusted nodes states map for system communication and execution.

2 The Framework and Network Structure of IA Distributed System

2.1 Wireless Network Overview

In this paper, the IA distributed system represented by a triple {A,S,N}[9], in which A represents an IA system consisting of cooperating and communicating IA that can migrate autonomously from node to node (Figure 1) , S is a set of m nodes on which the IA perform operations; and N is a network that connects nodes and assures IA mobility.

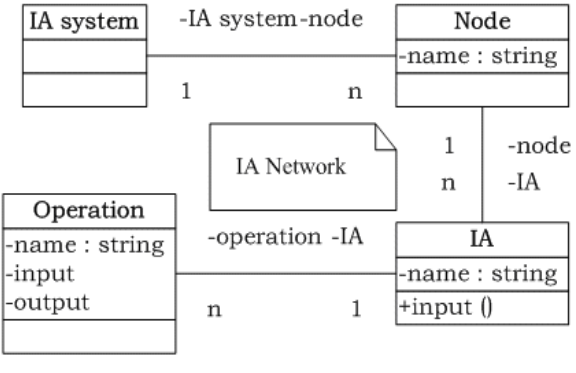


Fig. 1. The Class Diagram of IA System

The network of system is divided into there levels: The 1st-level, 2nd-level and 3rd-level. Their network nodes are called 1st-level nodes, 2nd-level nodes and 3rd-level nodes. The grade of those nodes is determined by its network level, 1st level nodes >2nd level nodes >3rd level nodes. And the higher the grade of a node is, the more powerful its authority is. Figure 2 is shown that 1st-level nodes are the most powerful nodes. A 1st-level node can manage 2nd-level nodes and 3rd-level nodes, and distribute tasks to them, when whose nodes are in its branch. The 2nd-level node should listen to the command of its 1st-level node, and also can manage its 3rd-level nodes. The 3rd-level node has the lowest power; it should listen to both of its higher level nodes. Therefore the high level node can manage low level node, and the low level node should listen to the command of its high level node.

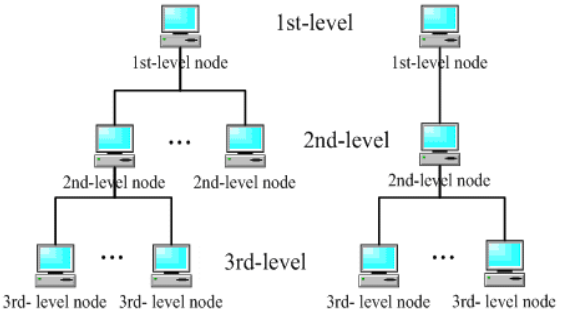


Fig. 2. The Logic of System Network

3 The Source of Algorithm

The means of confidence and reputation are explained above.

It is supposed that an event E need to be handled on one node. In this case, an IA, which is trusted by system, should to be searched. It is possible that there are several competitive IAs, for all of them can handle the event. On this circumstance, it is necessary that how to find the trusted IA. Generally speaking, the IAs take confidence and reputation as parameters to predict the trusted IA, and handle the event with it.

3.1 Confidence

If this IA has an experience of handling event E on this node, and the experience is stored in database. Its execute ability and execute time are taken for direct judgment whether it is competent to this event or not.

3.2 Reputation

If this IA has an experience of handling event E on other nodes, and the experience is stored in database. Its execute ability and execute time are taken for indirect judgment whether it is competent to this event or not.

As mentioned above, this paper, Confidence and Reputation are applied to a trusted monitoring in network. It proposed some new concepts as basal factor、Confidence、Reputation and monitoring trusted factor, and also explained their computational method. Then, a new network trusted monitoring algorithm is proposed.

Basal Parameters of Algorithm

Here, let's supposed that there is a node called A in IA distributed system, and just consider network monitoring to its sub-nodes. The number of sub-nodes is M, and D is one of them, as $D \in \{1,2,3,\cdots,M-1,M\}$ .assumed that i is the immediate monitoring times from A to D, t_i is the time when the monitoring of number i happened. Once a monitoring is realized, parameter τ and $N(t_i)$ are extracted for this trusted predictability, where τ is the monitoring delay, and $N(t_i)$ is the mutation parameter of node states. As time went on, $N(t_i)$ changed, and is always proportionate to the confident states of purpose node. After actual measurement, the accuracy of monitoring delay can be one thousandth (10-3) of a second. It has a great influence on prediction of trusted monitoring algorithm, if two monitoring nods have system time error. In such circumstances, time synchronization is a good method to avoid this error.

According to the trusted predictability above, and combined with the real especially system environment, each trusted factor is introduced below:

3.3 Basal Factor

When node A monitors node D by its monitoring IA, it will analysis the feedback collected from node D. where the extracted period is T. there are two judgments: normal set "1"; abnormal set "0". In this case, the basal factor is defended by two expressions.

At time t_i, when node D is judged to be "1" by monitoring IA, the basal factor will be:

$$p_{Di} = \begin{cases} 1-\frac{2\tau(N(t_i)-N(t_i-T))}{T}, t_i \geq T \\ 1-\frac{2\tau \times N(t_i)}{t_i}, t_i < T \end{cases} ; \qquad i \in \{1,2,\cdots,n\} \tag{1}$$

At time t_i , according to system setting, when node D is judged to be "0" by monitoring IA, the monitoring will be executed 3 times, so the basal factor will be:

$$p_{Di} = \begin{cases} 1-\frac{2\tau_1(N(t_{i1})-N(t_{i1}-T))}{T} \times \frac{2\tau_2(N(t_{i2})-N(t_{i2}-T))}{T} \times \frac{2\tau_3(N(t_{i3})-N(t_{i3}-T))}{T}, t_{ix} \geq T \\ 1-\frac{2\tau_1 N(t_{i1})}{t_{i1}} \times \frac{2\tau_2((t_{i2})}{t_{i2}} \times \frac{2\tau_3(t_{i3})}{t_{i3}}, t_{ix} < T \end{cases} \tag{2}$$

To node D,
where $P=\{p_{D1},p_{D2},\cdots p_{Di},\cdots p_{Dn}\}$, and $p_{Di}\in P_D$, $i\in\{1,2,\cdots,n\}$.

To node A,
where $P=\{P_1,P_2,\cdots P_M\}$, and $P_D\in P$, $P=\{P_1,P_2,\cdots P_M\}$.

3.4 Confidence

Assume node A and sub-node D, where $D\in\{1,2,3,\cdots,M-1,M\}$. So at time t_i, the Confidence about sub-node D is defended: it is the mean value of k times monitoring trusted factors before this monitoring, and the monitored node is D. If immediate monitoring time is less then confidence parameter ($i\le k$), the confidence will be the mean value of $i-1$ times monitoring trusted factors before this monitoring.

So the Confidence ($Con(i)$) will be:

$$Con(i)=\begin{cases}\dfrac{\sum_{j=i-k}^{i-1}Tr_D(j)}{k}, k\le i\\ \dfrac{\sum_{j=1}^{i-1}Tr_D(j)}{i-1}, k>i\end{cases} D\in O, O=\{1,2,3,\cdots,M-1,M\}, \tag{3}$$

3.5 Reputation

The Reputation of node D is defended: it is the mean value of k times monitoring trusted factors before this monitoring, and the monitored node is other node but not node D. If immediate monitoring time is less then confidence parameter ($i\le k$), the reputation will be the mean value of $i-1$ times monitoring trusted factors before this monitoring.

So the Reputation ($\mathrm{Re}\, p(i)$) will be:

$$\mathrm{Re}\, p(i)=\begin{cases}\dfrac{\sum_{D'\cup O'}\sum_{j=i-k}^{i-1}Tr_{D'}(j)}{k\times(M-1)}, k\le i\\ \dfrac{\sum_{D'\cup O'}\sum_{j=1}^{i-1}Tr_{D'}(j)}{(i-1)\times(M-1)}, k>i\end{cases} D'\in O', O'=\{1,2,3,\cdots D-1,D+1,\cdots,M\} \tag{4}$$

Trusted Monitoring Algorithm

The chapters above, new definitions are given about basal factor, confidence and reputation. There is a new definition, Based on different time:

$$Hi(i)=\left[v(i)\times Con(i)+(1-v(i))\times\mathrm{Re}\, p(i)\right] \tag{5}$$

Where $v(i)$ ($v(i) \in [0,1]$) is confidence trusted coefficient:

$$v(i) = \frac{Con(i)}{Con(i) + \mathrm{Re}\, p(i)} \tag{6}$$

So immediate monitoring trusted factor is defended as :

$$Tr_D(i) = x(i) \times p_{Di} + (1 - x(i)) Hi(i) \tag{7}$$

And $Tr_D(0) = 1$,where $x(i)$ ($x(i) \in [0,1]$) is immediate monitoring parameter:

$$x(i) = \frac{p_{Di}}{p_{Di} + Hi(i)} \tag{8}$$

By equation (5), immediate monitoring factor can be explained as:

$$Tr_D(i) = x(i) \times p_{Di} + (1 - x(i)) \left[v(i) \times \frac{\sum_{j=i-k}^{i} Tr_D(j)}{k} + (1 - v(i)) \times \frac{\sum_{D' \cup O'} \sum_{j=i-k}^{i} Tr_{D'}(j)}{k \times (M-1)} \right] \tag{9}$$

4 Simulation and Analysis

In the view of the computation of monitoring trusted factor, and combined with actual measurement monitoring delay τ ,the judgment states of monitored node can be divided kind: if $0.01 \le \tau \le 0.03(s)$, the judgment states will be set "1"; if $0.46 \le \tau \le 0.52(s)$, the judgment states will be set "1". Here, in system, assumed node A has 5 sub-nodes ($M = 5$). The test time will be $t = 3600 \times 3(s)$ this time. The extracted period $T = 600(s)$, and the monitoring period $T' = 60(s)$, $k = 60$ (extracting and analyzing k times monitoring trusted results before this monitoring).the state mutation time of sun-nodes is shown as table 1: assumed that at the beginning of test all of sub-nodes are normal (set "1"), and T1、T2、T3、T4、T5、T6、T7 are mutation time.

Table 1. The state mutation time of sub-nodes

Sub-node	T1/(s)	T2(s)	T3(s)	T4(s)	T5(s)	T6(s)	T7(s)
Node1	1800	2700	4500	6300	7800	9300	10800
Node2	200	2900	3700	5900	7200	9000	10800
Node3	350	909	1808	2807	5490	7155	10800
Node4	850	1260	3694	6800	7315	9840	10800
Node5	999	1999	2999	3999	4999	8900	10800

So the state of node 1 is shown as figure 3 in time t, and the mutation parameter of node 1 $N(t_i)$ is shown as figure 4.

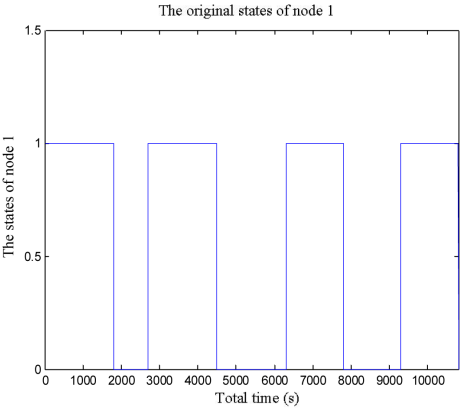


Fig. 3. The States of Node 1

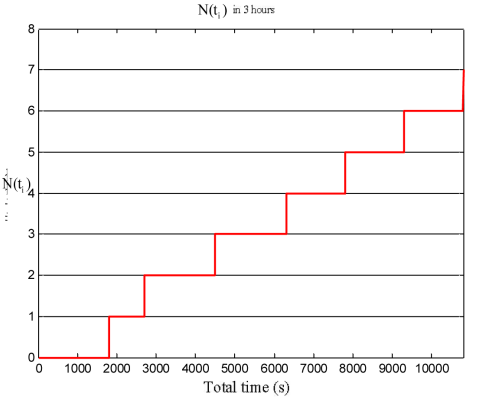


Fig. 4. The Mutation Parameter of Node 1

Firstly, According to the monitoring delay τ and immediate time t_i with network monitoring, the basal trusted factor p_{Di} will be obtained by equation (1)、(2). Secondly, confidence $Con(i)$ and reputation $\mathrm{Re}\, p(i)$ will be obtained by equation (3)、(4). Finally, the immediate monitoring trusted factor $Tr_D(i)$ can be obtained by equitation (9) with confidence trusted coefficient $v(i)$ and immediate monitoring parameter $x(i)$.

Owning to different monitoring delay for node A, its monitoring trusted factor is shown as figure 5, solid-line means $0.01 \le \tau \le 0.03(s)$ and $0.46 \le \tau \le 0.52(s)$; dashed-line means $0.1 \le \tau \le 0.3(s)$ and $4.6 \le \tau \le 5.2(s)$ dotted-line means: $1 \le \tau \le 3(s)$ and $46 \le \tau \le 52(s)$ .it shows that the monitoring delay has great influence on monitoring trusted factor. The shorter the monitoring delay is, the higher the monitoring trusted factor will be. What's more, the solid-line expresses monitoring trusted factor is given by actual measurement monitoring delay, and such phenomenon shows the monitoring algorithm is trusted.

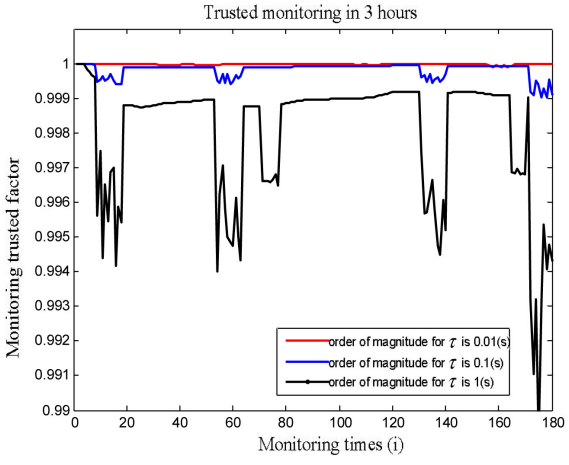


Fig. 5. Different Monitoring Factors with Different Monitoring Delays

5 Conlusion

In order to increase the reliability of network monitoring, this paper proposed a trusted monitoring algorithm in IA distributed system, developed by Python, with trusted computing technology. It is able to trustily predict the node states and realize trusted monitoring. The experimental result is shown that, the results can keep steady when node states are changed and it validates the good robust of algorithm. It provides a trusted nodes states map for system communication and execution.

References

[1] Fang, W., Zhijin, Z., Xueyi, Y.: A New Dynamic Network Monitoring Method based on IA. In: ISCSCT 2008, Shanghai, China, December 20-22 (2008)

[2] Trusted Computing Group. TCG Specification Architecture Overview specification Revision1.4[S] (August 2007), https://www.trustedcomputinggroup.org

[3] Trusted Computing Group. Main Specification Version 1.lb[S] (February 2002), https://www.trustedcomputinggroup.org

[4] Trusted Computing Group, TCG Infrastructure Working Group Reference Architecture for Interoperability (Part l) Specification Version1.0[S] (June 2005), https://www.trustedcomputinggroup.org
[5] Shen, C.X.: The study on trusted computering technology, http://www4.it168.com/jtzt/shenlan/tech/cedc2007/zhibo/ftp/k1.ppt
[6] Wu, X., Shen, Z., Zhang, H.: The Mobile Agent Security Enhanced by Trusted Computing Technology. In: International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2006, pp. 1–4. IEEE (September 2006)
[7] Wang, P., Zhang, Z.: A Computation Trust Model with Trust Network in Multi-Agent Systems. In: Proceedings of the 2005 International Conference on Active Media Technology, AMT 2005, pp. 389–392. IEEE (May 2005)
[8] Ramchurn, S.D., Jennings, N.R.: Devising Trust Model for Multi-Agent Interactions Using Confidence and Reputation. Applied Artificial Intelligence 18, 833–852 (2004)
[9] Fang, W., Zhijin, Z., Xueyi, Y.: IA Distributed System based on Python. Computer Engineering, 20 (2009)

A Study of the Impact of Electronic Processing Time on Performance in Optical Burst Switching Networks

G.M. Li, Z.K. Sun, and S.H. Geng

School of Information Engineering, SHANDONG University, China
gmli@sdu.edu.cn.

Abstract. In this paper, we mainly study the impact of electronic processing parameters on the network performance in optical burst switching (OBS) networks. The electronic processing mainly includes control packet processing time and switch configuration time. Our analysis shows that the control packet electronic processing time is the primary parameter closely related to the network performance. We also provide switching constraint resolutions to improve network performance.

Keywords: optical burst switching, electronic processing time.

1 Introduction

Wavelength division multiplexing (WDM) provides a tremendous transmission capacity with a single fiber. With the rapid development of WDM, the optical-to-electronic (O/E) and electronic-to-optical (E/O) conversion at each node becomes a critical bottleneck in terms of processing speed, cost, and power consumption. To resolve the bottleneck, there are three main candidates: optical circuit switching (OCS), optical packet switching (OPS) and optical burst switching (OBS). OCS uses two-way reservation to establish a point-to-point connection and no optical buffers are needed, but it suffers the low bandwidth utilization efficiency for bursty traffic. OPS offers much finer granularity and has much higher bandwidth utilization efficiency. However, it is an ideal choice at present since its required technologies, such optical buffer and optical processing, are immature. OBS provides the inter-granularity between OCS and OPS with the aim to combine the advantages of both. To cut down the long end-to-end delay associated with two-way reservation in OCS, OBS uses one-way reservation; to avoid the drawback of low bandwidth utilization efficiency for bursty traffic in OCS networks, OBS only allocates the network resource for the period of transmitting the data burst. On the other hand, to avoid the current technical limitations on OPS, OBS has no requirements on optical buffer and optical processing capability. Thus, OBS has been considered as a promising approach for the next generation network in the near term. Several testbeds have demonstrated the feasibility of OBS networks [1-3].

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1081–1089.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

In an OBS network, ingress edge nodes assemble packets into data burst (DBs). DBs are kept in the all-optical domain and have corresponding control packets (CPs) transmitted in separate channels along the same route. CP is sent ahead of its corresponding DB and this interval is named offset. CP is processed electronically at each intermediate node in order to schedule its pending DB to an output port without the necessity of optical buffer. At egress nodes, DBs are dis-assembled back into packets.

Existing work on OBS networks can be categorized into several major issues, such as burst assembly mechanism, resource allocation scheme, burst contention resolution and Quality of service (QoS) provision. Among them, resource allocation scheme is the primary one. Various schemes have been proposed and studied [1,4-7]. Several papers have investigated the relationships between various allocation schemes and network parameters [8-12]. However, the conclusions are inconsistent. In this paper, we aim to clarify the relationships between network input parameters and network performance parameters. Since electronic processing time on control packets is the most important input parameters, we focus on the impact of electronic processing time on other input parameters and network performance. In OBS networks, the electronic processing time mainly includes control packet processing time and switch configuration time. In Section 2, we analyze the impact of control packet processing time on network performance. Section 3 studies switch configuration time, and give switching constraint resolutions. Finally, we conclude in Section 4.

2 Impacts of Control Packet Processing Time

In OBS networks, since a control packet need allocate network resources for its corresponding data burst at each intermediate node from the source to the destination, it need be sent out in advance to perform electronic processing at multiple core nodes until it reaches the destination. The time gap between the control packet and the data burst is donated as offset. At source node, the minimum value of offset, the initial offset $(T_{off})_0$ is the sum of the control packet processing time $(T_{cp})_i$, i = 1,2,…,N at each node along the route and the maximum switch configuration time $\max\{(T_{sw})_i\}$ along the route , i.e..

$$(T_{off})_0 = \sum_{i=1}^{N} (T_{cp})_i + \max\{(T_{sw})_i, i = 1,2,\cdots,N\} \tag{1}$$

Usually, we may assume the control packet processing time $(T_{cp})_i$ at each node to be identical and the switch configuration time to be the same (T_{sw}). Thus, the initial offset of a route with H hops can be written as

$$(T_{off})_0 = H \times T_{cp} + T_{sw} \tag{2}$$

The control packet processing time is electronic-hardware dependent and is also determined by the computational complexity of the employed resource allocation scheme.

In OBS networks, each fiber link usually manipulates multiple wavelengths based on the WDM technology and the common assumption is that each node has full-range wavelength conversion capability. The scheduling problem is to assign the arriving data burst to an appropriate wavelength in the desired output port. According to the extent of utilizing information of previously scheduled data bursts, scheduling algorithms can be categorized into non-void filling and void filling algorithms. Typical non-void filling algorithms include First-Fit (FF) algorithm and the Latest Available Unscheduled Channel (LAUC) algorithm. Typical algorithms include the First-Fit with void filling (FFVF) algorithm, the Latest Available Unused Channel with void filling (LAUC-VF) algorithm.

We use simulations to investigate the performance of two types of OBS reservation schemes and various scheduling algorithms. We assume that the data burst traffic follows Poisson distribution, and the data burst length is exponentially distributed. A typical network topology, NSFNET, has been investigated, as shown in Figure 1. We assume each link has one dedicated wavelength as control channel and ten wavelengths as data channels. We also follow the popular assumption that each node has the capability of full range wavelength conversion. Each node in the topology may be an ingress/egress node or a core node, and all nodes are assumed to receive the same offered load in the network. When a new data burst arrives at an ingress node, it randomly chooses a destination from the rest of the nodes in the network and the shortest path routing is used. All simulations are sun sufficiently long such that the 95% confidence intervals are less than 1% of the results, thus, we omit them from all figures to improve readability.

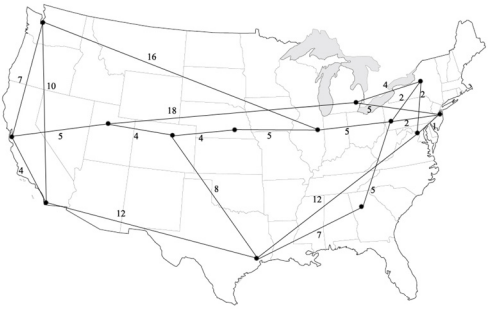


Fig. 1. The NSFNET topology, 1991 (the original map is available at ftp.uu.net/inet/maps/nsfnet/).

In Figure 2 where the overhead is equal to 0.1 times of the unit, we also observe that the deferred reservation scheme has slightly lower loss probability than the immediate reservation when the network is not heavily loaded (the load is below 0.8). This is due to the fact that the possibility of successfully reserving network resources after a delay is higher than that of reserving resources immediately. The second observation is that all scheduling algorithms of the deferred reservation have very similar performance, i.e., the void filling scheduling algorithms have no significant

advantage over non-void filling scheduling algorithms. The reason is that, when the overhead is very small, the initial offset is small and voids generated on the data channels are most likely too small to accommodate any subsequent data bursts.

When the overhead is large, as shown in Figure 3, the immediate reservation has worst performance. This is due to the fact that much more bandwidth is wasted during the offset interval. According to the figure, void filling scheduling algorithms have better performance than non-void filling scheduling algorithms. The reason is that, when the overhead is large, the offset is much larger and voids generated on data channels may be larger. Therefore, a void filling scheduling algorithm may be more likely to find more voids to accommodate new incoming data bursts, thus improving the network performance. For the non-void filling algorithms, LAUC has slightly lower burst loss probability than FF since LAUC assigns the new incoming data burst to the latest available channel and the bandwidth efficiency is higher than FF. For the void filling algorithms, FFVF and LAUC-VF have very similar performance. This indicates both algorithms effectively utilize the voids generated in the data channels; however, the way to utilize the void is insignificant.

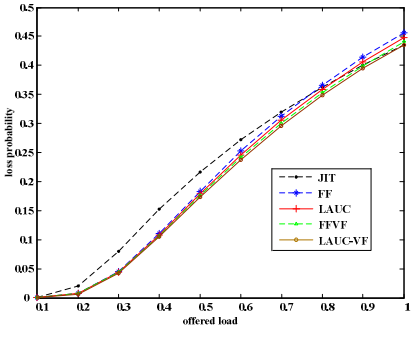


Fig. 2. Performance comparisons under different OBS protocols in NSFNET when the overhead is 0.1 times of the unit.

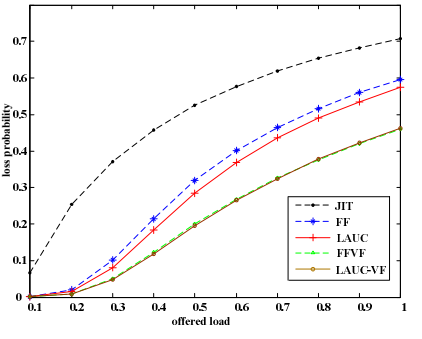


Fig. 3. Performance comparison under different OBS protocols in NSFNET when the overhead is 1 time unit.

In all, when the overhead is small, the deferred reservation scheme has slightly performance improvement than the immediate reservation in terms of loss probability, and the performance deference between void filling algorithms and non-void filling algorithms can be negligible; when the overhead is large, all deferred reservation schemes have better performance than the immediate reservation, and the void filling algorithms have the best performance in terms of loss probability.

3 Impacts of the Switch Configuration Time

3.1 Switch Configuration Time

In the original OBS, the difference between the per node control latency and the control packet processing time T_{cp} is neglected, which is equivalent to either

neglecting T_{sw} or treating T_{sw} as part of T_{cp}. However, fast optical switches which have switching time in nanoseconds or even picoseconds range are only available in small sizes at present, such as the switches based on electro-optical effects. Large optical switches (over 100 ports) normally use technologies such as micro-electro-mechanical system (MEMS) that requires reconfiguration time in milliseconds. Owing to the rapid increase of optical fiber transmission rate up to hundreds of gigabits per second and the necessity to keep a reasonable data burst size, the overhead due to optical switch configuration is no longer negligible. Treating the switch configuration time as a part of the control packet processing delay is also inappropriate.

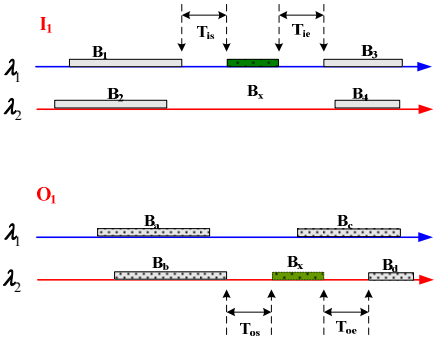


Fig. 4. Illustration of the time constraint due to the switch configuration time when a data burst B_x needs to be switched from input port I_1 in wavelength λ_1 to output port O_1 in wavelength λ_1.

If switching configuration time T_{sw} is negligible, all channels with idle period no less than the duration of the incoming burst are eligible to accommodate it. If T_{sw} is not negligible, the status of input/output port and input/output channels need to be taken into account with combination. Generally, a time gap no less than T_{sw} is required between data bursts. One example of the impact of T_{sw} on the burst loss is illustrated in Figure 4. At the input port I_1, data burst B_1 and B_3 has been scheduled on λ_1, and B_2 and B_4 has been scheduled on λ_2; at the output port O_1, B_a and B_c has been scheduled on λ_1, and B_b and B_d has been scheduled on λ_2. We assume a control packet has been processed at the node; after the offset, its data burst B_x will arrive at the node from the input port I_1 in λ_1. T_{is} and T_{ie} will be the time gap between B_x and B_1, B_x and B_3, respectively. The desired output of B_x is assumed to be O_1. If B_x can be successfully scheduled at the output port, T_{os} and T_{oe} will be the time gap between B_x and B_b, B_x and B_d, respectively. Considering the time consumed to configure the switch elements, scheduling data bursts from an input port to an output port follows two criteria:

S1: The scheduling is impossible if data burst B_b (B_d) has been scheduled from an input port and a wavelength channel other than I_1 and λ_2 and time gap $T_{os}\ (T_{oe}) < T_{sw}$;

S2: The scheduling is impossible if data burst B_1 (B_3) has been scheduled from an input port and a wavelength channel other than O_1 and λ_1 and time gap $T_{is}\ (T_{ie}) < T_{sw}$;

Criteria *S1* states the output channel switching constraint: for a wavelength channel at an output port, a node needs at least T_{sw} to set up a new connection from a wavelength channel of an input port different from the one currently connected. Similarly, criteria *S2* states the input channel switching constraint.

Most existing research on OBS assume that each node has full-range wavelength conversion capability and an internally non-blocking optical switch. Figure 5 and 6 show two typical core node architectures in OBS networks and the first one is usually treated as the default architecture. In the original OBS networks, they are functionally equivalent. However, we will show that the second architecture can be used to reduce burst loss due to large switching configuration time by analyses and simulations.

3.2 Switch Constraint Resolution

If the architecture is deployed as shown in Figure 5, the optical signal is switched first, and then converted to a suitable wavelength if needed. The output port switching constraint, criteria *S1*, is applied accordingly. The data bursts from different channels at different input ports intending for the same channel λ_i of output O_j must use the same wavelength converter dedicated to λ_i. Since the wavelength converter is connected to only one output port of the switch, the existing connection using this wavelength converter must be torn down first before a new connection can be set up. Thus the time gap between the data bursts at the same output channel must be at least T_{sw}.

A simple way to resolve output channel switching constraint is to change all *k-to-1* wavelength converters (WC) in Figure 5 to *k-to-k* wavelength converters so that no dedicated wavelength converter and switch output port are required for an output channel. However, this becomes expensive if *k* is large. Another method is to use the architecture in Figure 6 instead. Assume that the burst B_b in Figure 4 is from input port I_2 and transmitted in λ_2 at output port O_1, burst B_x also request λ_2 at output port O_1 from input port I_1. By deploying the architecture as shown in Figure 6, we can convert the signal in λ_2 first and establish an alternative internal path in the switch. This process is applicable since the control unit has the status of the input/output channels and wavelength converters. Thus, in the time flow of the desired output channel, B_b and B_x can be placed as close as the time for wavelength conversion.

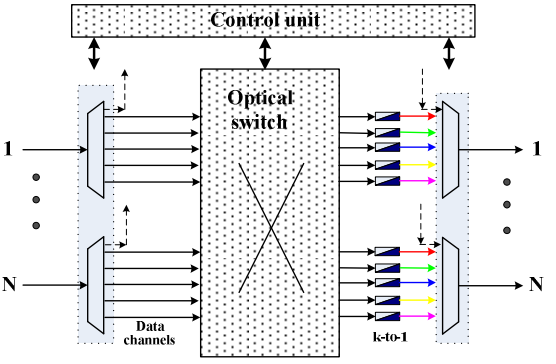


Fig. 5. A typical core node architecture. It mainly contains a control unit, an internally non-blocking optical switch, k-to-1 wavelength converters.

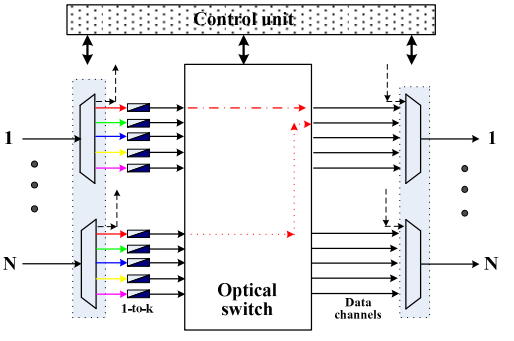


Fig. 6. An alternate architecture of a core node. The 1-to-k wavelength converters are connected to input ports.

For input port switch constraint, there is no simple way to overcome, as stated in scenario *S*2. A hardware solution will be to double the number of *1-to-k* wavelength converters and the inputs of the switch. At each output port of the de-multiplexer, a 3dB coupler is used to split optical signals to duplicated parts and each of the coupler port ε_1 and ε_2 has a *1-to-k* wavelength converter. For the example in Figure 2.12, we assume that the burst B_1 requests the output port O_2, B_x requests O_1, and $T_{is} < T_{sw}$. Since the node has prior knowledge of the switch status, it sets up two internal paths P_1 and P_2 in the switch, where P_1 is for the data burst B_1 and P_2 is for B_x. When B_1 arrives at the node, it is sent to the output O_2 by turning on the wavelength converter at ε_1 and through the internal path P_1. At the end of the transmission of B_1, the wavelength converter at ε_1 is turned off and that at ε_2 is turned on so that data burst B_x is now sent to O_1 through the internal path P_2. In principle, the time gaps between data bursts in an input port can be close to the response time of the wavelength converters.

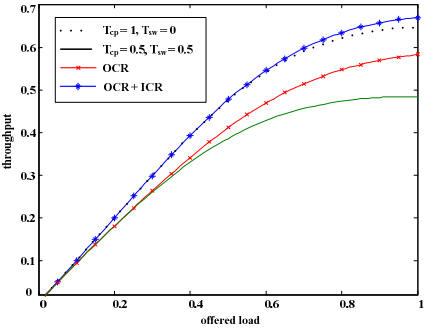


Fig. 7. The throughput with switching constraint resolutions in the topology of NSFNET with $\left[T_{cp}=0.5, T_{sw}=0.5\right]$. The dot line indicates the original OBS with $\left[T_{cp}=1, T_{sw}=0\right]$.

Figure 7 compares the throughput under different constraint resolutions in the topology of NSFNET. The dot line indicates the original OBS in which the control packet processing time is 1 time unit and the switch configuration time is assumed to be negligible. The solid line shows the throughput performance when the impact of switch configuration time ($T_{sw}=0.5$) is considered. We observe that neglecting the switch configuration time over-estimates the network performance. The time needed to configure the switch elements degrades the network performance greatly. The curve with crosses indicates OBS with output port switching constraint resolution (OCR), and the curve with asterisks shows OBS with both input port switching constraint resolution (ICR) and output port switching constraint resolution (OCR). According to the figure, with the output port switching constraint resolution, the negative impact of switch configuration time can be alleviated; with both the output port and input port switching constraint resolutions, the network performance can be further improved.

4 Conclusions

In this paper, the electronic processing time, one of the most important input network parameters in OBS networks, is identified firstly. We further study the impacts of the control packet processing time on the network performance by simulations. Then, we present the impacts of switch configuration time on the network performance. Switching constraint resolutions are proposed to alleviate the negative impacts of the switch configuration time on the network performance.

Acknowledgements. This research is supported in part by Independent Innovation Foundation of Shandong University (Project No. 2009TS105) and WEIHAI Science and Technology Development Plan (Project No. 2007063). The authors thank help from Prof. Victor Li, The University of Hong Kong, Dr. C.Y. Li and Prof. P.K.A. Wai, The Hong Kong Polytechnic University.

References

[1] Baldine, I., Rouskas, G.N., Perros, H.G., Stevenson, D.: JumpStart: a just-in-time signaling architecture for WDM burst-switched networks. EEE Communications Magazine 40(2), 82–89 (2002)

[2] Sun, Y., Hashiguchi, T., Vu, Q.M., Wang, X., Morikawa, H., Aoyama, T.: Design and implementation of an optical burst-switched network testbed. IEEE Communications Magazine 43(11), S48–S55 (2005)

[3] Kitayama, K., Koga, M., Morikawa, H., Hara, S., Kawai, M.: Optical burst switching network testbed in Japan. In: Proc. OFC/NFOE 2005, p. 3 (2005)

[4] Wei, J.Y., McFarland Jr., R.I.: Just-in-time signaling for WDM optical burst switching networks. Journal of Lightwave Technology 18(12), 2019–2037 (2000)

[5] Zaim, A., Baldine, I., Cassada, M., Rouskas, G.: Formal description of jumpstart (Just-In-Time) signaling protocol using EFSM. In: Proc. SPIE Optical Networking and Communications Conference 2002, pp. 160–173 (2002)

[6] Xiong, Y., Vandenhoute, M., Cankaya, H.C.: Control architecture in optical burst-switched WDM networks. IEEE Journal on Selected Areas in Communications 18(10), 1838–1851 (2000)

[7] Qiao, C., Yoo, M.: Optical burst switching (OBS) - a new paradigm for an optical Internet. Journal of High Speed Networks 8(1), 69–84 (1999)

[8] Xu, J., Qiao, C., Li, J., Xu, G.: Efficient channel scheduling algorithms in optical burst switched networks. In: Proc. IEEE INFOCOM 2003, pp. 2268–2278 (2003)

[9] Chen, Y., Wu, H., Xu, D., Qiao, C.: Performance analysis of optical burst switched node with deflection routing. In: Proceedings of IEEE ICC 2003, 2nd edn., pp. 1355–1359 (2003)

[10] Tan, S., Mohan, G., Chua, K.: Burst rescheduling with wavelength and last-hop FDL reassignment in WDM optical burst switching networks. In: Proceedings of IEEE ICC 2003, 2nd edn., pp. 1448–1452 (2003)

[11] Yoo, M., Qiao, C., Dixit, S.: QoS performance of optical burst switching in IP-over-WDM networks. IEEE Journal on Selected Arears in Communications 18(10), 2062–2071 (2000)

[12] Li, G.M., Li, C.Y., Li, V.O.K., Wai, P.K.A.: Novel reservation schemes in optical burst switching networks with fiber delay lines. In: Proceedings of CHINACOM 2008, pp. 454–458 (2008)

Study on Stayed-Cable Health Monitoring

Z.K. Sun, G.M. Li, and S.H. Geng

School of Information Engineering, SHANDONG University, China
gmli@sdu.edu.cn

Abstract. Structure health monitoring of cable-stayed bridges has been developing for decades but there are still no practical monitoring schemes for cracks in steel wires of stayed cables. Stress analysis based on finite-element analysis and fracture mechanics tries to find the relationship between cracks and redistribution of forces among cables and steel wires. Cable force monitoring systems based on magnetic flux sensor, FBG technology have been widely used in present. But to monitor cracks in cables, the existing force monitoring system is not sufficient, cable force monitor, wire force monitor and data processing center are needed.

Keywords: cable health monitoring, cable-stayed bridge, cracked wires.

1 Introduction

The cable-stayed bridge has been adopted in medium or long span bridges for decades. The stayed cable is a strand made of steel wires covered with protective layer and plays the important role of transmitting the kinds of loads from the beam to the tower, and finally to the foundation. However, at the same time the cables are the most vulnerable structural components to be damaged or ruined, because they usually work both in formidable environment and under fatigue loads. When the damage of the steel wires gets to a certain degree, the crack of the wire would happen. The successive danger due to the wire crack is that when one of the wires break up, the force in the wires will redistribute to the remaining wires. The force redistribution will speed up the damage procedure.

Engineers have been working on health monitoring for decades. Most of the cable monitoring systems are based on electromagnetisms and optical sensors which will be introduced in the following part. But there are still no practical systems to monitor cracks in steel wires effectively.

The relationship between cracks and forces in the cables implies that crack monitoring may work on the basis of force monitoring. The force monitoring system generates signal with sensors while the data processing system plays the role of signal processing and pattern recognition to decide whether there are cracks in a cable.

The remainder of this article is organized as follows. In Section 2 the cable system and its defects are summarized. In Section 3 the relationship between crack and the redistribution of tension is discussed. In Section 4 the existing wire crack detection

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1091–1098.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

ways and existing strain monitoring methods or systems are introduced. In Section 5 a monitoring method is put forward finally.

2 Summary on Cable Defects

Damages in cables start with the failures of its protective sheath which are caused by environmental effects and accidental hits and progress to corrosions by oxidizing reaction, and finally stresses in cables make the corrosive part of wires cracked.

A comprehensive literature review shows that the degree of defects in cables depends on the degree of protective sheath damages. Generally speaking, the higher part of the strand is always less damaged than the lower part, and the severely corrosive parts of strand are always located either near or lower than the damaged position of the protective sheath. The reason is that the corrosion of the cable wires will be induced once the protective function of the strand sheath fails. And the rainwater will be stored at the lower part of fissures on protective sheath which speeds up the electrochemical reaction. However, when the protective sheath works well, there is seldom strand damage, cracks or corrosion. In most cases, cracks are caused by the electrochemical reaction.

The vicious circle of the damage development between the protective sheath and the steel wires in it can be described as follows: Corrosion of the steel wires results in cracks on them, the cracked wires may damage the protective sheath in return, and this will make more wires of strand exposed directly to the environment. That is to say, the lesions of steel strands begin as corrosion and develop to cracks because of both electrochemical reaction and tension in wires.

The cracked wires will lose the ability of carrying tension and quit itself of its job. The force will be redistributed among all the cables and the rest of steel wires. Part III will discuss the redistribution of tension.

3 Cable Force Analysis Based on FEA and Fracture Mechanics

In a bridge, the real force in each cable may not be the same as the designed value because deviation usually occurs in manufacturing process and the installation work. For the same reason, the force in each wire or strand that made up of a cable may also not be the same. Under the corrosive environment, the wires holding more tension are apt to be damaged and even crack. While the force redistribution will result in or accelerate the damage of the rest of wires, strands or cables. The cable force analysis based on fracture mechanics theory and the finite-element analysis theory shows how the force is redistributed among the rest of working wires and cables so as to analyze the feasibility of monitoring cracks in wires of the cable. And the analysis may also offer a reference to help recognize and measure the damage or the crack in a cable.

3.1 Cable Force Analysis Based on FEA

The strains analysis based on finite-element analysis is to analyze the redistribution of strains after crack occurring using FEA software. Redistribution of strains takes place both among cables outside of the strand and among wires inside the strand.

In [3], Bai took a bridge on Wan-Yi highways as an example to analyze the redistribution of strains among cables after crack taking place. His model shows that strains in the cables nearer to the cable containing cracked wires change greater. The strains in the farthest cables almost stay what they were. The cable forces reduce gradually from the cable in which there is cracked wires to the one near the tower. And the cable forces increase gradually from the cable in which there is cracked wires to the one located at the mid-span of the cable-stayed bridge. Meanwhile the cable force of the damaged cable with cracked wires decrease and that cable will fracture and quit carrying loads when most of its wires crack.

In [12], Liang built a two dimensional finite element model of Shenzhen Bay Bridge by using the comprehensive commercial software ANSYS.. After the analysis he concluded that, 1) if 10% of the wires of a cable is damaged and quits working, the cable force of this cable decrease obviously; the forces in the cables located at the same side of the tower with that one with cracked wires increase, while the forces in the cables located at the other side of the tower decrease, but all these cable forces vary little; the total of the cable forces decrease too; 2) when the longest cable is completely damaged and quit working, its cable force will change larger than any other cases.

3.2 Wire Force Analysis Based on Fracture Mechanics

The analysis based on FEA shows redistribution of forces among the cables when wires or cables happen to break. To get the information of the redistribution of forces among wires in one damaged cable, analysis based on fracture mechanics should be carried out. In [11], Fan carried out the analysis of the influence factors which result in the damage and crack of the cables in cable-stayed bridges, based on the fracture mechanics and took Tongzilin Bridge over Yalong River as an example.. His conclusions are: 1) if the maximum value of stress intensity factor is less than the fracture strength, the crack in steel wire will not take place; 2) the forces in the rest wires will increase due to force redistribution. He also pointed out that the most important reason of the break of cables is the corrosion of steel wires while the direct reason of corrosion is the damages of shield.

Force analysis based on FEA and fracture mechanics is carried out to reveal the features which can be extracted to classify the working pattern of the cables. Part IV will introduce several existing force monitoring systems.

4 Cable Health Monitoring

Over latest several decades, though how to detect damages especially cracked wires in cables has been studying with the popularity of cable-stayed bridge, there is still no cracked wire monitoring system which can be employed to meet the demands of long-time and real-time monitoring. Cracked wires detecting systems have been used for a long time though they are not online systems. Recent years there have been many project cases employing force monitoring systems which meet the demand of long-time monitoring and are able to get real-time information. Cracked wires detecting methods and strain monitoring systems will be introduced in the following passages.

4.1 Crack Detecting Methods

Visual inspection. Engineers often check the cables with magnifying glasses or telescopes to find out fissures in the shield. Serious corrosion or crack can be observed visually with this method.

Windlass. This method takes use of windlass equipped with devices which can be used to detect both minor defect and reduces of cross sectional area of metal. The equipped windlass detects cracks and records field data moving along the cable.

Robot. Using robot to detect fissures is to collect field data with special designed robot equipped with devices similar to those equipped on windlass method. Robot also needs to move along the cable, but it is self-controlled.

The above detecting methods, especially the last two methods, are able to detect cracked wires and locate positions of crack. But real-time information cannot be obtained, to collect real-time information the best way is to employ sensors designed to detect cracks. Distributed sensors like fiber sensors cannot be employed because the contact surface is limited. Cracks are randomly distributed, direct detection of cracks is inadvisable. Also, to monitor cracks is not to locate their position, it is to find out the existence of cracks in the cable regardless positions. So other measurable features such as strain distribution in cables can be tested to judge the existence.

In part III, strain analysis lead us a crack monitor method that is to monitor the strain in cables and classify health condition accordingly. Some existing strain monitoring systems will be introduced here below.

4.2 Strain Monitor System Based on Magnetic Flux Sensor

MFS (magnetic flux sensor) works on the basis of the theory that permeability of ferromagnetic material would be changed according to the deformation of the material caused by stresses. That is to say, to monitor strains in the cable is to monitor the permeability transformation of the cable. Before MFS being used, permeability under different strains and temperatures should be tested to find out the relationship between permeability and strains under different temperatures. As soon as the relation is clear, MFS can be installed on the selected position of the cable system.

MFS systems have been used in Zhanjiang Bay Bridge and Yibin Yangtze River Bridge [6]. The monitor system in Yibin Yangtze River Bridge can be divided into two parts: the data collection subsystem and the monitor terminal. Eight MF sensors are connected to a switch box; boxes are linked with magneto elastic instrument which can send collected datum through GPRS to the monitor terminal where the datum was processed to predict the danger level according to the value of force.

4.3 Strain Monitor Systems Based on FBG

FBG sensors collect information by modulating the wavelength of fiber Bragg grating. While the wavelength of fiber Bragg grating depends on the period of grating n_{eff} and the effective index of the reverse coupling model Λ, as shown in equation (1).

$$\lambda_B = 2n_{eff}\Lambda \tag{1}$$

Any physical process that changes the two parameters may cause shift of FBG's wavelength. Of all physical processes that may cause the shift, the most observable two are stress and strain. Another great influencing factor is temperature variation. The sensors based on FBG also provide the capability of distributed detecting, and therefore, it has the ability of providing strain and stress information of the surface where fibers are paved.

The application of FBG in cable-stayed bridges includes cable force monitoring and wire force monitoring. The application of stress monitoring systems based on FBG is introduced below here.

1) The application of force monitoring system based on the FBG vibration sensors in the Second Yangtze River Bridge of Wuhan [9]:The system applied in the Second Yangtze River Bridge of Wuhan contains FBG vibration sensors, photoelectric signal processor, data acquisition units, computers and real-time inspecting software. Sensors are specially designed for the circumstances in which the bridge works (as shown in Figure 1). And the software has the ability to recognize the characteristic frequency of the cable, then calculate the force by equation (2).

$$T = \frac{4ml^2 f_n^{\,2}}{n^2} \tag{2}$$

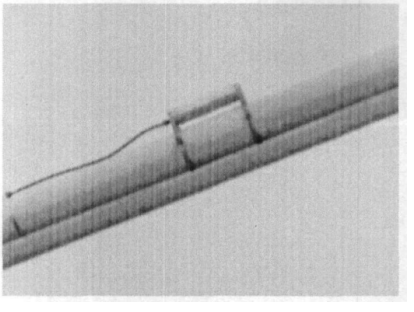


Fig. 1. FBG vibration sensor on a stayed cable (from [9])

Equation (2) indicates the relation between the tensions of a string and the characteristic frequency of the string. In the equation, l and m are the length and the linear density of the string, respectively; f_n is the nth order characteristic frequency.

The application of FBG in health monitor of Bengbu Huai River Bridge [18]: Fiber reinforced plastics (FRP) have found widespread usage in structures as the replacement of steel wires because of the following features of FRP: high strength to weight ratio, resistance to corrosion and ease of handling. Repair of disabled cables involves of the replacement of strands which are traditionally made of steel wires. The Bengbu Huai River Bridge is an old bridge built in 1989, the health monitoring system should be built to monitor the force in the cables.

Engineers employed FBG combined with FRP to produce "intelligent cable" with the aim of monitoring the cable force.

To reduce the influence of temperature, temperature sensors are needed. The structure of the intelligent cable is shown in Figure 2.

Fig. 2. The structure of the intelligent cable(from [18])

The blue solid dots in the graphic represent FRPs wires with FBG force sensors, the intelligent wires, the red solid dot represents the FRP with FBG temperature sensors, and the empty circles are steel wires, the big circle outside is the protective layer, the PE sheath.

The above passages introduced several existing methods for cable force monitoring. The MFS system and FBG vibration sensor system can be adopted to monitor forces of cables, and the intelligent cable system can be adopted to monitor force distributed in both cables and wires. The advantage of MFS and the FBG vibration sensor system lies in that they can be employed in the bridges no matter how long the bridges have been put into use, while the intelligent cable suffers the disadvantage that it has to be installed during the construction period or the renewal period when the damaged cable is replaced.

5 The Advisable Monitoring System for Cracks in Cables

There have been no thorough systems for monitoring the cracks of steel wires in cables of cable-stayed bridges by now. An advisable monitoring system is put forward here to monitor cracks in wires or cables with two subsystems and a processing center. One subsystem is to monitor the protective layer of a cable and the other to monitor the cable and wire forces, the processing center is to recognize cracks automatically based on force distribution.

Pattern recognition technology will be employed to process force datum. One problem is that there are no records about the crack pattern and the healthy pattern which are the training samples for classifier except information from models and simulations. To solve this problem, the monitoring system is obliged to collect training samples and predict or judge cracks at the same time. That is why the shield health monitoring system is necessary.

As mentioned before, the shield is the protective layer of steel strand and the surface of cable. The damages of a cable start with the damages of the shield. Comprehensive investigation shows that in most crack cases, crack is caused by electrochemical reactions. Corrosions of steel strand are the presentation of electrochemical reactions which are seldom reported in the unspoiled shield. Therefore the integrality of the protective layer means a circumstance against the electrochemical reaction. Once fissures in the shield are reported, crack test is necessary to ensure the safety of the

bridge. Distributed sensors will meet the demand of detecting the random occurrence of fissures. Distributed sensors based on FBG have found widespread usage in surface crack monitoring.

Cable force monitoring subsystem includes two parts: one is to monitor the force of cable and the other is to monitor the force distributed in wires. The wire force monitor will take the advantage of intelligent cable, and the cable force monitor will take the advantage of FBG vibration sensors to make the signal processing easier.

To found a pattern database, both field survey and theoretical analysis on structural mechanics are necessary. Theoretical analysis includes model test and theoretical calculation, model test is employed to obtain force information in damaged cables, while theoretical calculation is employed to avoid overload of forces. Field survey employs the shield monitoring subsystem and force monitoring subsystem to found a database considering real-time information about environment load, traffic load, cable forces and wire forces which will be compared and added to database founded according to simulations and model test. Field datum for the healthy pattern is collected before the shield monitoring system detect fissures, and field datum for the crack pattern is collected when the crack detecting systems such as crack detecting robots or windlass detect cracks of steel wires.

With the database consisting of training samples and test samples, the processing center can employ pattern recognition technology to classify wire cracks in cables.

6 Conclusions

This paper is a survey on cracked wires monitoring system. There have been no practical cracked wires monitoring system so far. Based on the literature review work, several stayed cable health monitoring systems based on cable force monitoring is introduced first. And an advisable cracked wire monitoring system is put forward finally based on the analysis of the existing monitoring systems.

Acknowledgements. This work is supported by Projects of Science and Technology Development Plan of SHANDONG province (No.2011GSF12007). This research is supported in part by Independent Innovation Foundation of Shandong University (Project No. 2009TS105) and WEIHAI Science and Technology Development Plan (Project No. 2007063).

References

[1] Tang, S.-Y., Huang, D.-G., Zhao, J.-W.: Common defects causes of concrete cable-stayed brdiges. Western China Communications Science & Technology 4, 60-61+68 (2008) (in Chinese)

[2] Zhou, Y., Song, J.-C.: The damage analysis of the cable-stayed bridge. Shanxi Architecture 1, 340–341 (2009) (in Chinese)

[3] Bai, X.-R.: Hazards and cause analysis of cable-stayed bridge. Chongqing Jiaotong University (2008) (in Chinese)

[4] Chang, B.-B.: Research on damage mechanism and prevention measures of inclined cable in cable-stayed bridge. Chongqing Jiaotong University (2008) (in Chinese)

[5] Yu, Z.-Y.: Study on the use of robots to check the cables. Guang Dong Gong Lu Jiaotong 4, 34–36 (2009) (in Chinese)
[6] He, L., Deng, N.-C.: Application of magnetic flux sensor in stay cable monitoring of Yin Bin Yangtze River Bridge. Technology of Highway and Transport 1, 95–97 (2010) (in Chinese)
[7] Zhao, Y., Ansari, F.: Embedded fiber optic sensor for characterization of interface strains in FRP composite. Sensors and Actuators 100(2-3), 247–251 (2002)
[8] Li, M., Liao, Y.-B.: Fiber optic sensor. Wuhan University Press (2008) (in Chinese)
[9] Zhang, D.-S., Li, W., Guo, D., Hu, J., Luo, P., Jiang, D.-S.: Real-time monitor system of bridge cable force based on FBG vibration sensor and its application. Chinese Journal of Sensors and Actuators 12, 2720–2723 (2007) (in Chinese)
[10] Wang, D.-H.: Insitu monitoring theory and method for stayed cables dynamic response based on piezoelectric sensors. Piezoelectrics & Acoustooptics 1, 80–84 (2001) (in Chinese)
[11] Fan, H., Ke, Y., Tang, T.: A study on cable breakage and fracture of long span cable-stayed bridge. Highway 4, 107–111 (2010) (in Chinese)
[12] Liang, Z., Li, N., Zheng, C., Liu, X.-D.: Sensitive analysis of cable damage based on monitoring deflection and cable tension. Highway 3, 39–43 (2009) (in Chinese)
[13] Li, S.: Study and application about cable force testing method based on fiber optical grating sensing principle. Wuhan University of Technology (2009) (in Chinese)
[14] Zhang, X.-S., Du, Y., Sun, B.: Study in monitoring of cables' tension with fiber brag sensor. China Safety Science Journal 13(7), 77–80 (2003) (in Chinese)
[15] http://wiki.zhulong.com/baike/detail.asp?id=230532 (in Chinese)
[16] Hou, L.-Q., Ou, J.-P., Zhao, X.-F., Li, H.-W.: Identification of modal parameters and cable tension of Harbin Sifangtai Bridge. Journal of Vibration and Shock 5, 106–110+207–208 (2009) (in Chinese)
[17] Tang, Y.-M., Zhang, H.: Large-span Bridge Cable Damage and Health Monitoring. Bridge Construction 5, 79–82 (2002) (in Chinese)
[18] Lv, E.-R.: The Research of Engineering Application of Fiber Bragg Grating Technique in Health Surveillance of Beng Bu Huai He Bridge. Jilin University (2009)

Research on the Information Library for Component Testing

Ying Jiang[1,2,3,*], Lin Jin[4], Xiao-Dong Fu[1,2], and Yu-Ping Liu[1]

[1] Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China
[2] Yunnan Key Lab of Computer Technology Application, Kunming 650500, China
[3] Key Laboratory of High Confidence Software Technologies (Peking University), Ministry of Education, Beijing 100871, China
jy_910@163.com
[4] Advanced Digital Technology Co. Ltd, Beijing 100089, China

Abstract. The major objective of Component Based Software Development (CBSD) is to reduce the cost and time of software development by reusing available components. Testing is an important phase in CBSD. During the process of component testing reuse, the storage and management of test information play a key role. As a result, it is important to research the issues about the information library for component testing. This paper presents an infrastructure of component test information library. Firstly, the architecture of the library is established. Then the component model for test information description is extended based on BIDM. Towards the main functions of test information library, the classification schemes, retrieval strategy, verification and evaluation method are proposed. As a support for component developers, reusers and the third party, the information library for component testing and its functions will be helpful for component testing and reuse.

Keywords: CBSD, COMPONENT TEST INFORMATION LIBRARY, DESCRIPTION MODEL, CLASSIFICATION SCHEME, VERIFICATION AND EVALUATION METHOD.

1 Introduction

In order to improve the productivity and quality of software, component-based software development (CBSD) is a feasible way [1, 2]. Software systems are assembled from black-box and white-box components using the method of CBSD. The main activities during the process of CBSD include component production, component management and component composition. The aim of component management is to store and retrieve reusable software components, which can help to develop the software systems composed of components for the developers, reusers and third party. Component library is the infrastructure of component management.

* Corresponding author.

Z. Du (Ed.): Intelligence Computation and Evolutionary Computation, AISC 180, pp. 1099–1108.
springerlink.com © Springer-Verlag Berlin Heidelberg 2013

To ensure the quality of new software systems, the quality of components should be ensured firstly. Due to the differences between components and traditional software, new testing methods are needed to make the components reusable and reliable [3]. Test information is the information that produced or used during the testing process, including the test plan, test data suites, test adequacy criteria, bug report and etc.

At present, the reuse is concentrated in component's code or function. However, the reuse of test information is an important part of component reuse. Usually, the developer of a component should test it in different application domains before publishing component. Then the reuser will re-test component again before component is reused in the new application. Some testing work executed by reuser is as the same as that of developer, which will cost plenty of extra test resources. In fact, developer's test information is useful for the reuser to test component. Thus, the test information produced in the development process can also be reused during the reuse process. When the reuser reuses the test information of the developer, the cost of testing and reuse can be reduced.

Gradually, some users pay attention to the issue of component test information. Even though the developer provides component test information, the type and format of test information may be various. How to query and retrieve test information effectively has became the key problem while reusing test information. Moreover, the validity of test information may not be ensured easily, which will enhance the difficulty of quality evaluation of component for the reuser. For the purpose of storing and retrieving test information, it is necessary to establish the information library for component testing.

In this paper, we mainly address the series problem of information library for component testing. We propose an infrastructure of information library for component testing, which can support test information's description, classification, maintenance, verification, retrieval, feedback and evaluation. The rest of this paper is organized as follows. Section 2 discusses research work related to ours. Section 3 describes the architecture of the information library for component testing. Section 4 proposes an extended component model for test information description. Section 5 presents the description, classification and retrieval of component test information. The verification and evaluation method for test information are proposed in Section 6. Finally, we conclude this paper and present our future work in Section 7.

2 Related Work

Software testing is one of the important techniques that assure the software quality. Recently, more and more users begin to notice the necessaries of test information management. Many companies and researchers rely on the test management tool to manage software test information. Using these tools, the focus of management aims at every action during the whole test cycle, including the test requirement, test plan, test execution and so on. TestDirector and ClearQuest are the typical test management tools. These product tools can support the traditional software testing process. In order to meet the special requirement of different domain, some researchers also

propose specific test information management, which usually manage test data or results [4, 5].

During software testing, plenty of test information will produced, such as information about test user, test environment, test data and test evaluation. Generally, many tools or researches manage test information from the view of software testers. The test information for non-testers can not be managed in these tools.

Comparing with the testing of traditional software and object-oriented software, some characteristics of components, such as heterogeneity, likely unavailability of source code and documentation of some components, and evolution, challenge the testing of component-based software. Because of the above difference between component and software module, component testing usually will produce more information than software testing.

Test reuse is an important part in the process of component reuse. Whereas, the current software test information management can not support test reuse effectively. The requirement of test information management and reuse from the reuser or third party do not been considered.

Our information library for component testing aims at solving some problems about test information management and reuse. In order to facilitate information management of developers, reusers and the third party, an infrastructure of the information library for component testing is proposed. Furthermore, the main functions of the information library for component testing are defined in detail.

2.1 The Architecture of the Information Library for Component Testing

Component management tool is named as component library management system, including the component description, classification, store, maintenance, retrieval and etc. Be similar to component management, component testing management should provide some functions for different users.

The management of test information is a useful assistant during component testing. It need systematic solutions and relative tools support testing engineers and quality assurance engineers to create, update, and maintain diverse test information, including test cases, test scripts, test data, test results, and discovered problems [6].

In order to support component testing and reuse, we design an information library for component testing, whose architecture is shown in Figure 1.

There are seven main functions in the architecture of the information library for component testing.

(1) Test Information Description. Test information provided by the developer or reuser will be described following the extended component model in Section 4.
(2) Test Information Classification. Test information provided by the developer or reuser will be classified and stored in component test information library according to the description and classification method defined in Section 5.
(3) Test Information Maintenance. Test information should be maintained to accord with the management rules by the different users and the administrator.
(4) Test Information Verification. The administrator verifies test information using the approach in Section 6.

(5) Test Information Retrieval. The different users can retrieve the test information that meets their requirement.
(6) Test Information Feedback. The component reusers provide the feedback information about component test information, which will be useful for the other reusers and developers.
(7) Test Information Evaluation. The administrator evaluates test information according to the attributes and feedback information. The detail method is presented in Section 6.

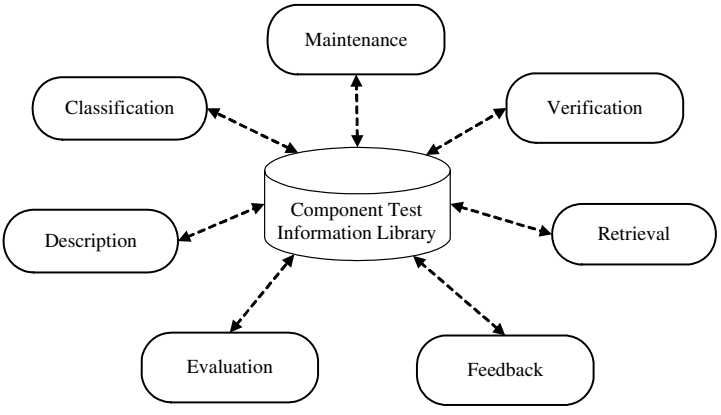


Fig. 1. The architecture of the information library for component testing

2.2 Extended Component Model for Test Information Description

During the process of CBSD, the component model is the basis of component reuse and development. A component model defines the ways to construct components and regulates the ways to integrate and assemble components, which can support component interactions, composition, and assembly. In addition, a component model also defines the mechanisms for component customization, packaging, and deployment [6]. Thus, building the right component model is the first important step while developing and reusing components. From the different view, Mei et al. [7] discuss the four kinds of component models, including the model for macroscopical description of components, the model for component description and classification, the model for component specification and composition, and the model for component implementation.

The model for component description and classification are the data models of the component library, which tries to describe components in a comprehensive form. They need to describe all necessary component information that is helpful for users to look up, understand, select, adapt and use components [7]. The representative description and classification model includes BIDM (Basic Interoperability Data Model) [8] and REBOOT (Reuse Based on Object-Oriented Technology) [9].

BIDM is a data model standard used in the interoperation among software reuse library and defines the basic core structure to describe component and relative

resources, including RIGObject, Library, Organization, Asset and Element. According to BIDM, some basic information of component can be defined, such as component name, publisher, publish date, component description and so on. However, there is not certain information about component testing in BIDM. In order to support the testing reuse and management, test information needed to add in the component model.

Based on the structure of BIDM, we extend the Element in BIDM. As shown in Figure 2, Code, Document and ComponentDescription are derived from Element.

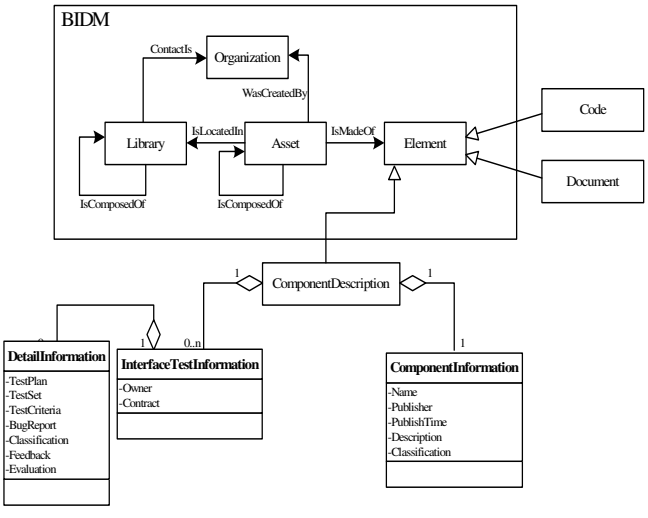


Fig. 2. The extended BIDM with test information

Code is the resource or executable codes of components. Document is the additional document attached with components. ComponentDescription describe different kind of information, including the basic information and test information of components.

Based on the extended BIDM, ComponentDescription is consisted of two part information, including basic information (ComponentInformation) and test information (InterfaceTestInformation、DetailInformation) for each component interface. InterfaceTestInformation include the information of component developer or reuser (Owner) and contract defined by them. DetailInformation include test plan, test set, test criteria, bug report, classification, feedback and evaluation.

Through adding test information in BIDM, the ability of component management can be enhanced from the aspect of component understanding. The reuser can understand component through the attributes and classification. The effective test set of the developer can help the reuser to decrease repetitious testing, which will increase the test efficiency. At the same time, the test information of the reuser shows the quality and reusable degree of component at some degree.

Except for understanding and testing, component test information is useful to retrieve component that meet the requirement.

After extending the model for component description and classification, component test information and process will be integrated with component management process.

3 The Classification and Retrieval of Component Test Information

3.1 The Classification of Component Test Information

To understand component and decrease the reuse cost, it is necessary to define a better classification scheme of component. Component test information is a kind of component attribute information. In fact, there are many similarities between component attribute and testing information. So we use component classification method to classify component test information.

There are four typical classification of component library summarized by Frakes et al. [10], including attribute-value classification scheme, enumerated classification scheme, faceted classification scheme, and free text classification scheme. In attribute-value classification scheme, a collection is described in terms of a set of attributes and their values. And a subject area is broken into mutually exclusive, usually hierarchical, classes in enumerated classification scheme. Faceted classification schema is the base technology for the domain analysis method proposed by Prieto-Diaz [11]. In a faceted classification scheme, a subject area is analyzed into basic terms that are organized as facets. Objects are then classified by synthesizing the facet term pairs in the classification scheme. The development of facets is usually done by identifying important vocabulary in a domain and then grouping like terms together into facets [10]. In free text keyword indexing [12, 13], terms are automatically extracted from documentation.

We define four kinds of classification scheme in our information library for component testing. The faceted classification scheme is the main classification for test information. The main facets of component test information are test method, test category, test phase and test environment. The detail description tree based on these four facets can be seen in Figure 3.

As can be seen from Figure 3, a facet is consisted of a set of terms, and a component can be depicted by one or more terms.

The facets cannot be unchangeable. The administrator need maintain the facets along with the development of component testing and the requirement of the users.

3.2 The Retrieval of Component Test Information

Based on the above faceted classification scheme, each facet represents a kind of partition for test information in library. So the users can define the retrieval condition from the different views that is also helpful to understand component.

After combing the faceted, attribute-value, enumerated, and free text classification scheme, the users can use many methods to retrieve components based on different classification scheme [10].

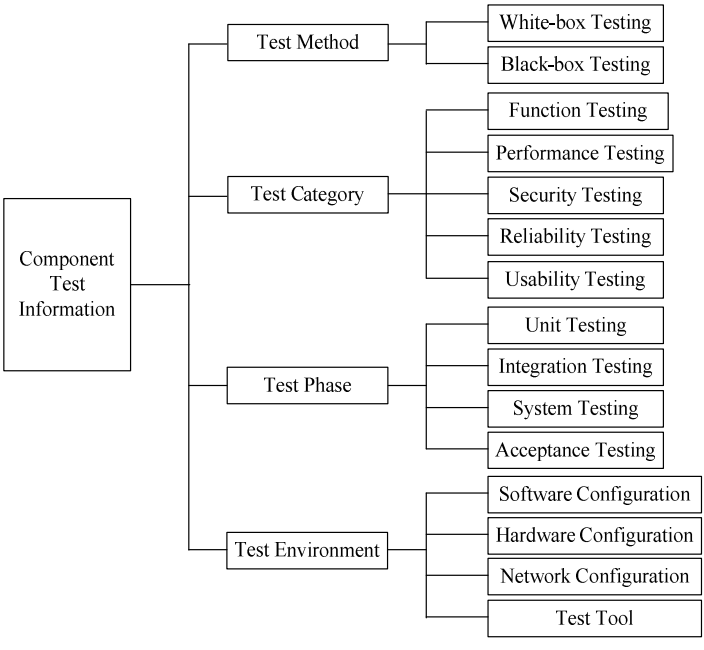


Fig. 3. Component test information description tree based on faceted classification

4 The Verification and Evaluation Method for Component Test Information

4.1 The Verification for Component Test Information

Components need to be evaluated qualitatively and quantitatively before they are stored in component library. The main purpose of evaluation is to indicate the requirement that component can meet and forecast the income that component will obtain [14]. The process of component verification can insure the quality and normative of components in library. At present, most component libraries do not provide automatic verification mechanism. The administrator of component library manually insures the qualification of submitted component. Sometimes, there is external support system for component verification.

Be similar to component verification, component test information also requires be verified before entering the library. Because component test information can reflect the quality of components, the reuser or third party will estimate components through test information. It is necessary to assure the quality of test information.

In our component test information library, the verification of test information is the combination of artificial and automated method. We have developed a testing platform for components [15], which can support some automated verification of test information. With the help of component testing platform, the administrator verifies test information based on the following steps:

(1) execute the test set according to the test plan;
(2) check the actual test results with relevant test criteria in test information;
(3) review the bug report whether match the actual testing.

After finishing the above process, the administrator will gain the real test results of component. These verification results will be the important factors of test information evaluation.

4.2 The Evaluation for Component Test Information

Table 1. The Weight Value of Evaluation Factors for Component Test Information

Factors		**Sub - Factors**	
Name	**Weight Value**	**Name**	**Weight Value**
Reusability	0.249	Usability Operability Suitability Interoperability Compliance	0.225 0.239 0.162 0.246 0.128
quality of bug report	0.118	Undersatndability Completeness Accuracy	0.321 0.218 0.461
test complexity	0.183	Complexity Learnability Operability Configureability	0.362 0.119 0.271 0.248
test efficiency	0.125	Time Behavior Resource ehavior	0.5 0.5
reuser feedback	0.192		
administrator evaluation	0.133		

After the verification done by the administrator, test information will be stored in the library. The quantitative evaluation for test information can provide help for the different users. During the assessment of component test information, the considered

factors include the reusability, quality of bug report, test complexity, test efficiency, user feedback and administrator evaluation. The users and administrator usually give the ambiguous feedback and evaluation result, such as fine, medium, poor and other. So the users' feedback and the administrator's evaluation results show some degree of ambiguity. Furthermore, some quality factors of component test information also have a certain ambiguity.

While evaluating test information quality, the quantitative result is viewed as objective and effective compared with the qualitative result. In order to present a quantitative evaluation result for test information, the fuzzy theory and Analytical Hierarchy Process method is used to calculate the weight value of all factors and sub-factors. Based on our previous research [16], the content and weight value of evaluation factors for component test information is regulated. The current weight value is listed in Table 1.

Using the weight value in Table 1, the administrator can give a quantitative evaluation value for the specific test information according to the user's feedback and the information obtained during the real testing process.

5 Conclusion and Future Work

The library of component test information and its functions can provide effective test information for component testing, which will facilitate component testing and reuse. The different roles, including the developers, reusers and third party, can use the library during component testing and reuse. Through evaluating the quality of component based on the effective component test information, it is helpful to select and reuse components.

In the future, we will complete the functions of test information library. The definition of component test information will be refined. Moreover, the facets in the classification will be adjusted according to the requirement of component testing. We will also focus on the retrieval method of test information, and improve the validity and efficiency of retrieval.

Acknowledgements. This research is sponsored by the National Science Foundation of China No. 60703116 and 61063006, the Ministry of Education Research Foundation of Key Laboratory of High Confidence Software Technologies No. HCST200903, and the Science Research Foundation of Education Department of Yunnan Province No. 2010Z017.

References

[1] Gaffney, J.E., Durek, T.: Software reuse-key to enhanced productivity: some quantitative models. Information and Software Technology 31(5), 258–267 (1989)

[2] Mili, H., Mili, F., Mili, A.: Reusing software: issues and research directions. IEEE Transactions on Software Engineering 21, 528–561 (1995)

[3] Weyuker, E.: Testing component-based software: A cautionary tale. IEEE Software 15(5), 54–59 (1998)

[4] Ding, Y.-F., Jiang, S.: Software test information management system to support collaborative test. Computer Engineering 31(6), 81–83 (2005)
[5] Liu, C.: Software testing management based on defect tracking system. Journal of Yulin Teachers College 3, 140–144 (2007)
[6] Gao, J.Z., Tsao, H.-S.J., Wu, Y.: Testing and Quality Assurance for Component-Based Software. Artech House, Norwood (2003)
[7] Mei, H., Yang, F.-Q.: Building enterprise reuse program – A model-based approach. Science in China (Series F) 45(3), 161–173 (2002)
[8] Institute of Electrical and Electronics Engineers. IEEE Std 1420.1, IEEE Standard for Information Technology-Software Reuse-Mata Model for Reuse Library Interoperability: Basic Interoperability Data Model, BIDM (1995)
[9] Sindre, G., Conradi, R., Karlsson, E.-A.: The REBOOT approach to software reuse. System Software 30, 201–212 (1995)
[10] Frakes, W.B., Pole, T.: An empirical study of representation methods for reusable software. IEEE Transaction on Software Engineering 20, 617–630 (1994)
[11] Prieto-Diaz, R.: Implementing faceted classification for software reuse. Communications of the ACM 34(5), 89–97 (1991)
[12] Frakes, W., Baeza-Yates, R.: Information Retrieval: Data Structures and Algorithms. Prentice-Hall, Englewood Cliffs (1992)
[13] Frakes, W., Nejmeh, B.: An information system for software reuse. In: 20th Ann. Haw. Int. Conf. Syst. Sci., Kona, HA, USA (1987)
[14] NATO Communications and Information Systems Agency Components: NATO Standard for Management of a Reusable Software Component Library (1991)
[15] Jiang, Y., Hou, S.-S., Shan, J.-H., Zhang, L., Xie, B.: An Approach to Testing Black-box Components Using Contract-Based Mutation. International Journal of Software Engineering and Knowledge Engineering 18(1), 93–117 (2008)
[16] Jiang, Y., Lin, J., Liu, Y.-P.: The Management Framework of Component Test Information. In: 2011 International Conference on Energy and Environmental Science, pp. 3539–3546 (2011)

Author Index

Bi, Xiaoming 153, 173
Bi, Ye 753
Bo, Li 855

Cao, Lijia 13
Cao, Yujun 849
Cen, Kuang 825
Chai, Jiwen 547
Chai, Yunpeng 679
Chen, Chun-Liang 375
Chen, Hua 61
Chen, Jian 535
Chen, Juan 197, 499
Chen, Junliang 797
Chen, Min 861
Chen, Mingrong 25
Chen, Shiqiang 453
Chen, Xi 983
Chen, Xiang-tao 937
Chen, Xueguang 355
Chen, Yunping 753
Cong, Chen 139

Dai, Lu 869
Deng, Fan 805
Ding, Haitao 625
Ding, Lixin 869
Ding, Wen 203, 481
Du, Guotong 255
Du, Jing-yi 657
Du, Wei 797
Du, Zhihui 679
Dutkiewicz, Eryk 945

Fan, Wenjun 679
Fan, Yuezhen 243
Fan, Zuo-e 513
Fang, Gengfa 945
Fang, Li 453
Fang, Pei 905
Fang, Wu 1071
Farshchi, Seyyed Mohammad Reza 1045
Fu, Xiao-Dong 1099

Gao, Zhan-Feng 237
Geng, S.H. 1081, 1091
Geng, ZengJian 575
Gu, Haijun 965
Gu, Hongjiu 475, 493
Gu, Wen-jin 513
Guo, Jianqing 87
Guo, Konghui 625
Guo, Lei 1021
Guo, Xiaofeng 215

Han, Bing-Xin 237
Hao, Aimin 327
He, Wei 369
He, Xiang-yang 937
He, Zhikun 37, 381
He, Zishu 917
Hong, Liu 355
Hou, Chunjie 31
Hu, Guangzhou 423
Hu, Liang 153, 161, 173, 185, 191
Hu, Liming 887
Hu, Mingsheng 355
Hu, Yunan 47
Hua, Shujun 133
Huan, Tian 779
Huang, Chengzhong 161, 185, 191

Huang, Guorui 633
Huang, Manyuan 185

Ji, Hongcheng 805
Jia, Jun 401
Jia, Weihua 845
Jiang, Fachao 243, 249
Jiang, Hongan 753
Jiang, Lingli 825
Jiang, Ying 1099
Jie, Zhang 487, 905
Jin, Lin 1099
Jing, Qiu 467
Jun, Li 439

Kan, Shulin 929
Kang, Hui 791
Kong, Dehui 723

Lai, Wei 407
Lei, Jiping 849
Lei, Shen 93
Li, Chiyuan 1063
Li, Fachun 881, 895
Li, Fule 607
Li, G.M. 1081, 1091
Li, Guohui 733
Li, Haifeng 557
Li, Haixia 761
Li, Haoru 423
Li, Ji 911
Li, Jie 223
Li, Kunlun 7
Li, Qiuying 557
Li, Rui 381
Li, ShuLiang 875
Li, Tianli 67, 395
Li, Wei-feng 657
Li, Xiangping 255
Li, Xiaoda 1001
Li, Xin 723
Li, XinMing 715
Li, Yangqing 945
Li, Yi 715
Li, Zhiyong 841
Lia, Ran 953
Liang, Haozhe 733
Liang, Hu 209, 923
Liang, Shao 1071
Liao, Xue 575
Lin, Chengyan 87
Lin, Nan 215
Lin, Wen 255
Ling, Xia 813
Linmao 785
Liu, Entuo 243, 249
Liu, Fenlin 535
Liu, Guangbin 37
Liu, Jinfeng 1021
Liu, Jinjun 831
Liu, Keding 415
Liu, Re 527
Liu, Shanmei 547
Liu, Ting 369
Liu, Wei 625, 965
Liu, Wenchang 73
Liu, Yanjun 109
Liu, Yanshu 841, 849
Liu, Yehui 461
Liu, Yinan 13
Liu, Yiqin 431
Liu, Yu-Ping 1099
Liu, Zhanyi 369
Liu, Zheng 61
Liu, Zhi 575
Lu, G.Y. 339, 347
Lu, Wenpeng 761
Lui, Xiangyang 1001
Luo, Zhuojun 837

Ma, Changxia 753
Ma, Yunfei 1015
Mao, Guojun 689
Mei, Lin 271
Meng, Lingrui 31

Nie, Zuoling 1
Niu, Wenjie 567

Pan, Chang-peng 513
Pei, Tao 805

Qi, Zhao-Hui 237
Qian, Shao 699
Qin, Ling 929

Ren, Qingchang 665
Rong, Hua 25

Shao, Ming-zhu 657, 1009
Shen, Rensheng 255

Shen, Ying-Ming 237
Shi, Xian-jun 513
Shi, Xingjun 773
Shi, Ziming 791
Song, Jian 959
Su, Yuhai 1
Sun, Bin 305
Sun, Bo 1009
Sun, Peiliang 767
Sun, Shibing 845
Sun, Z.K. 1081, 1091

Tan, Jialei 1015
Tan, Shaosheng 861
Tang, Shoulian 953
Tang, Yong 547
Tao, Jin 125
Tao, Zhang 1029

Wang, Chong-jin 1
Wang, Guodong 557
Wang, Haocheng 281
Wang, Huaiqun 461
Wang, Hui 363
Wang, Jian 53, 271
Wang, Jianping 861
Wang, Jianxin 1015
Wang, Kejia 423
Wang, Kening 297
Wang, Lixin 381
Wang, Lizhe 911
Wang, Molin 243, 249
Wang, Pan 7
Wang, Shuliang 355
Wang, Tong 1015
Wang, Wei 297
Wang, Wentao 327
Wang, XiaoMei 715
Wang, Xin 117
Wang, Xiuling 281
Wang, Xiu-Ping 527
Wang, Yiqiang 327
Wang, Yueping 101
Wang, Yuqin 215
Wang, Yutian 363
Wang, Zhen 25
Wang, Zhiwu 355
Wangyani 313
Wangyong 785
Wei, HaiLong 575
Wei, Wei 767
Wei, Zhang 1039
Wei, Zhenxiang 261
Wu, Binying 453
Wu, Jin 591
Wu, Peng 401
Wu, Tao 983
Wu, Weihe 327
Wu, Yangbo 875

Xi, Ye 615
Xia, Han 1063
Xia, Ming 767
Xiao, Shufeng 819
Xie, Kai 339, 347
Xie, Qiuen 707
Xie, Rui-yun 1009
Xie, Wenliang 79
Xiong, Lei 1035
Xiong, Shunqing 583, 599
Xu, Dianjun 1055
Xu, Jingqi 505
Xu, Jun-Ming 53
Xu, Ke 1063
Xu, Ping 215
Xu, Shukui 733
Xu, Wenping 355
Xu, Xirong 53
Xue, Ruojuan 761

Yaghoobi, Mahdi 1045
Yan, Bin 535
Yan, Hu 147
Yan, Jiang 447
Yan, Yongjun 407
Yan, Yusong 983
Yang, Jian 37
Yang, Li 47
Yang, Maoqiang 231
Yang, Shihan 547
Yang, Shizhong 665
Yang, Wen 161
Yang, Yanxin 117
Yang, Yi 689
Yang, Yihuai 945
Yang, Yingyuna 743
Yang, Yu 975
Yang, Zhigang 647
Yao, Yuan 271
Yao, Zhen-qiang 271

Yao, Zhiqiang 1063
Yi, Can 841
Yi, Jianbing 707
Yi, Li 287
Yi, Zhi'an 101
Yin, Guangli 699
Ying, Jiang 521
Yu, Cao 167, 179
Yu, H.Q. 339, 347
Yu, Qiu 467
Yu, Zhengyong 993
Yu, Zongguang 975

Zeng, Fan-ming 305
Zeng, Liang 743
Zeng, Ying 535
Zhan, Xianghui 1001
Zhang, Chao 13
Zhang, Chunzhi 423
Zhang, GongXuan 715
Zhang, Lei 31
Zhang, Li 607
Zhang, Qin 363
Zhang, Ruihua 591
Zhang, Shengxiu 13
Zhang, Tianle 679
Zhang, Wei-dong 305
Zhang, Xiao 255
Zhang, Xiongxing 297
Zhang, Xueping 117
Zhang, Yang 271
Zhang, Yong 723
Zhang, Yonghua 911
Zhang, Zhanqiang 753
Zhangbo 319
Zhao, Hua 305
Zhao, Xiaohui 965
Zhao, Xijing 37, 381
Zheng, Chunjiao 671
Zheng, Weifa 79
Zhou, Haifeng 389
Zhou, Jianliang 407
Zhou, Lijie 261
Zhou, Shuming 53
Zhou, Weihong 583, 599
Zhou, Yang 743
Zhou, Yunbo 975
Zhu, Tianxiang 401
Zhu, Xuezhen 773
Zhu, Yao-jie 271
Zou, Donglan 875
Zou, Huan 117
Zou, Lin 917