

Lecture Notes in Electrical Engineering

Volume 140

David Jin and Sally Lin (Eds.)

Advances in Electronic Engineering, Communication and Management Vol. 2

Proceedings of 2011 International Conference
on Electronic Engineering, Communication
and Management (EECM 2011),
held on December 24–25, 2011, Beijing, China



Springer

Prof. David Jin
International Science & Education Researcher Association
Wuhan
China
E-mail: dayang1818@163.com

Prof. Sally Lin
International Science & Education Researcher Association
Guang Zhou
China
E-mail: julia499780828@163.com

ISBN 978-3-642-27295-0

e-ISBN 978-3-642-27296-7

DOI 10.1007/978-3-642-27296-7

Lecture Notes in Electrical Engineering ISSN 1876-1100

Library of Congress Control Number: 2011943324

© 2012 Springer-Verlag Berlin Heidelberg

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

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

Typeset by Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

International Science & Education Researcher Association (ISER) puts her focus on studying and exchanging academic achievements of international teaching and scientific research, and she also promotes education reform in the world. In addition, she serves herself on academic discussion and communication too, which is beneficial for education and scientific research. Thus it will stimulate the research interests of all researchers to stir up academic resonance.

EECM2011 is an integrated conference concentrating its focus upon Electronic Engineering, Communication and Management. In the proceeding, you can learn much more knowledge about Electronic Engineering, Communication and Management of researchers all around the world. The main role of the proceeding is to be used as an exchange pillar for researchers who are working in the mentioned field. In order to meet high standard of Springer, LNEE series, the organization committee has made their efforts to do the following things. Firstly, poor quality paper has been refused after reviewing course by anonymous referee experts. Secondly, periodically review meetings have been held around the reviewers about five times for exchanging reviewing suggestions. Finally, the conference organization had several preliminary sessions before the conference. Through efforts of different people and departments, the conference will be successful and fruitful.

EECM2011 is co-sponsored by International Science & Education Researcher Association, Beijing Gireida Education Co. Ltd and Wuhan University of Science and Technology, China. The goal of the conference is to provide researchers from Electronic Engineering, Communication and Management based on modern information technology with a free exchanging forum to share the new ideas, new innovation and solutions with each other. In addition, the conference organizer will invite some famous keynote speaker to deliver their speech in the conference. All participants will have chance to discuss with the speakers face to face, which is very helpful for participants. During the organization course, we have got help from different people, different departments, different institutions. Here, we would like to show our first sincere thanks to publishers of Springer, LNEE series for their kind and enthusiastic help and best support for our conference, especially to Mr. Ditzinger Thomas for his great help. Secondly, the authors should be thanked too for their enthusiastic writing attitudes toward their papers. Thirdly, all members of program chairs, reviewers and program committees should also be appreciated for their hard work.

In a word, it is the different team efforts that they make our conference be successful on December 24–25, 2011, Beijing, China. We hope that all of participants can give us good suggestions to improve our working efficiency and service in the future. And we also hope to get your supporting all the way. Next year, in 2012, we look forward to seeing all of you at EECM2012.

October 2011

ISER Association

Committee

Honor Chairs

Prof. Chen Bin	Beijing Normal University, China
Prof. Hu Chen	Peking University, China
Chunhua Tan	Beijing Normal University, China
Helen Zhang	University of Munich, China

Program Committee Chairs

Xiong Huang	International Science & Education Researcher Association, China
LiDing	International Science & Education Researcher Association, China
Zhihua Xu	International Science & Education Researcher Association, China

Organizing Chair

ZongMing Tu	Beijing Gireida Education Co. Ltd, China
Jijun Wang	Beijing Spon Technology Research Institution, China
Quanxiang	Beijing Prophet Science and Education Research Center, China

Publication Chair

Song Lin	International Science & Education Researcher Association, China
Xionghuang	International Science & Education Researcher Association, China

International Committees

Sally Wang	Beijing Normal University, China
LiLi	Dongguan University of Technology, China
BingXiao	Anhui University, China
Z.L. Wang	Wuhan University, China
Moon Seho	Hoseo University, Korea
Kongel Arearak	Suranaree University of Technology, Thailand
Zhihua Xu	International Science & Education Researcher Association, China

Co-sponsored by

International Science & Education Researcher Association, China
VIP Information Conference Center, China

Reviewers of EECM2011

Chunlin Xie	Wuhan University of Science and Technology, China
LinQi	Hubei University of Technology, China
Xiong Huang	International Science & Education Researcher Association, China
Gangshen	International Science & Education Researcher Association, China
Xiangrong Jiang	Wuhan University of Technology, China
LiHu	Linguistic and Linguistic Education Association, China
Moon Hyan	Sungkyunkwan University, Korea
Guangwen	South China University of Technology, China
Jack H. Li	George Mason University, USA
Marry Y. Feng	University of Technology Sydney, Australia
Feng Quan	Zhongnan University of Finance and Economics, China
PengDing	Hubei University, China
Songlin	International Science & Education Researcher Association, China
XiaoLie Nan	International Science & Education Researcher Association, China
ZhiYu	International Science & Education Researcher Association, China
XueJin	International Science & Education Researcher Association, China
Zhihua Xu	International Science& Education Researcher Association, China
WuYang	International Science & Education Researcher Association, China
QinXiao	International Science & Education Researcher Association, China
Weifeng Guo	International Science & Education Researcher Association, China
Li Hu	Wuhan University of Science and Technology, China
ZhongYan	Wuhan University of Science and Technology, China
Haiquan Huang	Hubei University of Technology, China
Xiao Bing	Wuhan University, China
Brown Wu	Sun Yat-Sen University, China

Contents

Constructions and Application of Network Learning Platform	1
<i>Wenjun Liao, Guohong Gao, Ying Wang, Mingli Guo, Tong Jiang</i>	
A Blind Adaptive Matching Pursuit Algorithm for Compressed Sensing Based on Divide and Conquer	7
<i>Wenbiao Tian, Guosheng Rui, Zheng Fu</i>	
Multi-resolution Model Applied to Knowledge Base System for Personalized Service	13
<i>Zhenzhen Yi, Ke Zhao, Kai Li, Shuxing Du</i>	
The Construction of the Security Ability System in Coal Enterprise	19
<i>Liu-an Kong, Wen-jing Wu</i>	
Discussion on Certain Problems and Countermeasures on Expansion of Management Advisory Services of Accounting Company in China	25
<i>Chenyun Ye, Lingyan Kou</i>	
Research on Control System of Inspecting and Adjusting Off Pattern for Rotary Screen Printing Machine	33
<i>Aiqin Sun, Jidai Wang, Dongyue Zhang</i>	
Analysis of Test Scheme for a Bayesian Plan of Qualification Test in Binomial Case	39
<i>Zhiqian Ren, Zhimao Ming</i>	
Wireless Sensor Networks Fusion Based on the Hybrid of Ordering Filter and Adaptive Weight	45
<i>Changjiang Wang, Liqin Fu, Yongmei Zhang</i>	
Process Control and Intelligent Algorithms of Network Optimization Problems under Uncertain System	51
<i>Jing Zhang</i>	

A Hybrid Clustering Based on ACO and Single-Pass	57
<i>Wen Xiong</i>	
Advanced Sensing Device for Gesture Recognition	63
<i>Yong Mu Jeong, Ki-Taek Lim, Seung Eun Lee</i>	
Research on the Reactive Power Control of Distributed Photovoltaic System	67
<i>Zhiqiang Jia, Weiping Luo</i>	
Application of Gateway in Fault Location System of Belt Conveyor	73
<i>Shaojie Yin, Jinsheng Sun</i>	
A Contract Performing Scenario Model for Web Services Collaboration . . .	81
<i>Denghui Zhang, Ji Gao</i>	
A SVM Approach for MCs Detection by Embedding GTDA Subspace Learning	89
<i>Xinsheng Zhang, Minghu Wang, Fan Yu</i>	
Differential Evolution Using Local Search for Multi-objective Optimization	95
<i>Youyun Ao</i>	
The Research about Fuzzy Control of Electric Generator Stator Bar-Ends Cure-Heating	103
<i>Jie Bai, GaoJun Li, XiaoDong Tan</i>	
The Study on Phase-Control of Vacuum Circuit Breaker for Electric Railway Substation	109
<i>Jie Bai, KeLiang Shu, XiaoDong Tan</i>	
General Information Diffusion Method in Risk Analysis	115
<i>Qiong Li, Jianzhong Zhou, Donghan Liu</i>	
Multi-agent Based Bank Queuing Model and Optimization	121
<i>Qicong Zhang</i>	
Simulation of Agent-Based of Intelligent Public Transit System	129
<i>Qicong Zhang</i>	
Design and Simulation of Fuzzy Decoupling for Combustion Process of Circulating Fluidized Bed Boiler	137
<i>Hao Tong, Xiaopeng Zhao, Guozhuang Liang</i>	
The Application of Ant-Colony Clustering Algorithm to Earthquake Prediction	145
<i>Xiaoyan Shao, Xiang Li, Lingling Li, Xinru Hu</i>	

Rapid LED Optical Model Design by Midfield Simulation Verified by Normalized Cross Correlation Algorithm	151
<i>YongJiang Di</i>	
Wireless Data Acquisition and Control System Based on DCS	157
<i>Genghuang Yang, Abreham Fekibelu, Feifei Wang, Shigang Cui, Li Zhao</i>	
ARM and DSP Based Device for Power Quality Monitoring	163
<i>Genghuang Yang, Feifei Wang, Shigang Cui, Li Zhao</i>	
Chinese Ecological Risk Early-Warning Indicator System and Empirical Research	169
<i>Zeng Yongquan</i>	
The Optimization on the Multiperiod Mean Average Absolute Deviation Fuzzy Portfolio Selection	175
<i>Zhang Peng, Pan Jin</i>	
$[r, s, t]$-Colouring of Join Graphs $S_n + O_m$	181
<i>Mingzhong Mo</i>	
Study on Product Design for Seniors	187
<i>Chunqing Yang</i>	
A Time-Efficient FPGA-Based FFT Processor	191
<i>Xudong Wang</i>	
The Research on the Impact of Project Communication to Project Performance	195
<i>Zhongqing Wang, Qiang Hu</i>	
Time-Frequency Filtering for LFM Based on Time S-Transform and Time Inverse S-Transform	203
<i>HuaShan Chi, HongXing Wang, Min Zhang, PeiHong Zhao</i>	
A BP-LZ77 Compression Algorithm Based on BP Network	211
<i>HongWei Yan, HuiShan Lu, Qiang Gao</i>	
Modeling and Optimization of Frequency Hopping Sequence Designation in Asynchronous Network	217
<i>DaoCheng Lei, YongLin Yu, Gang Wang, RunNian Ma</i>	
Chinese Demographic Risk Early-Warning Indicator System and Empirical Research	225
<i>Zeng Yongquan</i>	
A Sparse Recovery Method of DOA Estimation for Uncorrelated Sources with Multi-snapshot	231
<i>MaoWei Yin, DongYang Xu, ZhongFu Ye</i>	

Compensation of Double Closed Loop Control Systems with Time-Delay and Data Packet Losses	237
<i>HuiSheng Gao, XiaoRui Li, Jun Wang</i>	
Optimization Design of Improving Signal Integrity for Enhancing Differential-Mode Transmission Lines	243
<i>Wen-Jong Chen, Ming-Huang Wu</i>	
Coordinate Transformations in Satellite Navigation Systems	249
<i>Pengfei Zhang, Chengdong Xu, Chunsheng Hu, Ye Chen</i>	
The Verification Method Research of Mining Right Boundary Based on GIS Technology in Shanxi Province	259
<i>Bei Jia, Huijuan Li, Baoping Wang</i>	
Research on Evaluation Model of Digital Library Service Quality	267
<i>YanHui Liu, DaLi Xie, XiaoBo Wang</i>	
Bayesian Model Averaging for Lasso Using Regularization Path	273
<i>Mei Wang, Erlong Yang</i>	
Optimization Design of Parameters with Hybrid Particle Swarm Optimization Algorithm in Multi-hole Extrusion Process	279
<i>Wen-Jong Chen, Dyi-Cheng Chen, Wen-Cheng Su, Fung-Ling Nian</i>	
Decision-Making Model of Business Continuity Management	285
<i>Gang Chen</i>	
Research on Energy Supply Modes of Electric Vehicle	291
<i>Changxian Cheng, Jianhua Yan</i>	
Why Does Firm Choose Internet Marketing Channel? A Firm Perspective and Contingent Framework	297
<i>Maohong Liu, Xiaoliang Feng, Minxue Huang</i>	
Study on Diluted Symmetric Discrete-Time Hopfield Network	303
<i>Li Tu, Yan Wang</i>	
Research on the BM Algorithm Based on the Intrusion Detection System ...	309
<i>Xuewu Liu, Lingyi Peng, Chaoliang Li</i>	
Modeling and Analysis of DNA Mutation Effects on Protein Structure with Finite State Machine	317
<i>Rui Gao, Wen-Song Hu, Cheng-Qiu Zhang</i>	
Design and Implementation of Solar Thermal Utilization Information System Based on GIS	325
<i>Yu Zhang, Shijun He, Shutai Zhang, Zhonghua Chen, Wenjun Zhou, Yanyan Dai, Shiting Zhao, Fan Bai</i>	

Research of Lip Contour Extraction in Face Recognition	333
<i>Kunlun Li, Ruining Xin, Miao Wang, Hui Bai</i>	
Public-Keys-Based Entity Authentication Protocol with Digital Signature for Cognitive Radio Networks	341
<i>Chao Wang, JiWen Guo</i>	
An IPA Analysis of Effective Teaching Activities in China	347
<i>Haiyan Kong</i>	
A Robust Control Scheme for Nonlinear Systems	353
<i>Limin Liu</i>	
A Research for Bilingual Courses of Embedded Systems	359
<i>Limin Liu</i>	
Research on Negative Inventory Balances Management of Modern Manufacturing Industry under ERP Environment	365
<i>LinKun Li, KaiChao Yu, YiMing Li</i>	
Design of Automatic Control System of Fructose Crystallization	371
<i>Jian Chu, Qing Lu</i>	
Text Clustering Method Based on Improved Genetic Algorithm	377
<i>ZhanGang Hao, Tong Wang, XiaoQian Song</i>	
Layout Optimization Method of Traffic Information Data Acquisition System Based on Linear Programming	383
<i>Wei Li</i>	
Application of MUSIC Algorithm to Scenario-Based Training Database Construction Method of ISAR Images	389
<i>Sang-Hong Park, In-O Choi, Jae-Heung Ye, Young-Jun Kim</i>	
The LFT Formulation of Dynamical Systems with the Uncertainties of Parameter Variations	395
<i>Jieh-Shian Young</i>	
The Design and Implementation of Wireless Network System for Temperature Control Based on ARM	403
<i>Tianhao Hu, Kunchang Chen, Xinli Wu, Dichong Wu, Canlin Mo, Renwang Li</i>	
Research of Interactive Product Design for Virtual Tourism	411
<i>Zhongyan Hu, Zaihui Cao, Jinfa Shi</i>	
Rapid Implementation of Communication-Related Simulation Equipment on the Open Web-Based Virtual Lab	417
<i>Yue Zhao, Fuan Wen</i>	

The Principle and Application Exploration of Kademlia's Structured P2P	425
<i>Xiaochun Shi, Guoqing Liu, Jinchun Feng</i>	
Automatic Color Temperature Control Circuit	431
<i>Tai-Ping Sun, Chia-Hung Wang, Huei-Jyun Lin, Kai-Yan Li</i>	
Study on Integrated Circuit Design for Multi-color LED Pulse-Width Modulation	439
<i>Tai-Ping Sun, Chia-Hung Wang, Fu-Shiung Tsau, Jia-Hao Li</i>	
An Embedded Controller for Smart Products	447
<i>Limin Liu</i>	
Fuzzy Comprehensive Evaluation on the Quality of Vehicle Equipment Maintenance Based on Membership Degree Transformation Algorithm	453
<i>XiaoLing Wei, ZhiWen Chen, ZhenZhong Zhou</i>	
Moving Vehicle Detection Based on Union of Three-Frame Difference	459
<i>Caiyuan Chen, Xiaoning Zhang</i>	
Improving the Learning Performance by Providing Intellectual Suggestions	465
<i>Mohammad Sarem, Mostafa Bellafkih, Mohammad Ramdani</i>	
The Study of Substation Automation Remote Communication for Multi-granularity Access Control System	473
<i>Lichun Shi, Haimei Xiao</i>	
Parameter Estimation of Sinusoidal Signal under Low SNR	481
<i>Mingjian Sun, Xiaowei Zhang, Fanming Meng</i>	
Contour Algorithm Based on GVF Snake	491
<i>Fan Zhang, Shengwei Li, Gang Zheng, Rui Li</i>	
The Design of Partial Electric and Control System for 500kV Substation ...	497
<i>Haimei Xiao, Lichun Shi</i>	
Research on the Secure Access of E-Commerce	505
<i>Longjun Zhang</i>	
A Secure Authentication Scheme of Wireless Mesh Network	511
<i>Longjun Zhang</i>	
Research on Capacity Allocation in a Supply Chain System Based on TOC	517
<i>Kaijun Leng, Yuxia Wang</i>	
Research of Data Acquisition Controller for DO Based on FPGA	525
<i>WeiSheng Zhong, YaPing Wang</i>	

Research on the Simplification of BOM Maintenance under the Condition of ERP	531
<i>Jia Zhang, Kaichao Yu</i>	
Simulation and Analysis of the Common-Emitter Circuit in Analog Electronics Teaching	537
<i>Xiao-Hong Wu, Bin Wu, Jun Sun, Zhen-Yu Wang</i>	
Research on the Evaluation of the Third Party Reverse Logistics Suppliers	543
<i>QingKui Cao, JingHua Zhang</i>	
Simulation Analysis of Traffic Capacity for Ticket Gates of Metro Station	549
<i>Xin Han, Xie He, Beihua Cong</i>	
Analysis of Compensation Strategies for Dynamic Voltage Restorer Based on DSFR-PLL Control	555
<i>Xiaoying Zhang, Jun Yan, Zhiwei Wen, Dan Jin</i>	
Design of H.264 Encoder Based on ADSP-BF561	563
<i>Qinghui Wang, Guian Wu, Ying Gao</i>	
Optimized Implementation of H.264 Integer Transform Based on Blackfin 533	569
<i>QingHui Wang, HuaPing Xue, JingJing Yang</i>	
Power Line Communication System Design for LED Transmission	575
<i>Tai-Ping Sun, Chia-Hung Wang, Jia-Yu Chen</i>	
The Design of Multi-parameter and Portable Monitor for Sleep Apnea Status	583
<i>XiaoMei Liu, Ying Lian, ShaoChun Chen, ShanShan Tang</i>	
The Design of Body Sensor Network Based on CC430	589
<i>XiaoMei Liu, ShanShan Tang, ShaoChun Chen, Ying Lian</i>	
Historical and Status Quo Analysis on the Production and Consumption of China's Electric Power Energy	595
<i>Jie Yao, Jin Zhu, Wei Gao</i>	
A High Performance Audio Watermarking Algorithm Based on Substitution in Wavelet Domain	601
<i>Huan Hao, Liang Chen, Yipeng Zhang, Yinguang Zhang, Bo Zhang</i>	
Generalization Model of State Observers Based on Delay-Quantization Method	609
<i>Leiming Liu, Chaonan Tong</i>	

Design of Doherty Amplifier in Digital TV Transmitter	615
<i>Chen Jun, Su Kaixiong</i>	
Fusion Method of Feature Weight Based on Evidence Support Matrix	621
<i>Guochu Chen, Rui Miao, Yue Li</i>	
Implementation and Application of DSP-LCD Driver Based on Framebuffer	627
<i>Qinghui Wang, Xiaoying Chang, Ansong Feng</i>	
A Novel Energy Efficient Cluster-Based Algorithm for Wireless Sensor Networks	633
<i>Bin Jiao, Youping Xiong</i>	
A 3D Temperature Field Reconstruction Algorithm Based on Acoustic Travel-Time Tomography	639
<i>Hua Yan, Kun Li, ShanHui Wang</i>	
Retraction: Evaluating Gigabit Switches Using Perfect Communication	645
<i>JianMing Zhao, YongNing Guo</i>	
Application of Virtual Prototype Technology to Simulation Test for Airborne Software System	653
<i>Yang Wang, Lifeng Wang, Zewei Zheng</i>	
Retraction: Deconstructing the Ethernet	659
<i>Zhong Chen, YongNing Guo</i>	
Comment on a Research and Analysis Four-Prime RSA	669
<i>Yongning Guo, Chenglian Liu</i>	
An Improved FastIMM Algorithm Based on α-β and α-β-γ Filters	677
<i>Junchen Sha, Jianping Xing, Zhenliang Ma, Liang Gao, Can Sun, Juan Sun</i>	
An Approach on Feature Selection of Cascaded Support Vector Machines with Particle Swarm Optimization Algorithm	689
<i>Ruihu Wang, Zhangping Hu, Liang Chen, Jing Xiong</i>	
Quantitative Analysis of Survivability Based on Intrusion Scenarios	701
<i>Chengli Zhao, Zhiheng Yu</i>	
Facial Expression Recognition Based on Local Binary Patterns and Least Squares Support Vector Machines	707
<i>Xiaoming Zhao, Shiqing Zhang</i>	
Facial Expression Recognition Using Kernel Locality Preserving Projections	713
<i>Xiaoming Zhao, Shiqing Zhang</i>	
Author Index	719

Constructions and Application of Network Learning Platform

Wenjun Liao¹, Guohong Gao², Ying Wang³, Mingli Guo², and Tong Jiang²

¹ Institute of computer and Information Engineering
Xinxiang University, Xinxiang, China, 453003

² School of Information Engineer
Henan Institute of Science and Technology, Xinxiang, China, 453003

³ Xinke College
Henan Institute of Science and Technology, Xinxiang, China, 453003
914747841@qq.com

Abstract. The research study is a kind of learning way with self-learning, problem-centered approach to learning research. Network environment based on J2EE platform research considers study learning theory as basis. It makes use of advanced technology and J2EE platform, JSP and JavaBean technology to build the system and takes advantage of SQL database to have management background, which has the advantage of advanced, security, manageability and maintainability in the development. After testing, learning platform meets the application requirements.

Keywords: Research study, Learning platform, JSP, Personalization.

1 Introduction

Research study is a learning activity with researching nature of the problem it looks on issues as a carrier and regards research study as the main way of learning. Research study is a strong practical education and teaching activities, it is not just limited to books for students imparting knowledge, but let students participate in practical activities and carry out research and exploration in practice, so that they can learn studying and acquiring skills[1].

Web-based research study is a conformity between computer network and the research study, and is a research study supported by network. It builds a dynamic, open, interactive, informative greatly learning environment for students. Web-based research study to borrow WEB technology and Internet can realize independent inquiry, collaborative learning, personalized learning and other learning[2].

Currently, Web-based learning model of continuous exploration and study, WebQuest is often used abroad with the general pattern, there is no national model of a uniform standard, but the scope is defined broadly. With the ongoing of Web-based learning theory and practice of research, domestic and school-related educational software company has been trying to research the development of learning platforms, such as classroom research study of the sky platform, Tsinghua Tongfang research study Platform [3].

2 Requirements Analysis and Related Technologies of System

2.1 Requirements Analysis

Requirements analysis is the beginning of software development, which is essential to the entire development process and software product quality. Through the survey of Henan Science and Technology the College of Information Engineering Educational Technology, it defined the basic needs of users and identified the students, teachers and administrators three user roles.

(1) Students: Students want to choose research topics from a variety of ways to facilitate access to and management of research resources they need. In the research process, their own difficulties can receive guidance and help from teacher or other students. They can communicate and exchange research topics with the instructor and other students through a variety of ways such as chat rooms, BBS, etc. With a variety of learning tools such as Word, Excel, concept maps to deal with collected data of their own other research topics, meanwhile, students can make use of some multimedia tools such as Flash, Dreamweaver, etc to produce their own electronic works so that they have a better performance and description of their research results. Usage of existing network transport technologies such as the World Wide Web, FTP and so on makes their research results publicized, so that more users provide guidance for subsequent research and reference through user browsing and evaluating feedback on the research results.

(2) Teachers: Teachers want to understand the characteristics and the basic information of students, such as learning styles, attitudes, and level of networking operations so as to provide important information for students choosing topic and carrying out the and research topics. Besides, teachers know information on the students research projects in time, so that it is easy to communicate and exchange with students to guide students and issues management. They also provide students with issues related to resources, case studies and curriculum. Finally, they guide the students ways of acquiring resources, access methods, data collection and analysis, research thesis writing and publishing research results[3].

(3) Administrators: Administrators hope that the information can be easily published and modified other operations; they hope to facilitate the management of each module on the platform and acquire relevant platform for users comments or suggestions to adjust and optimize the platform.

2.2 Development Related Technologies

According to the router bearing function clustering division, the router is usually Currently web-based learning platform of network research is very popular, it is based on Browser/Server (browser/server) architecture to develop, so-called B/S mode. It consists of a browser, Web server and database servers. The structure of Network environment based on J2EE for the research study shown in Figure 1. standards-based J2EE JSP and JavaBean technology are mainly used in the development of this system.

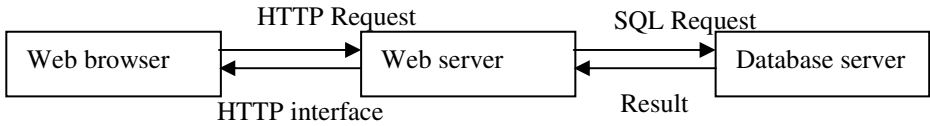


Fig. 1. Platform structure

3 Module Design

According to the above functional needs analysis, the platform mainly consists of registration and login module, the student characteristics, navigation, my laboratory, research, resource centers, teacher style, chat rooms, forums, and backstage management function module. Platform module structure shown in Figure 2.

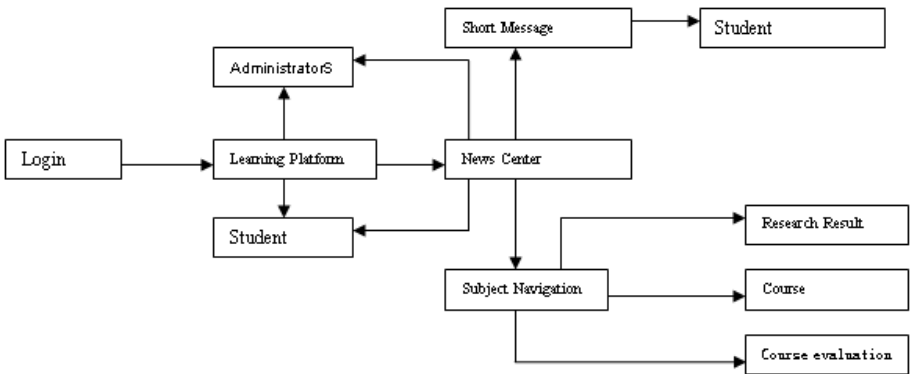


Fig. 2. Platform module

Features related to the core module are as follows:

(1) Characteristics of Students

Characteristics of Students is an important part of web-based research learning implement. Research in Web-based learning, students characteristics not only includes the factors of intellectual, such as intellectual level, knowledge base, cognitive abilities and other characteristics; characteristics of non-intellectual factors ,for example, interest, motivation, emotion, will, character and Learning styles, etc, but also includes the students computer skills, technology status and ability to cooperate, cooperation, awareness and other features.

(2) Subject Navigation

Subject Navigation module not only can view related major detailed information such as professional training objectives, curriculum, etc, but also choose courses related to learning resources for courses related downloads. It is easy for teachers to add the curriculum and modify other management. Teachers may also need the help of questions quickly and creating paper templates of the students for self-diagnostic tests. Platform

automatically statistics students answer and analyzes all the knowledge students have commanded. Teachers can adjust teaching strategies according to the situation of students commanded all knowledge so as to guide students in learning better.

(3) Communication Center

Research in Web-based learning, teachers and students are in the state of space-time separation, so it is important that teachers and students how to exchange. The module's main function is to provide convenient communication tool for the teachers and students. Real-time communicative tools includes chat rooms, NetMeeting, Windows Messenger and so on. Students can research the problems encountered in the process through chat rooms. Students also can take advantage of chat rooms on their own topics of interest to study the hot issues discussing with other students in order to find the problem.

4 Construction of the Database Platform

Database construction is very important in the system module. It is an important part of three-tier architecture[4]. Database selection and building construction is related to the success of the platform. Thus, we chose SQL SERVER 2005 database software to build back-end database. The database system provides enterprise-class database development and management of the environment card, which has higher security, availability, and manageability.

Research study is mainly responsible for the database platform of data storage. Because of more data table involved in the database, the user use the platform to require a large number of tables in the database for complex queries. In order to ensure data security and reduce the complexity of the query, view is used to display the data. Part of the data table and its description of the function in the database is shown in Table 1.

Table 1. Related data sheet

Table Class	Table Name	Explain
Student Information sheet	stuinfo	Storing information for the students, registered user and log in using
Teacher Information sheet	teainfo	Storage teacher information, log on using
Administrator Information sheet	admininfo	Information storage administrator, log management and use of background maintenance
Subject information sheet	courseinfo	Information stored in the subject
Survey	quest	Survey related to information storage
Subject resources	source	Issues related to the contents of storage resources
Research diary	daily	Log storage research
Course	coustudy	Store course-related steps, instructions, etc
Course evaluation	evaluate	Evaluation information storage courses
Short Message	message	Storage platform timely release of the short message
Record learning activities	actrecord	Record store learning activities

5 Conclusion

Network environment is based on J2EE Platform research study learning theory. Taking into account the temporal and spatial separation of the teachers and students, the platform introduces the module of characteristics students. Building a web-based research study of the dynamic assessment model, it was evaluated from the multiple dimensions of the learning process, research, capacity development. Development platform of the research study attempts to research information technology and integrated training mode. Advanced technology of JE22 is used in the system development platform, and JSP and JavaBean technology is also used to build the system. The usage of SQL database management background has advantages in the development of advanced, security, manageability and serviceability.

References

1. Li, X.-Y., Yang, G.-Q.: Design and implementation of access network technology network teaching platform. *Information Technology* (1), 102–105 (2011)
2. Zeng, Z.-M.: Research and Implementation of Art Design Network Teaching Platform Based on J2EE. *Modern Computer* (3), 158–161 (2010)
3. Xu, C., Wang, Y.: Design and Applications of Web Instruction Platform. *Network Security Technology & Application* (11), 75–77 (2010)
4. Tang, Z., Liang, X., Luo, D.: Constructions and Application of the English Study platform in the Network Age. *China Computer & Communication* (07), 41–45 (2009)

A Blind Adaptive Matching Pursuit Algorithm for Compressed Sensing Based on Divide and Conquer

Wenbiao Tian*, Guosheng Rui, and Zheng Fu

Signal and information processing provincial key laboratory in Shandong,
Naval Aeronautical and Astronautical University, Yantai, Shandong 264001, China
twbi5si@gmail.com

Abstract. The existing reconstruction algorithms are mainly implemented based on the known sparsity of original signal. In this paper, a new algorithm called blind adaptive matching pursuit (BAMP) is proposed in this paper, which can recover the original signal fast in the case of unknown sparsity based on the divide and conquer method. Firstly, the range of sparsity is determined by trial and error test. Secondly, the efficient support set is screened out rapidly by adaptive grouping and support extension. Last but not least, reconstructed signal is obtained by pruning. The results of simulation show that the new algorithm can reconstruct signal faster and get better precision than other similar algorithms in the same conditions.

Keywords: Compressed sensing, signal processing, blind sparsity, adaptive reconstruction.

1 Introduction

Around 2004 Emmanuel Candès, Terence Tao and David Donoho discovered important results on the minimum number of data needed to reconstruct a signal even though the number of data would be deemed insufficient by the Nyquist–Shannon criterion, i.e. Compressed Sensing (CS) [1, 2]. CS theory asserts that one can recover certain signals and images from far fewer samples or measurements than traditional methods use. To make this possible, CS relies on two principles: sparsity, which pertains to the signals of interest, and incoherence, which pertains to the sensing modality, and it can be demonstrated that the number of these compressive measurements can be small and still contain nearly all the useful information. The theory is widely applied to CS Radar, biomedicine, wireless sensor network and remote sensing image processing etc [3, 4].

Combining the idea of self adapting like SAMP [5] with rapid grouping test as is done in CoSaMP [6], blind adaptive matching pursuit (BAMP) is proposed based on the divide and conquer (D&C) estimation of sparsity.

* Born in 1987, Ph.D candidate of Naval Aeronautical and Astronautical University (NAAU). His research interests are in communication signal processing with emphasis on compressive sampling and analog-to-information conversion.

2 CS Reconstruction Theoretic Framework

An $N \times 1$ discrete time signal vector \mathbf{x} is K -sparse or compressible in some sparsity basis matrix Ψ (where each column is a basis vector ψ_i), i.e., only $K \ll N$ of the expansion coefficients \mathbf{a} representing $\mathbf{x} = \Psi\mathbf{a}$ are nonzero. The theory of CS recovery demonstrates that the signal can be recovered from $M = cK \ll N$ non-adaptive linear projections onto a second basis $\Phi : M \times N$ that is *incoherent* with the first (Ψ), where c is a small *overmeasuring* constant. Thus, rather than measuring the N -point signal \mathbf{x} directly, the M linear projections $\mathbf{y} : M \times 1$, $\mathbf{y} = \Phi\mathbf{x}$ are acquired. The recovery of the signal \mathbf{x} can be achieved using optimization by searching for the signal with ℓ_0 -sparsest coefficients \mathbf{a} that agrees with the M observed measurements in \mathbf{y} ,

$$\hat{\mathbf{x}} = \Psi\hat{\mathbf{a}} \quad \hat{\mathbf{a}} = \arg \min \|\mathbf{a}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \Phi\mathbf{x}. \quad (1)$$

While solving this ℓ_0 -optimization problem is NP-hard, equivalent solution can be acquired by solving an easier ℓ_1 -optimization [7],

$$\hat{\mathbf{x}} = \Psi\hat{\mathbf{a}} \quad \hat{\mathbf{a}} = \arg \min \|\mathbf{a}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \Phi\mathbf{x}. \quad (2)$$

There are three kinds of recovery algorithms: (1) iterative greedy algorithms like MP [8], OMP [9], StOMP [10], CoSaMP, SAMP, etc.; (2) convex relaxation algorithms as BP [7] and Iterative thresholding [11], etc.; (3) combined algorithms such as Chaining Pursuit [12], HHS Pursuit [13], etc. BP-class algorithms, which start from a “full” model (i.e., a representation of the object in a basis) and then iteratively improve the “full” model by taking relatively useless terms out of the model and swapping them for useful new ones [7], are global optimization ones and a bit more complex than greedy iteration. In contrast, iterative greedy algorithms start from an “empty model” and build up a signal model an atom at a time, at each step adding to the model only the most important new atom among all those not already in the model. The existing reconstruction algorithms are mainly implemented based on the known sparsity of original signal, while the sparsity is difficult to be pre-informed in practical conditions. Recently, many papers [14, 15] have studied blind sparsity reconstruction algorithms by stage-wised and regularized method in order to estimate the true supporting set of the approximated signal.

Based on D&C sparsity estimation, blind adaptive matching pursuit (BAMP) is proposed. BAMP Combines the idea of self adapting like SAMP with rapid grouping test as is done in CoSaMP and is essentially a greedy iterative algorithm of solving ℓ_0 -norm minimization problems with unknown sparsity.

3 Blind Adaptive Matching Pursuit Algorithm

At each iteration, greedy iterative algorithms choose the column of measurement matrix Φ that is most strongly correlated with the remaining part of projections \mathbf{y} . Then its contribution is subtracted off to and the residual is iterated on. One hopes

that, after m iterations, the algorithm will have identified the correct set of columns. Let $|\text{supp}(\mathbf{x})|$ be the size of \mathbf{x} support set, i.e.

$$\|\mathbf{x}\|_0 = |\text{supp}(\mathbf{x})| = |\{j : x_j \neq 0\}|. \quad (3)$$

Given a sparse signal \mathbf{x} , $\Lambda = \text{supp}(\mathbf{x})$

$$\mathbf{x}|_{\Lambda} = \begin{cases} x_i, & i \in \Lambda \\ 0, & i \in \Lambda^c. \end{cases}$$

We say that a signal \mathbf{x} is k -sparse, when $\|\mathbf{x}\|_0 \leq k$. If the sparsity estimation is larger than the truth value ($K \geq k$) in the process of trial and error, the support set of recovery signal $\text{supp}(\hat{\mathbf{x}})$ will be extended constrainedly. That the relative error σ is greater than 1 after several times iterations is the necessary condition of support domain overestimation, where

$$\sigma = \|\hat{\mathbf{x}} - \mathbf{x}\|_2 / \|\mathbf{x}\|_2.$$

To the contrary, if the support domain is underestimated, the relative error $\sigma \leq 1$. Thus σ can be regarded as an offline criterion of sparsity estimation, however, the signal truth value \mathbf{x} is unknown in the process of iteration. The online sparsity estimation is discussed next.

Based on D&C Sparsity Estimate, BAMP combining the idea of self adapting like SAMP with rapid grouping test as is done in CoSaMP is proposed, algorithm flowchart as Fig. 1.

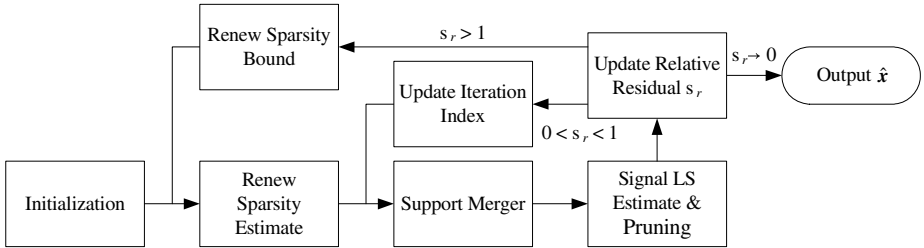


Fig. 1. BAMP algorithm flowchart

To identify the signal sparsity, we need to determine *which* interval the sparsity belongs to by D&C (the outer iteration). The idea behind D&C is to estimate sparsity and eliminate half of values in current range which is incorrect by trial and error test. At each inner iteration, we choose the columns of measurement matrix that are most strongly correlated with the remaining part of the residual. Then we merge them with the current support set and estimate signal by least squares. Last, we prune it to obtain next approximation and iterate on the relative residual. Pseudocode for BAMP appears as follows:

```

BAMP (  $\Phi, \mathbf{y}$  )
Input: Measurement matrix  $\Phi_{M \times N}$ , measurement vector  $\mathbf{y}_{M \times 1}$ 
Output: Target signal estimation  $\hat{\mathbf{x}}_{N \times 1}$ 


---


 $\mathbf{r}_0 = \mathbf{y}$ ;  $K_b = M$ ;  $K_a = 0$ ;  $\hat{\mathbf{x}}_0 = 0$ ;  $n = s = 1$  //Initialization
WHILE (  $K_b - K_a > 1$  ) //Outer Iteration
{
 $K_s = (K_a + K_b)/2$  //Sparsity Estimate
WHILE ( 1 ) //Inner Iteration
{
 $\mathbf{v}_n^{(s)} = \Phi \mathbf{r}_n^{(s)}$ 
 $\Lambda = \text{supp}(\mathbf{v}_{2K_s}) \cup \text{supp}[\hat{\mathbf{x}}_{n-1}^{(s)}]$  //Support Merger
 $\mathbf{b}|_{\Lambda}^{(s)} = \Phi_{\Lambda}^+ \mathbf{y}$ ;  $\mathbf{b}|_{\Lambda^c}^{(s)} = 0$  //Signal LS estimate
 $\hat{\mathbf{x}}_n^{(s)} = \mathbf{b}|_{K_s}^{(s)}$  //& Pruning
 $\mathbf{r}_n^{(s)} = \mathbf{y} - \Phi \hat{\mathbf{x}}_n^{(s)}$  //Update Relative
 $\sigma_r = \|\mathbf{r}_n^{(s)}\|_2 / \|\mathbf{y}\|_2$  //Residual  $r$ 
IF  $\sigma_r > 1$ 
 $K_b = K_s$ ;  $s = s + 1$ ; BREAK
END
 $n = n + 1$  //Update Iteration Index
}
IF  $\sigma_r \rightarrow 0$ 
BREAK
ELSEIF  $\sigma_r < 1$ 
 $K_a = K_s$ ;  $s = s + 1$ ; BREAK
END
}

```

As is shown in the algorithm, whether the sparsity should be adjusted is decided self-adaptively by relative residual, that is, whether next trial of sparsity or next recovery iteration will be introduced in. The sparsity is hence unnecessary to be known as a prior knowledge of signal reconstruction and is fast approximated by trial and error test. BAMP avoid halting criterion failure resulting from the step is too small (far less than signal sparsity) and over matching that recovery accuracy is achieved by many iterations caused by inappropriate threshold, and these problems may emerge from SAMP, RAMP and other similar algorithms. Back-track sifting ensures signal support set to be ascertained fast and precisely.

4 Experiments Results and Discussions

In this experiment set, we compare the performance of the SAMP, the RAMP and the BAMP in the practical compressed imaging scenario. 256×256 test images *Lena* is chosen and the sparsifying matrix is the discrete cosine transform matrix. The original image is divided into 256 blocks (16×16) and each block is reshaped into a column for next processing. Gaussian random matrices were used as the sampling matrix. The initial values of variable step sizes are set to be $M/2\log(N)$ in both RAMP and SAMP. The visual reconstructions of the image *Lena* 256×256 from $M=N/2$ measurements is shown in Fig. 2.



Fig. 2. The comparison of the *Lena* image recovery under $M/N=0.5$ for diversity algorithms

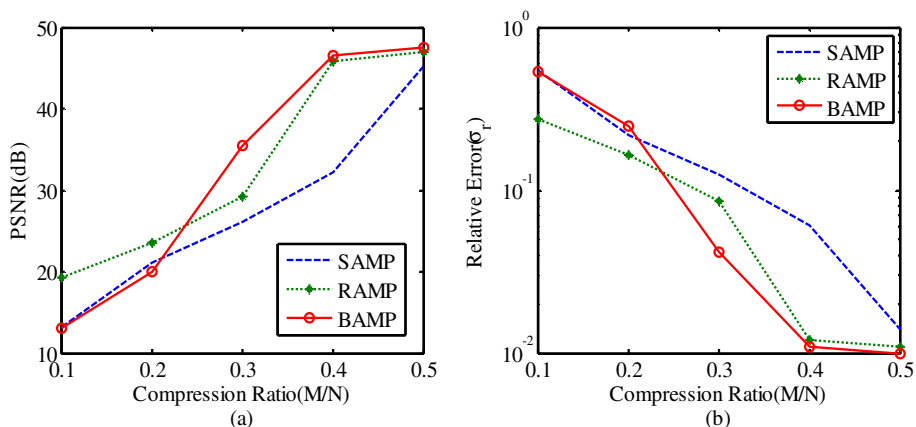


Fig. 3. Experiment results of the three algorithms (a) PSNR vs. Compression Ratio (b) Relative Error vs. Compression Ratio

Fig. 3(a) and Fig. 3(b) demonstrate the results for PSNR and relative error comparison, respectively. As can be seen, for PSNR comparison, performance of the BAMP exceeds that of all other algorithms, especially under high compression ratio. For relative error comparison, the BAMP is better than other algorithms while $M/N \geq 0.3$.

5 Conclusions

In this paper, a novel greedy pursuit algorithm, called the blind adaptive matching pursuit, is proposed and analyzed for reconstruction applications in compressed sensing. As its name suggests, this reconstruction algorithm is most featured of not requiring information of sparsity of target signals as a prior. The results of simulation show that the new algorithm can reconstruct signal faster and more precisely than other similar algorithms in the same conditions.

References

1. Baraniuk, R.G.: Compressive Sensing. *IEEE Signal Processing Magazine* 24, 118–121 (2007)
2. Donoho, D.L.: Compressed sensing. *IEEE Trans. on Information Theory* 52, 1289–1306 (2006)
3. Donoho, D.L., Tsaig, Y.: Extensions of compressed sensing. *Signal Process.* 86, 533–548 (2006)
4. Guangming, S., Danhua, L., Dahua, G., Zhe, L., Jie, L., Liangjun, W.: Advances in Theory and Application of Compressed Sensing (in Chinese). *Acta Electronica Sinica* 37, 1070–1081 (2009)
5. Do, T.T., Gan, L., Nguyen, N.S.: Sparsity adaptive matching pursuit algorithm for practical compressed sensing. In: *Asilomar Conference on Signals, Systems and Computers* (2008)
6. Needell, D., Tropp, J.A.: CoSaMP: Iterative Signal Recovery from Incomplete and Inaccurate Samples. *Commun. ACM* 53, 93–100 (2010)
7. Chen, S.S., Donoho, D.L., Saunders, M.A.: Atomic Decomposition by Basis Pursuit. *SIAM J. Sci. Comput.* 20, 33–61 (1998)
8. Mallat, S., Zhang, Z.: Matching pursuits with time-frequency dictionaries. *IEEE Trans. on Signal Processing* 41, 3397–3415 (1993)
9. Tropp, J.A., Gilbert, A.C.: Signal recovery from random measurements via orthogonal matching pursuit. *IEEE T. Inform. Theory* 53, 4655–4666 (2007)
10. Donoho, D.L., Tsaig, Y., Drori, I., Starck, J.: Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit (2006)
11. Blumensath, T., Davies, M.E.: Iterative thresholding for sparse approximation. *Fourier Anal.* 14, 629–654 (2008)
12. Gilbert, A., Strauss, M., Tropp, J., Vershynin, R.: Algorithmic linear dimension reduction in the l_1 norm for sparse vectors., <http://www.math.ucdavis.edu/~vershynin/papers/algorithmic-dim-reduction.pdf>
13. Gilbert, A.C., Strauss, M.J., Trop, J.A., Al, E.: One sketch for all: Fast algorithms for compressed sensing. In: *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pp. 237–246. Association for Computing Machinery (2007)
14. Zongnian, Z., Rentai, H., Jingwen, Y.: A Blind Sparsity Reconstruction Algorithm for Compressed Sensing Signal (in Chinese). *Acta Electronica Sinica* 39, 18–22 (2011)
15. Yaxin, L., Ruizhen, Z., Shaohai, H., Chunhui, J.: Regularized Adaptive Matching Pursuit Algorithm for Signal Reconstruction Based on Compressive Sensing (in Chinese). *Journal of Electronics & Information Technology* 32, 2713–2717 (2010)

Multi-resolution Model Applied to Knowledge Base System for Personalized Service

Zhenzhen Yi, Ke Zhao, Kai Li, and Shuxing Du

College of Mechanical&Electrical Engineering, Xidian University, Xi'an
710071, China
yizzhen@yahoo.com.cn

Abstract. This paper focuses on problem-solving expert system to generate the more personalized solutions. The multi-resolution model is combined into the knowledge base organization of expert system. Firstly, the combination possibility of multi-resolution model and knowledge expert system is discussed and the multi-resolution organization structure and representation of knowledge base is proposed. Secondly, the choice process of multi-resolution model based on user model is given. The new system based on it can automatically select appropriate model and carry out abstract reasoning and concrete reasoning to meet different requirements of users. Finally, the application example in the intelligent problem-solving expert system verifies the proposed model is feasible.

Keywords: Multi-resolution Model, Personalized Problem Solving, Knowledge Base Modeling, User Information Model, Dynamic Selection Process.

1 Introduction

Expert system (ES)[1] is a program system of automatic solving problem in a certain field of system. For some fields of problem solving, more attention is given to the answer sequence of problem-solving process, such as planning expert system, teaching expert systems[2]. Such systems even more need a personalized user-oriented service[3][4]. So it is necessary to develop the more personalized expert system. For the same question asked by different knowledge levels of users, it can give an appropriate answer with different level of detail.

Resolution is described as the level of detail that a model describes the real world, which is described in detail the number of models. A widely accepted definition [5] is given by U.S. SISO/SIW: "Resolution is the accuracy and detail the model describes the real-world, in the modeling or simulation, also known as granularity." Multi-resolution modeling is defined to establish different degrees of model resolution in the same system or process, and maintain consistency of the main features these models describe. According to a kind of mechanism, the system can select a suitable resolution rate model to meet the system requirements, when it works. Currently researches on multi-resolution modeling focus on distributed simulation system[6][7][8].

Expert system usually deals with complex problems with uncertainties and unstructured, difficult to use algorithms to solve and implement in specialized fields. According to the function in problem solving, the expert knowledge can be divided into three levels: data level and knowledge base level and control level. Data-level knowledge of specific issues is provided by the initial facts and middle-process conclusions, final conclusions during the problem solving process. The knowledge of Knowledge base level is the knowledge of experts, and its quality and quantity determines the level of an expert system performance. The knowledge of control level, also called meta-knowledge, is the knowledge about how to use the first two kinds.

For giving the solution sequence in different levels to different user, expert system not only depends on organization structure of the knowledge base, but also need add user model into the control level. Moreover, expert system can reason hierarchically in abstract and concretion. Multi-resolution modeling of the knowledge base for the expert system provides a hierarchical structure.

The interest of the paper is to combine problem-solving ES with multi-resolution model. To meet the personalized needs of users at different levels, user information model is established and on its basis, the transformation model between multi-resolution modules is proposed and selection process of approximate resolution in the dynamic reasoning process is given. In the follow sections, a multi-level user-oriented knowledge base system of multi-resolution is introduced in detail.

2 Structure of Multi-resolution in Knowledge Base

Knowledge Base(KB)is the core of the expert system. Knowledge representation and knowledge organization affects the efficiency of inference engine's reasoning, as well as enrichment and renewal of knowledge and finally, affects intelligence level of the whole ES[1]. To some extent, with different representation model, the knowledge base should have different organization. Here we organize the system function modules in order. According to the requirements of dynamic object executed, it selects appropriate knowledge base. Each function module establishes a variety of different resolution models.

A piece of knowledge can be expressed as a production rule. And here, we adopt a triple (X, Y, S) where X is a set of condition of production rules; Y is a set of conclusion. S is a set of reliability of rules. The expression of a rule is described as $X_1 \wedge X_2 \wedge \dots \wedge X_n \rightarrow Y_1 \wedge Y_2 \wedge \dots \wedge Y_m$. ($m, n \in \text{integer}$). In resolution model(RM), $X_{\text{low}} \subseteq X_{\text{high}}$ and $Y_{\text{low}} \subseteq Y_{\text{high}}$. The relation set R of X_{low} and X_{high} , Y_{low} and Y_{high} , $X_{\text{low}} \rightarrow Y_{\text{low}}$ and $X_{\text{high}} \rightarrow Y_{\text{high}}$, must be logically analyzed and organized to implement the conversion between abstract reasoning and concrete reasoning.

Fig.1 illustrates a function knowledge module with a symbol A of the multi-resolution hierarchical model. Each module can have multiple levels, and each level can have different types of property. Layer 0 is the lowest level of resolution model(RM) with the most coarse knowledge granular for the most abstract reasoning.

The higher the level and the higher the resolution, the more detailed reasoning process and the more detailed problem solving sequence.

In the structure shown in Fig.1, the maintenance operations in KB are convenient such as adding rules, deleting rules and changing the rules. If the pre-classification of knowledge is rational, these actions only need operate within one or a number of modules. When a new classification occurs, it is easy to select the appropriate resolution model and to add expert knowledge into an appropriate module.

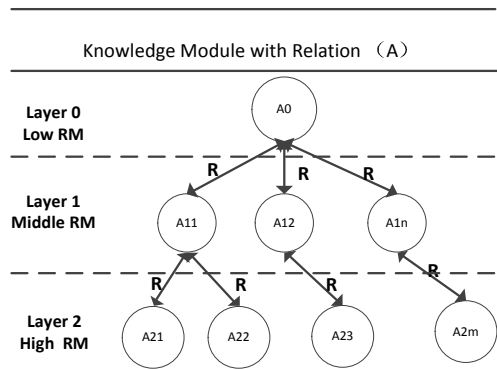


Fig. 1. Structure of Multi-Resolution Model

3 Dynamic Choice of Multi-resolution Model Based on User Model

3.1 User Model

There are several resolution models in KB. But there is only one resolution model during the process of reasoning in the system. So ES needs select the appropriate RM according to user model.

User model is the basic properties of the user and the product related attributes abstracted from the formation of the model. Here we use a message property array (R) to represent the user information. Each user has multiple properties and property values can be specific values or a range of values. The user's knowledge level can be graded according to property values of a user. For a user k , suppose R_k has N values, $R_k = (R_{k1}, R_{k2}, \dots, R_{kN})$ of which, $R_{ki} \in (1, 2, \dots, N_i)$. All R_k ($k = 1, 2, \dots, K$) form the user information space. Users selection set group is expressed as (R_k, P_k) . P_k is a collection RM_k of one or more resolution model selected by the user k based on user mapping model, $RM_k \in (1, 2, \dots, J)$. J depends on the number of existing models used when solving the corresponding problem. So we can get a user's total selection set:

$$L = \{(R_1, RM_1), (R_2, RM_2), \dots, (R_k, RM_k)\}$$

The system guides the choice of RM by Individual user information, then attaining more suitable problem-solving sequence for the user. An example of user model in math intelligent tutoring system illustrates in Table1. 1) Each user has a ID and has basic information such as user type set $P1 = \{\text{teachers, parents and students}\}$, use level $P2 = \{\text{VIP, general user-level users}\}$. 2) Users' math knowledge is classified into a knowledge-level $L64$ and the knowledge level is forward compatible, L said that all knowledge, parents and teachers have this property, of course, this property is optional, a user is not completely fixed. 3) $FUN_1, FUN_2, FUN_3, FUN_4, \dots, FUN_n$ that the system can provide a variety of functions. 1 means users can use this FUN_i and 0 means that the user cannot use.

Table 1. Example of user model information

User ID	User Type	User level	User knowledge structure	FUN1	FUN2	FUN3	FUN4	...
K1	st	general	≤L5	0	1	1	1	...
K2	st	general	≤L10	0	1	1	1	...
K3	TE	VIP	L	1	1	0	0	...
K4	PA	general		1	1	1	1	...

3.2 Selection Process of Multi-resolution Model Depending on User Model

In the dynamic reasoning process, it is necessary to ensure that the same functional module between different resolutions does not conflict. Fig.2 shows a two-tier model for the resolution which is contacted by the module interaction resolver (IR) and consistency enforcer (CE) module .IR Stores the correspondence relation and the reasoning control strategies when a high-resolution model or a low-resolution model is called. If the system determines a need to adopt high(or low)resolution model according to user model which does not match the current resolution, IR runs and deletes(or generates) some process information to ensure right reasoning. CE module stores relations of attributes and information between high-resolution models and low-resolution model. Once the resolution changes, the reasoning process information can be right changed. The resolution of KB is obtained by dynamic selection in the reasoning process .The steps are as follows:

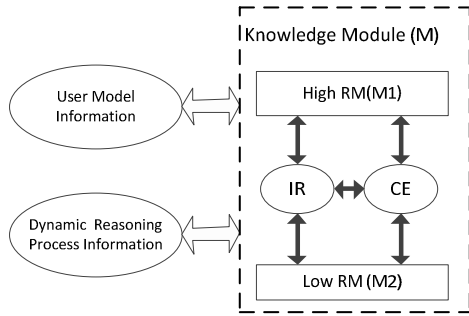


Fig. 2. Reason Model of A Knowledge Module in Multi-Resolution System

Step1: the user inputs issues to deal with.

Step2: according to this user’s database information and user information analysis module, the system automatically ask the user the current state and analyze information through the user interface to form the specific user model information and table records.

Step3: problem-solving expert system successively reads user model information and user input issues to be addressed,.

Step4: according to user type, user-level, functional properties in user model information, determine to select which functional subsystems for processing.

Step5: each subsystem deals with the question in turn. If the subsystem is information search processing, then enter the step6; if the reasoning subsystem is an expert system, then skip to step 7. If all the processing subsystems is over, then skip to step 8.

Step6: search the database for information to submit the feedback. Skip to step5 and enter the next subsystem.

Step7: reason in the low resolution model on the user question. during the reasoning process, when the resolution is inconsistent with the resolution according to the associated matrix based on the level of the user knowledge and information, record the middle reasoning information. After low resolution inference is complete, acquire the appropriate information and CE information on IR according to record information, then convert between the low resolution information and high-resolution information. The output of problem-solving sequence of subsystem is brought into the next subsystem. Skip to step5.

Step 8: the solutions feedback to the user through the user interface.

4 An Application Example in an Intelligent Tutoring System

Intelligent Automatic Problem Solving Agent—the core of Remote Intelligent Tutoring System [9]— is an expert system and developed based on knowledge-based automatic reasoning technology. It can automatically solve the math problem proposed by students and can simulate the teacher tutoring the students. Abundant knowledge about the mathematical axioms, theorems, definitions and so on are stored in the reason knowledge base. Its service object is students of different grades.

Students in various stages of knowledge learn different contents. The higher their grade, the more knowledge they have mastered. Therefore,

the solution to the same mathematical problem is different in the use of mathematical methods, the axioms, theorems, definitions and properties. The system adopting the knowledge base system architecture based on multi-resolution model above to implement different answers to the question of the students in different stages of learning. Through trying out tentatively, correct rate is 100% and rationality rate is 90%. It verified the proposed model is feasible.

Fig.3 is an example that shows the transformation between low resolution model and high resolution model in an automatic problem-solving geometric reasoning system. During the process of de-aggregation an aggregation, the information of lines, angles, points and others changes in the middle of reasoning processing. The static connotation relationship R between the two model is explicit and is stored in the IR. So it is easy to determine what are changed and how to change.

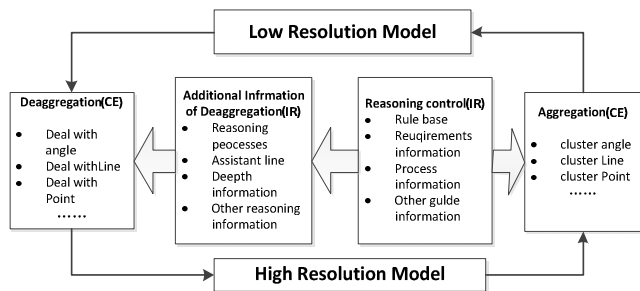


Fig. 3. Example of Transformation Structure in Automatic Problem Solving System

5 Conclusion and Further Research

From the different resolutions and different point of view, analysis and dealing with the problem is an important way of thinking of mankind, as well as an effective means to deal with complex issues. High-resolution model can grasp the details, and low-resolution model can better reveal the macro things, and the nature of the property. Besides, research and development of the system must take full account of the user's own characteristics. Such system can achieve a personalized user-oriented service. In this paper, the structure of KB expert system based on multi-resolution model is proposed for multi-level users. The hierarchy of multi-resolution model is given to represent inner hierarchy of a knowledge module in KB. Using dynamic user information, the system can guide the selection of multi-resolution model of knowledge base during the reasoning process to meet the different requirements of users. The model present in this paper has applied to a plane-geometry intelligent tutoring system for junior school students with different grades or scores to provide multi-level tutoring information, and the result got broad recognition. Further research on the model applied to other fields is required.

References

1. Yang, X., Zhu, D.-Q., Sang, Q.B.: Research and Prospect of Expert Systems. *China Journal of Application Research of Computers* 24(5), 4–9 (2007)
2. Wo, B., Ma, Y.: *Expert System*. University of Beijing Technology and Industry, Beijing (2001)
3. Zeng, C., Xing, C.-X., Zhou, L.: 'A survey of Personalization Technology. *Chinese Journal of Software* 13(10), 1952–1962 (2002)
4. IBM high volume web site group. Web site personalization[EB/OL], <http://www-900.ibm.com/websphere/hvws/personalize.shtml> 2000-01/2004-10
5. Pace, D.K.: Dimensions and Attributes of Simulation Fidelity. In: *Proceedings of 1998 Fall SIW* (1998)
6. McGraw, R., Treshansky, A.: MRMAideTM: A Technology for Implementing Mixed Resolution Models. In: *Proceedings of the 2002 Fall SIW* (2002)
7. Cat, K., Gardner, K., Gryder, T.: Military Modeling Framework (MMF): A Framework Supporting Composability and Multi-Resolution Modeling. In: *Proceedings of the 1998 Fall SIW* (1998)
8. Gottschalk, T.D., Davis, D.M.: Application of Proven Parallel Programming Algorithmic Design to the Aggregation/De-aggregation Problem. In: *Intersevice/Industry Training, Simulation & Education Conference* (2006)
9. Yi, Z., Zhao, K., Li, Y., Cheng, P.: Remote Intelligent Tutoring System Based on Multi-Agent. In: *Proceedings of The 2nd International Conference on Information Engineering and Computer Science*, Wuhan, China, vol. 3, pp. 958–961 (2010)

The Construction of the Security Ability System in Coal Enterprise

Liu-an Kong¹ and Wen-jing Wu²

¹ Henan University of Urban Construction, Pingdingshan, Henan, 467000

² Henan Polytechnic University, Jiaozuo, Henan, 454000

Abstract. Based on entrepreneurship theory and modern safety management theory, the system model of security capacity in coal enterprise was established according to system theory, and furthermore its index system was determined. Taking characteristic vectors of paring comparison matrices as the weights of every factor, the influence sequence of the assessment factors was determined. Results showed that: the coal enterprise safety ability in China mostly lies on human factors, followed by technique, management, and then environmental factors. According to the evaluation and analysis results, in order to improve the safety of coal enterprise ability, some effective actions was put forward.

Keywords: Coal enterprise, System of security capacity, Analytic hierarchy, Evaluation model.

1 Introduction

Coal resources has always been our country's primary energy, therefore, the coal company security issue has become a top priority of national security production[1]. From the angle of analysis and study, how to improve the safety of coal enterprise ability level and the actual operation needs a "security capability evaluation index system". The coal enterprise safety production capacity is technology, management, and the environment and other factors. The coal enterprise security ability of the decision factors and human factors more complicated, and the correct understanding of various factors on the influence of coal enterprise security ability is to carry out good safety management and safety evaluation of the foundation. This paper put forward the ability to build coal enterprise security systems conceptual framework, and use analytic hierarchy process of coal enterprise security ability on the ability of a coal enterprise security analysis, and strive for the safety assessment of the coal business and management decisions.

2 The Security Capabilities in Coal Enterprise

2.1 The Entrepreneurship Theory

Entrepreneurship theory is a summary for a series of articles and books having a common research concern which are still differential, and they all agree that "the primary important significance of their over all strategy is the company-owned special

asset, namely: knowledge-related, can not see or feel it, but can let you feel, difficult to trade and in the enterprise internal departments divisible assets"[2]. Because enterprise ability theory is a summary of the numerous works, it has many branches, and among them the most important one is "enterprise resource-based theory" represented by Werner Felt in 1980s [3] and core competence theory" represented by Prahalad and Hamelas in 1990s [4]. The core competence theory can be seen as the latest developments in the theory of enterprise capabilities that the organization's cumulative knowledge constitutes the company's core capabilities (or core competencies), clearly put forward the nature of the enterprise capabilities of knowledge. The enterprise ability theory mainly concern about the nature of the firm, boundary problem and the competitive advantages of enterprises. The solution of these problems can provide the rule and principles for enterprise strategy decision making.

2.2 Safety Management Theory

Now the study of the theory of the safety management has been more comprehensive and systematic. The early safety management being equalized to accident management has been transited to a modern management of hidden trouble. System of modern management theory is one of the most basic principles, and it is to show people engaged in management work how to make use of systematic theory, viewpoints, methods and full system analysis to achieve the goal, which is the optimal management in the system view, theory and methods to understand and deal with the problems in management.

Through the above enterprise ability and safety management summary of the knowledge, the author holds that, enterprise security ability is the power system, is the enterprise in certain safety culture atmosphere; use all sorts of safety knowledge and resources, through the reasonable organization, coordination, and control, in the form of a kind of real or potential reaction. The foundation of improving production safety is to improve the security capabilities.

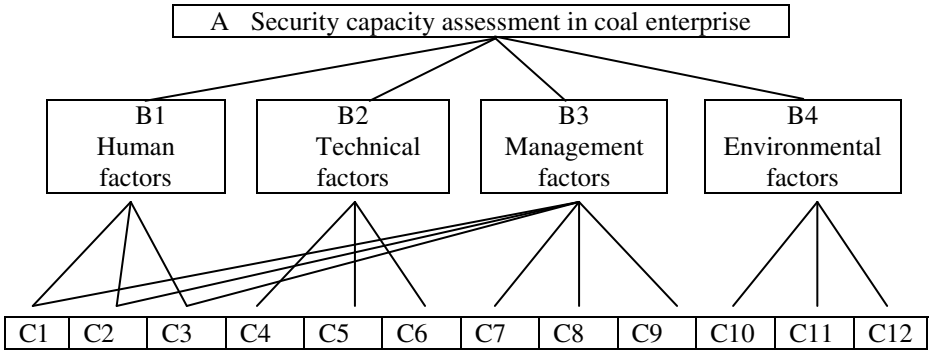
3 System Model of Security Ability in Coal Enterprise

3.1 Clear the Problem, Determine the Evaluation Index

The index system should be able to reflect the basic conditions of coal enterprise and the main security features, to reflect the security of the system has the ability for the target. China's coal production is mainly underground, because of complex geological conditions of coal occurrence, often by the gas, water, fire, dust, roof and other natural disasters, threats, combined with the majority of coal production, backward technology, poor quality of personnel, security loopholes in supervision and management, resulting in weak safety [5]. From the synthetic analysis of coal enterprise security capability from the perspectives of human, machine and environment, we can see that the deciding and influence factors of coal enterprise security ability mainly include the human factors, technological factors, management and environmental factors that affect four aspects [6]. The human factor refers to their basic quality, safety awareness and safety behavior; Technical factors is to point to the coal enterprise production technology and safety technical level; Management factor is the safety management system construction and state of key indicators; Environmental factors is mainly refers to the occurrence of coal geology condition[7].

3.2 The Establishment of a Hierarchical Model

According to the basic steps of AHP, the establishment of coal capacity assessment level of enterprise security model, shown in Figure 1.



Note: C1-Level of cultural quality. C2-Skill level. C3-Safety education and training. C4-Mining process. C5-Safety equipment. C6-Information technology. C7-Safety procedures. C8-Regulation. C9-Safety supervision strength. C10-Gas content. C11-fire, water and coal dust. C12-Roof and rock mass

Fig. 1. Security capabilities assessment hierarchy model in coal enterprise

3.3 Tectonic Judgment Matrix and Consistency Examination

Using 1-9 scale method (see Table 1) for pair-wise comparison, and reference expert opinions, to determine the relative importance of each factor and to give the corresponding value, to construct the hierarchy of all the matrix, the right solution vector and the consistency test.

Table 1. Meaning of judgment matrix 1-9 scale

Serial number	Importance level	Cij assignment
1	i, j two elements are equally important	1
2	i element than j element slightly important	3
3	i element obvious important than j element	5
4	i element than j element strong important	7
5	i element than j element extremely important	9
6	i element a bit not important than j element	1/3
7	i element obvious not important than j element	1/5
8	i element strong not important than j element	1/7
9	ielement extremely not important than j element	1/9

Note: Cij assigned 2,4,6,8,1/2,1/4,1/6,1/8 to determine the level of two middle values.

A-B to determine the composition between the matrix form as follows:

$$A = \begin{pmatrix} 1 & 2 & 3 & 2 \\ 1/2 & 1 & 2 & 1 \\ 1/3 & 1/2 & 1 & 1/2 \\ 1/2 & 1 & 2 & 1 \end{pmatrix}$$

Find eigenvalues : $W=(\omega_1, \omega_2, \omega_3, \omega_4)= (0.424,0.227,0.122,0.227)$, $\lambda_{\max}=4.0104$.
 $CI= (\lambda_{\max}-n) / (n-1)=0.003467$, Look up table has $RI=0.90$, $CR=CI/RI=0.003852<0.1$,
 Therefore, the consistency of the comparison matrix is acceptable.

B-C to determine the composition between the various forms are as follows:

$$B1 = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 2 & 1 & 2/3 \\ 3 & 3/2 & 1 \end{pmatrix}$$

The characteristic value of the matrix for: $W=(\omega_1, \omega_2, \omega_3)= (0.167,0.333,0.500)$,
 $\lambda_{\max}=3$. $CI= (\lambda_{\max}-n) / (n-1)=(3-3)/(3-1)=0$, $RI=0.58$, $CR=CI/RI=0<0.1$, So this
 judgment matrix consistency is acceptable.

$$B2 = \begin{pmatrix} 1 & 2/3 & 2 \\ 3/2 & 1 & 3 \\ 1/2 & 1/3 & 1 \end{pmatrix}$$

The characteristic value of the matrix for: $W=(\omega_1, \omega_2, \omega_3)= (0.333,0.500,0.1667)$,
 $\lambda_{\max}=3$. $CI= (\lambda_{\max}-n) / (n-1)=(3-3)/(3-1)=0$, $RI=0.58$, $CR=CI/RI=0<0.1$, So this
 judgment matrix consistency is acceptable.

$$B3 = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/3 & 1/2 & 1/3 \\ 2 & 1 & 2/3 & 1 & 2 & 1/2 \\ 3 & 3/2 & 1 & 2 & 3 & 1 \\ 3 & 1 & 1/2 & 1 & 2 & 2/3 \\ 2 & 1/2 & 1/3 & 1/2 & 1 & 1/2 \\ 3 & 2 & 1 & 3/2 & 2 & 1 \end{pmatrix}$$

The characteristic value of the matrix for: $W=(\omega_1, \omega_2, \dots, \omega_6)= (0.069,0.158,0.261,$
 $0.169,0.100,0.243)$, $\lambda_{\max}=6.0906$. $CI=0.01812$, $RI=1.24$, $CR= 0.014613<0.1$, So
 this judgment matrix consistency is acceptable.

$$B4 = \begin{pmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 2/3 \\ 1/3 & 3/2 & 1 \end{pmatrix}$$

The characteristic value of the matrix for: $W=(\omega_1, \omega_2, \omega_3)= (0.550,0.210,0.240)$,
 $\lambda_{\max}=3.0735$. $CI= (\lambda_{\max}-n) / (n-1)=0.03675$, $RI=0.58$, $CR=CI/RI=0.063362<0.1$, So
 this judgment matrix consistency is acceptable.

It can be seen, all single-sorted $CR < 0.1$, that the consistency of each comparison matrix is acceptable.

3.4 The Total Level of Sorting and Consistency Check

The above is a set of elements to get a layer on an element in its weight vector. To get the lowest element, especially in the sort of programs for the target weight, the total order is needed. Important top-down sort of total power under the weight of a single criterion for synthesis.

Influence of coal enterprise security all the factors are always ability sort. As shown in table 2.

Table 2. Evaluation of coal enterprise security overall sorting capabilities

C layer \ B layer and the weights	Human factors	Technical factors	Management factors	Environmental factors	C-level factors, the weight of the total order
		0.424	0.227	0.227	0.122
Cultural quality level	0.167	0.000	0.069	0.000	0.0865
Skill level	0.333	0.000	0.158	0.000	0.1771
Safety education and training	0.500	0.000	0.261	0.000	0.2712
Mining process	0.000	0.333	0.000	0.000	0.0756
Safety equipment	0.000	0.500	0.000	0.000	0.1135
Information technology	0.000	0.167	0.000	0.000	0.0379
Safety procedures	0.000	0.000	0.169	0.000	0.0384
Regulations	0.000	0.000	0.100	0.000	0.0227
Safety supervision strength	0.000	0.000	0.243	0.000	0.0552
Gas content	0.000	0.000	0.000	0.550	0.0671
Fire, water and coal-dust	0.000	0.000	0.000	0.210	0.0256
Roof and rock mass	0.000	0.000	0.000	0.240	0.0293

Hierarchy total sort to the consistency of inspection, from high to low test, but also to have the latest study, in the AHP method is no test the consistency of the hierarchy total sort. That is, in practice, the total ranking consistency test can often be omitted [8].

3.5 The Analysis of the Result

Based on the above analysis results, the human factor is the primary factor to affect the safety of China's coal enterprises, followed by technical factors and management

factors, and environmental factors are ranked in the final. Visible technology and environment factors and not the most important, the person's factor is the decision of the key factors of coal enterprise security.

4 Conclusions and Suggestions

Through the above analysis, the human factor is the ability of coal enterprise security system is the most important factor. The coal enterprise security staff are the subject of the safety of the employees, cultural quality, safety concept, security behavior is inseparable from the people factors. Accordingly, the first thing we should set up is the idea of "people oriented", and makes the people as the main body of the security system of coal enterprise. Develop in line with the actual work of the staff code of conduct and operational safety procedures, continue to strengthen the constraints on employee safety behavior, and gradually achieve the institutionalization of management to employees of self-discipline to managing change and build "people-oriented" culture of safety systems. In established "people-oriented" security system, we should also strengthen the coal enterprise employee training work, and improve the overall quality of the staff.

Through the implementation of the national production safety laws and regulations, policies, and analysis, research and evaluation coal enterprise in production and construction of the existence of the various safety factor, from personnel, organization, technology, management and training, and many other aspects to take reasonable and effective measures to establish a coal enterprise security ability system. Through this system, we can eliminate and control the production of coal existing in the process of construction hazards, prevent accidents or coal enterprises to maximize the scope and extent of control the effects of the accident, the coal companies to achieve optimal security status for the building of coal production protection.

Acknowledgement. PhD Research Fund of Henan University of Urban Construction.

References

1. Wang, S.: Security practice and reflection of the development. Coal industry press, Beijing (2010)
2. Fox, N.J., Knudsen, C.: Code section Anchorage, Li Donghong translation, Universal business - business-oriented capacity theory. Northeast financial university press (1998)
3. Wernerfelt, B.: A Resource—based View of the Firm. Strategic Management Journal (1984)
4. Prahalad, C.K., et al.: The Core Competence of the Corporation. Harvard Business Review (1990)
5. Zhang, C., et al.: Factors on one million tons of coal mine the mortality of regression analysis and its application. Chinese Production Safety Science and Technology (January 2005)
6. Jing, Q., et al.: Based on the analytic hierarchy process (AHP) of coal mine safety production ability index system research
7. State Administration of Production Safety Supervision and Management. "Mine safety" Eleventh Five-Year "plan"
8. Du, D., et al.: A modern comprehensive evaluation method and case selection. Tsinghua university press (June 2008)
9. Satty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)

Discussion on Certain Problems and Countermeasures on Expansion of Management Advisory Services of Accounting Company in China

Chenyun Ye and Lingyan Kou

School of Economics and Management,
North China Electric Power University
No.2 Beinong Rd., Huilongguan, Beijing, 102206, China

Abstract. Under our national condition, as the economic opening degree of our country is strengthened day by day after accession to WTO, business competition of Chinese's enterprises becomes fiercer and fiercer, the management advisory demand increasing sharply from society. With the view of the west country, it is development trend of accounting business that the management advisory services enlarges rapidly to meet the needs of enterprises' competition. The main contents of this paper is that author analyses some backgrounds and problems exist in management consultation services of accounting company in China firstly, and then author argue some countermeasures of improving difficult condition of management consultation services of accounting company in China. So that our China's accounting company can fully utilize the internal resources, improve the audit services quality, reduce the risk of audit services, coincide with the accounting company's development tide.

Keywords: Accounting companies, Management Advisory services, Problems and Countermeasures.

1 Introduction

In the course of globalization competition of the 21st century, management consultation plays an important drive role in improving western enterprise's management level and competitiveness, has become one of the fastest developing services industries in advanced western countries too.

What is called management consultation, according to explanation from a survey report of the United Nations about management consultation in 1993, it is a new-type trade that involves very wide scope, it permeates through every aspect in economic management life day by day. Domestic scholars generally believe the management consultation services that the accountant is engaged in belongs to operation of accounting company's consultation, the important content of the services business. It is a kind of consultation services activity that the accounting company accepts enterprise's commission, uses the scientific method, diagnoses the problems existing in administration and management of enterprises on the basis of investigation and analysis, proposes improving measures and instructs its implementation, helping enterprises to

improve administration and management in order to promoting enterprises' economic efficiency. And according to present practice in foreign countries, management advisory services that accountants undertake includes enterprise advisory work and financial consultation business mainly. The former includes concretely the long-term strategy programming, organization frame, stock-right design, cost management, achievement management, information technological consultation, employee salary evaluation services etc., the latter includes global company's financing, enterprise annexation, purchase, stripping services, business reorganization, assets assessment, financial variety consultation, etc.

With the progress of society and economy in the world, the enhancement of business and Industry administration demand and people's deepening of this understanding in near future, the intension of management consultation services is being expanded abundantly, increasingly and constantly. Thus, how to deal with the expansion of management consultation services of the Accounting Company has become a very important a question for discussion that our accountants and scholars should be confront with it earnestly.

2 The Background of Accounting Company's Developing Management Advisory Services in China

With the rapid development of socialism market economy in China recent years, more and more enterprises participates in domestic and international market competition, moreover, as the economy opening degree of our country is strengthened continuously after the accession to WTO, enterprise competition is fiercer and fiercer, in the meaning time, demand of the management consultation from the society increases sharply. The recently statistics shows that the business of management consultation in our country has kept the increase rate of more than 50% all the time since 1999, we can see that management consultation in our country has already demonstrated the booming development situation, it makes people gratified. Meanwhile, the existing problems also have made people very anxious, within the domestic market, the foreign advisory companies have got hold of advantage position, the domestic market development of management consultation is still insufficient, and its whole still retains relatively low-level, there is no ability to contend with management advisory corporations of western countries at present, especially after joining WTO, this situation expose more outstandingly. For the accounting occupation of our country that gets involved in this field is almost blank, it is very disproportional to the whole accountant development level, and causes the social resources wasted badly. On the one hand, because the audit demand in the accounting profession is seriously insufficient in our country, it makes the trade competition become fiercer day by day, on the other hand, enterprise management consultation has great demand, and accounting professions that need talents, business and organizational advantage terribly have slow reaction in this respect, falling into such predicament cannot but provide food for thought and hold our own wrists.

3 The Certainty and Feasibility That Accounting Companies Are Engaged in Management Advisory Services

Seen from the west, the management advisory services increasing rapidly is to meet the needs of enterprise competition and accounting business development trend. First of all, the quick increase in management advisory services is an inevitable result that the western business management theory develops, which caused by this the structure inside enterprises is recombined. Secondly, with the indetermination enhancement of the management environment, such as the technological change accelerating, intensification of the internationalization degree of economy, product cycle being shortened, assets securitization trend strengthening, and computer crimes increase with more concealment comes from electronic trade popularization and information technology using extensively in administration and management, and a series of factors even more, enterprises face the greater experience risk. Under this situation, enterprises pay attention to the risk management of experience day by day, and will inevitably rely on the exterior management consultant experts, thus promote the development of the management consultation services. Thirdly, as audit receives severe challenge of information society again, high lawsuit risk not descending impels the business of auditing to expand difficultly, and by contrast, the managerial consultation belongs to the value-added services, it helps accounting company make a generous profit with its risk being relatively little, and develop rapidly. From the fifties of the 20th century, the management consulting services income of the western accounting profession began to increase, and have kept the powerful growth. Possessing the advantage as weather, favorable terrain, people also lets its rapid development be more inevitable. Under the present situation, it is a rational choice for accounting company to embark on managerial consultation services. Only the accounting company catches this opportunity can it be promoted to development rapidly.

The feasibility that accounting companies are engaged in managerial advisory services mainly reflects as follows.

A. Legal basis. According to article 15 in chapter three "business scope and rule" of "Law on Certified public accountants of the People's Republic of China" that really gives the accountants this right, namely stipulate that the accounting company can undertake accounting consultation, accounting services, here the accounting consultation and the accounting services item embody the management advisory services.

B. Feasibility on time. As noted previously, it is the timeliness business that the accounting company audits, most of the time, the resource is in the state of leaving unused in the accounting company, so the certified accountant has abundant time to improve the management advisory services for customers.

C. The feasibility on business operation. On the one hand, according to the certified accountant's own quality, he is a specialized personnel with higher professional level and abundant experience, possessing basic ability that should have to offer management consulting services for enterprises; on the other hand, from the general character between management advisory services and routine audit business of

certified accountants, the certified accountants totally has the ability to offer high-quality, high-efficient management consulting services too.

D. Feasibility on practice. In our country, some accounting Company's business at present already has begun to wade into the services field of managerial consultation. Though the scale and range are relatively small, it offers the beneficial try for the accounting company in our country to launch the management advisory services. Comparatively popular method is that the accounting companies amalgamates with management advisory corporations or unite in order to offer the pluralistic services at present, this method is suitable for some large-scale accounting companies mainly. For medium and small-scale accounting companies, it may be a wise choice for them to utilize their limited resource, develop the management advisory services by themselves, and offer specialization services.

4 The Analysis of Certain Difficulty or Problem That Accounting Companies Development Management Advisory Services in China

For accounting companies in our country, it is not only of great significance to launch the of management advisory services as stated, but also it is exist much feasibility on the realistic aspect, so it has already been the trend of times to expand the managerial consultation services. However, in the course of opening up the of management advisory services, the accounting companies in our country may meet many new problems as follows :

4.1 The Idea of Accounting Companies Lags Behind in Expanding Other Business

Most accounting companies in our country are only limited to auditing services while offering the business, and stagnate in offering other business. On the one hand, it relates to the fact that the accounting company forms the background at present in our country. The accounting companies produce through being cultivated by government all alone, it is the routine business for them to carry out the auditing business for customers and their own. Because of their non-marketization of original background with the accountant's firms or limit companies, it caused them non-marketlization in the course of development. Though the accountant's firms or limit companies have performed the auditing business for many years, but they still do not have a natural market-based relation of supply and demand with the customers.

4.2 Dull Demand Services Object from Accounting Trade in Our Country

Government department is the most demand person for the accounting profession serve, but the administrative authority of most companies has no inherent motive force to produce the demand for the accounting services yet. Because our accounting companies mainly launch legal business, such as auditing, capital verification, etc.

Concerning legal business, the capital verification, industrial and commercial annual check serve business and administration department; the annual check of foreign currency serves the department of foreign exchange control; the accounting company's report audit of state-owned enterprises serves the responsible department of enterprises; the customer demand for strengthening management consultation for the accounting company is insufficient.

4.3 Accounting Company Haven't Yet Provided Effective and Appropriate Services for Various Enterprises

Comparing with some small-scale and not standardization advisory companies, our accounting company has certain advantages to launch the of management advisory services in such aspects as personnel, organization and capital, but it still has certain disparity in need for offering abundant and effective management consulting services. Because of short time establishment with our accounting company, an inaccurate orientation and narrow services objects, causing it to rely mainly on the auditing business, the experience is still not abundant in other services. On personnel's structure, all is the accounting company and audit respect personnel, only this is far from enough, and the actual need is to offer such experts of various fields as extensive advisory services, tax, law, insurance, administration and management, etc. On organizational framework, mostly it is in the forms of limited responsibility for our accounting companies, although it can disperse the risk, but will cause decision-making slowly. Moreover as a result of special and changeable characters of the management advisory services, we often need loose and flexible organizational forms even more, and high-efficient and rich pertinence in the decision-making.

4.4 Our Accounting Company Haven't Short of Enough Professional Reputation Answer for Challenge of Management Consultation Services

In the course of their development, our accounting companies have not set up good job reputation yet, have not formed the brand advantage either, they are uncompetitive in offering management consulting services. Our accounting trade is not a market-based result, but a subsidiary body of government department in a long period all along, has not really incorporated the market, so for a long time, it has not established the good job image in the public mind, and can not form its own brand in the market competition either, thus it has no advantage in offering management consulting services.

4.5 The Domestic Accounting Company Have Few Advanced Management Techniques at Present

The accounting company launches less national exchange and cooperation, has not fully studied and used the experience that the foreign accounting company has been engaged in the management advisory services for reference. Our accounting trade seldom exchanges with international accounting company in offering management consulting services at present, the five international accounting companies have already set up their branches in some cities of our country, is it more, on account of little

intercommunion, it causes our accounting company scarcely understand the advanced knowledge in foreign countries in this respect, this is unfavorable to grow our accounting company.

5 On Some Pertinent Countermeasures of Accounting Companies Develop Management Advisory Services in China

Difficult and problems that our accounting company facing is engaged in the management advisory services, the writer of this article proposes adopting the following countermeasures:

First of all, Accounting company should change idea, base on market, understand customer inherent need really, offer added-value services of going beyond its appreciation, and regard expanding the of management advisory services as its important strategy. *With the constant perfection of the governance structure of state-own enterprise corporation and structure change of quoted company*, especially after joining WTO, enterprises of our country move towards standardization gradually. Our accounting company should change its backward idea too, participate in the market competition actively, advance its own marketization process, consolidate the cooperation with the customer, thus can occupy the advantage position in the keen competition of services field of management consultation.

Secondly, The accounting company should improve talent and institutional framework vigorously, improve on the management quality level of advisory services. In composition of trained personnel, we can introduce the professional personnel of other trades, pay attention to the integration among all kinds of specialized personnel. Through many modes of uniting, cooperating, annexing or purchasing etc. with other professional organizations, the university and scientific research institutions, utilize the specialized personnel of various fields to improve the services level of management consultation; strengthen training former employees to promote the management advisory services ability of the accountants. In organizational framework, the accounting company should adopt the flexible organizational forms such as partnership, limited company, even individual proprietorship, etc. personal, in order to meet the needs of management advisory services. We can fully utilize various kinds of resources via uniting, cooperating, acquisition and merger, recombination, etc. with other organizations ; can set up the services departments of special management consultation or various kinds of business groups inside the organization; set up the information resource bank where include the customer management file, so as to offer management consulting services abundantly and deeply.

Thirdly, The accounting company should launch the marketing tactics energetically to the customer, to establish the advantage of brand. The present accounting company has not formed its brand advantage yet, it relates to accounting company's own prestige, the important reason lies in the accounting company's marketing tactics not doing its best. So in order to strengthen the public knowledge of itself, and to launch the extensive management advisory services, it has already been very urgent choice that the accounting company should launch the marketing tactics vigorously to the general

public. We should pay attention to adopting the positive marketing tactics in the course of concrete services, meet the demand of customers to the maximum extent, thus to strengthen the competitiveness in launching the management advisory services of accounting company.

6 Conclusion

With the development of China capital markets and the modern enterprise system, coupled with the impact of current financial crisis, the contradictions and problems of the governance structure of accounting companies in China will inevitably continue to show sudden, if we do not speed up treatment, it will be seriously hit and impact on future reforms, norms and healthy development for China audit market system. As has been mention above, some conclusion can be drawn are as following:

(1) Accounting companies should take effective measures to copy with balance of market Competition and management consultation services quality, in doing so, our accounting companies in China can win more higher ration parts of the worldwide auditing and management servicess market in near future.

(2) Every certified public accountants in variety districts and cities are improve quality of personal ethics and talent or levels of professional auditing or management without delay through all means of methods as soon as possible.

(3) The association of certified public accountants in China should stipulate for the criterion of management advisory services, the related departments should promote standardization toward management advisory business, work along both lines to improve the quality of management advisory services of accounting company, for facilitating normal development of the services trade of whole management consultation healthily. The association of certified public accountants can make the behavioral norm of management advisory services, allow accounting personnel to have regulations to abide by, can also weigh its operation quality effectively, to make accounting company unlikely to damage independence and job image during launching management advisory services, and the market supervision department strengthens and standardizes the services industry of management consultation through making the laws and regulations, and it helps to improve the present situation that fish and dragons mix up together at services market of management consultation, and create conditions that the accounting company involves in the management consulting trade, for the sake of launching the fair competition.

Acknowledgements. This work is partially supported by Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality (PHR20100512). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

1. Xu, R., Li, X.: Carrying Out the Deep Change of the Chinese Accounting System—The Contemporary Accounting Investigation On the Frontier Problems. The version of 2001 on Economic Publisher of Chinese Public Finance (2001)
2. Zhao D.: The Analysis On Finance Economic Behavior and Efficiency. The printed draft of Ph.D. degree thesis, in The Southwest University of Finance and Economics (1998)
3. Spender, J.C.: Some Frontier Activities around Strategy Theorizing. *Journal of Management Studies* 28(3), 67–69 (2000)
4. Pike, R., Dobbins, R.: Investment Decisions and Financial Strategy, pp. 107–109. Philip Allan Publishers Ltd. (2003)
5. Kaplan, R.S., Norton, D.P.: Using the Balanced Scorecard as a Strategic Management System. *Harvard Business Review*, 122–124 (1996)
6. Kapla, R.S., Norton, D.P.: The balanced scorecard: Translating strategy into action, pp. 56–57. Harvard Business School Press (2003)
7. Michael, E.P.: Competitive Strategy Techniques for Analyzing Industries and Competitors, pp. 36–40. Harvard Business School Press (2004)

Research on Control System of Inspecting and Adjusting Off Pattern for Rotary Screen Printing Machine

Aiqin Sun, Jidai Wang, and Dongyue Zhang

College of Mechanical and Electronic Engineering,
Shandong University of Science and Technology, Qingdao, China
saq80@sdust.edu.cn, djdwang8911@sina.com,
jinyu1987829@163.com

Abstract. Based on the problem that the control system of rotary screen printing machine can not on-line inspect and real-timely adjust off pattern in the process of printing, an automatic inspecting registration system is designed by images of pattern position marks on both sides of fabric in wax dying. The images of pattern position marks before printing are captured by a CCD image sensor, and the real distance about the current pattern and the error between the measured value and the standard value are calculated, which are used to estimate the registration precision and produce the control pulse for regulating the rolling speed. After printing, the effect of dynamic adaptation is given by another CCD image sensor and according to the feedback the secondary adjustment is made. In this way, the closed-loop registration system for rotary screen printing machine is realized that will raise the quality and the productivity of printing.

Keywords: Automatic registration, off pattern, rotary screen printing machine, image sensor.

1 Introduction

The wrong pattern is unavoidable in the process of printing with rotary screen printing machine because of the wear and the loosening of mechanical parts or the deformation of fabric structure [1-3]. The quality and the productivity of the printing for the rotary screen printing machine would be seriously affected if wrong pattern was not adjusted in time. So the off pattern regulation is very critical fact to the final printing accuracy of rotary screen printing machine. Traditional inspection method by the artificial eye is commonly applied on the printing, which is hard for workers and reduces the accuracy of the printing, so it is very difficult for improving the qualification in the conditions of intricate pattern or high-speed printing. With the continuous development and advancement of the computer control technique, many kinds of rotary screen printing machines with automatic inspecting registration control system were shown in the last ten years [4-7]. In the present automation inspecting registration systems, the registering error is mainly inspected and adjusted by the special man-made pattern mark or zero position label on the rotary screen rollers or on the roller driving motors. The off pattern error is indirectly acquired according to the distance between the motor

labels detected through photoelectric encoder or the deviation of the marks printed on fabric with image sensors [6], so control signal for adjustment is generated. Recently, a new method of directly inspecting the printing pattern is applied in a new printing registration system named the “Hawk Vision-WPOA” wax [7], developed by Changzhou Hongda Electric Co., Ltd. The registering accuracy is highly improved with this new system. However, the feedback part of adjustment is not considered in these registering systems, which require manual intervention in many cases, so they are insufficient for the enterprise actual demand.

A novel automatic inspecting registration method based on the images of pattern position marks on both sides of the fabric in wax dyeing is introduced in the paper. This new method has advantages of not only on-line prediction and adjustment of off pattern in printing, but also compensation by real-time detecting the effect of adjustment with adopting closed-loop feedback control.

2 Principle of Off Pattern Measurement

Some marks, just like letters or other patterns, are printed at regular intervals on both sides of the fabric during wax printing. The distance between the two marks printed at regular interval can be measured and named one printing cycle. The mark of the printing cycle is used as the actual position of ground pattern in this system. The registration precision of rotary screen printing is measured and controlled by directly capturing the images of these marks of print cycles.

The process of the measurement and control for the rotary screen printing is shown in Fig. 1.

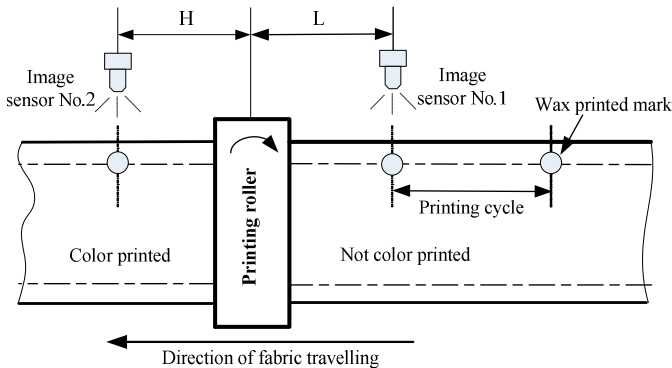


Fig. 1. Layout of image sensors

As shown in Fig. 1, two image sensors named No.1 and No.2 respectively are used in the system. The No.1 sensor is installed at a printing cycle (assumed a printing cycle as L) ahead of the printing roller, and the No.2 at H behind the printing roller.

The No.1 sensor is for acquiring the dynamic error of the actual printing cycle to estimate the off pattern value and adjust the speed of the printing rollers, and the No.2 sensor is for acquiring the effect of dynamic adjustment to compensate next time. Both image sensors are triggered according to the numbers of output pulses from the rotary encoder placed on the main motor shaft, and the trigger frequency should be higher than that of capturing all the marks.

3 The Deviation Calculation of Printing Cycle

A pattern with clear geometry shape and obvious recognized feature should be chosen as the standard mark from all the figures on both sides of the fabric. As shown in Fig. 2, the solid dot is located in the center of the field of CCD camera's field as ideal position of the mark. The hollow dot is marked actual position in the camera's field. The difference between the actual mark position and standard position is denoted as Δ . When the actual position is just at the ideal center, $\Delta=0$; when it is ahead of the center, Δ is positive; when it is behind the center, Δ is negative. The last dynamic error of the mark in the field obtained by No.1 image sensor is assumed as Δ_0 and the current error is assumed as Δ_1 . Suppose that the output pulse number of the encoder is N in two successive capturing the target mark and the distance that the fabric travels forward per pulse is δ . Then the deviation between the value of the printing cycle just going into color printing and the value of the standard is as follow:

$$\Delta L = N \cdot \delta + \Delta_0 - \Delta_1 - L \quad (1)$$

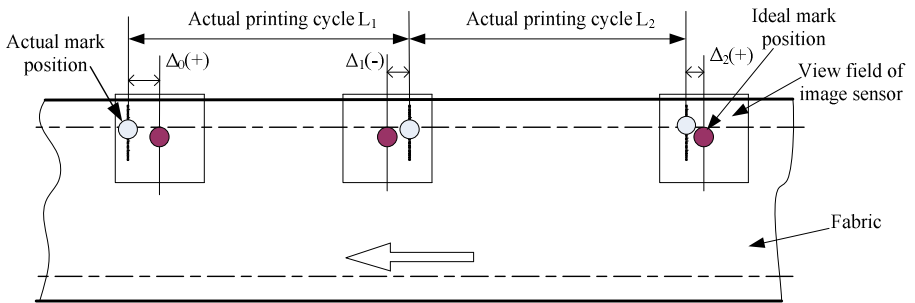


Fig. 2. Schematic diagram for calculation of actual printing cycle

It is supposed that the deviation of the actual position measured from the No.2 sensor is named as ΔH , and the trigger number of output pluses from the encoder is named as M . The error (assumed as ΔQ) between the actual distance from the mark to the printing roller and the standard distance can be easily calculated according to the pulse number (M) and the deviation (ΔH) captured by the No.2 sensor:

$$\Delta Q = M \cdot \delta + \Delta H - H \quad (2)$$

4 Design and Adjusting

As shown in Fig. 3, the automatic control system of adjusting off pattern is composed of an image acquisition and processing unit, a main controller and upper machine operation system.

The image acquisition and process unit includes image processor, optical lens, area CCD camera and LED light source, and so on. The pattern image set is captured by CCD camera with optical lens and then the image signal collected is sent to the image processor. In designing, a ring white reflective LED light source and the image processor by Keyence with image gathering card are selected, the latter can be used to process images with a large number of pixels with complex algorithm calculations and to support four cameras running simultaneously and many types of interfaces.

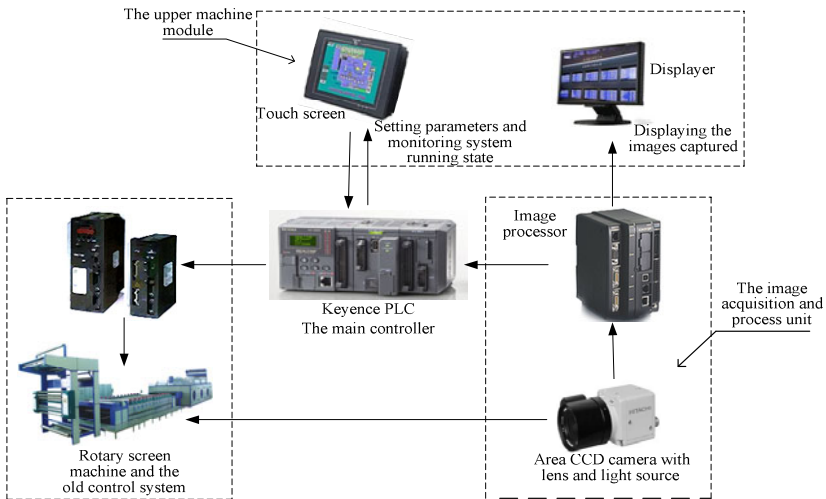


Fig. 3. Schematic diagram for on-line control system of inspecting and adjusting off pattern

The main controller is used to accomplish automatic and manual mode speed regulation of printing roller motors according to the error signal collected with the image acquisition and processing unit. In automatic mode, it fulfills pulse counting of the encoder, calculations of printing cycle error and data storage. And in manual mode, it directly sends control pulses to manual adjusting button of the old control system for rotary screen printing machine.

The upper machine operation module consists of the touch screen and the image displayer. It adopts RS-485 bus to make connection and achieve communication between the touch screen and the main controller. The touch screen monitor performs system initialization, monitors the actual running state of the whole system, and modifies control parameters of the main controller.

5 Technical Features

The system introduced in the paper has the following features:

1) It has a function of traditional adjustment with manually accelerating and decelerating as well as automatic control.

2) It has the ability to adopt subdivision technology for the large deviation of the actual printing cycle. For example, the maximum adjustable error for system is set 3mm according to experience and actual conditions. If the deviation value is more than 3mm the control system makes maximum adjustment of 3mm at this time and the residual deviation is left to adjust step by step, the maximum adjustment being 3mm every time.

3) It considers the case of duplicate capturing of the same mark. In programming, the PLC controller only stores the first position error of the same image from the image processor, for that the experiments prove that the calculated results are equal by different error of the duplicate captured image.

4) The problem of missing acquisition of some marks in printing cycle is solved with the numbers of the image collected corresponding to the actual measurement and then with treatment of the error and pulse transformation by the main controller.

6 Conclusion

The system introduced in the paper has realized closed-loop registration control by both predicting and adjusting for the off pattern error and feedback detecting of adjustment result. And it is easy to connect with the old control system to perform manual adjustment. The experiment shows that the system is reliable and stable and the precision of registration is up to $\pm 0.8\text{mm}$ at the max speed of 60m/min. It is proved that the product quality and production ratio are improved greatly with the system and the new solution method is provided with the system for engineers working on off pattern checking.

Acknowledgments. The paper is sponsored by the Research Project of “SUST Spring Bud”(06540040411).

References

1. Wu, X.: Present situation and development of printing technology. *Dyeing & Finishing* (12), 44–47 (2003)
2. Zhang, T., Liu, Y., Guo, Y., Liu, Z., Xue, C.: The broken design analysis of the circle net independence transmission printing machine. *Progress in Textile Science & Technology* 48(4), 37–38 (2005)
3. Ye, X.: A study on mechanical factors affecting register precision of rotary screen. *Dyeing & Finishing* (3), 29–30 (2004)
4. Liu, S., Qu, P., Fei, P.: The applications of machine’s visual function textile inspection. *Journal of Textile Research* 24(6), 89–91 (2004)

5. Jing, J., Kang, X., Li, P.: Detection of the fabric printing quality based on the machine vision. In: ICICTA 2010, vol. 2, pp. 75–78 (2010)
6. Huang, Z., Qiu, J.: Research of rotary screen printing machine automatism patterning system based on machine vision. *Progress in Textile Science & Technology* (3), 48–50 (2006)
7. Gu, J., Qin, Y.: Innovation automatic registering technique for rotary screen printing machine. *Dyeing & Finishing* (19), 26–27, 34 (2010)

Analysis of Test Scheme for a Bayesian Plan of Qualification Test in Binomial Case

Zhiqian Ren and Zhimao Ming

School of Mechatronics Engineering & Automatization, National Univ. of Defense Technology,
Changsha, China, 410073
rzq_rtl@nudt.edu.cn

Abstract. A Bayesian plan of reliability qualification test is discussed in combination with the results obtained from the actual tests. On the premise that the test sample size is quantitative, a scheme of reliability qualification test in binomial case is formulated by using prior information efficiently. From the view of the meaning posterior risk actually concerned by users and producers, the difference and connection between the Bayes reliability qualification test and evaluation scheme has been analyzed, which could be the guidance for test scheme selection.

Keywords: Binomial distribution, bayesian method, reliability qualification test, producer's risk, user's risk.

1 Introduction

Design and produce a new-type product usually undergoes several steps-theoretical prototype, initial prototype and final production. According to the current military standard, it needs a large sample size for longtime testing, which increase the testing cost dramatically[1]. Thus, this method is not suitable for the evaluation of complex and costly small-sample weapon equipments, because the producers and users want to test the product reliability and acceptance with the given sample size. In order to solve this problem, it is essential to take full advantage of the testing information and make the system type approval test rational[2].

From the view of meaning posterior risk, this paper proposes the test scheme for a Bayesian plan of reliability qualification test, and analyzes how to design the optimal reliability qualification test in small-sample case without decreasing test precision and increasing test risk.

2 Bayesian Plan of Reliability Qualification Test

Binomial model is widely used in reliability analysis of large scale and complex systems[2]. Therefore, we use the binomial distribution Bayes qualification test for analyzing the reliability and acceptance of complex system.

2.1 Determining Prior Distribution

Assume that the success rate of binomial distribution is R , and the failure rate is $q = 1 - R$. The probability of s successes in n samples follows a binomial distribution, as shown in Equ. (1):

$$P(S = s) = \binom{n}{s} R^s (1 - R)^{n-s}, \quad s = 0, 1, \dots, n \tag{1}$$

In order to minimize the sample size, we use the Bayesian plan of reliability qualification test. Firstly, determine the prior distribution of the success rate R . Usually, Beta distribution is used as the prior distribution for binomial case[2], so we use the Beta distribution of R :

$$\pi(R) = \frac{R^{a-1} (1 - R)^{b-1}}{\beta(a, b)}, \quad 0 \leq R \leq 1 \tag{2}$$

Where a and b are the prior hyper-parameters, and:

$$\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)} \tag{3}$$

Usually, we use empirical Bayes method proposed by Martz & Waller[5], or the method of prior quantile [3] to determine the prior hyper-parameter a, b .

2.2 Bayesian Plan of Qualification Test

According to Equ.(1) and Equ.(2), the posterior distribution of R is:

$$\pi(R | D) = \frac{L(R)\pi(R)}{\int_{\Theta_R} L(R)\pi(R)dR} = \frac{R^{n-f+a-1} (1 - R)^{b+f-1}}{\beta(n - f + a, b + f)} \tag{4}$$

Assume that qualification scheme is (n, c) , if the system reliability is R , then the acceptance probability of the qualification test is:

$$P(Ac | R) = \sum_{f=0}^c \binom{n}{f} R^{n-f} (1 - R)^f \tag{5}$$

The rejection probability is:

$$P(Re | R) = \sum_{f=c+1}^n \binom{n}{f} R^{n-f} (1 - R)^f \tag{6}$$

Suppose that we use the primary sampling scheme (n, c) , n is the number of tests, c is the maximum acceptable fault numbers, the acceptance quality is R_0 , the

limiting quality is R_0' , the producer's risk and user's risk are α , β , respectively. Once the discrimination ratio D , the upper limits R_0 and the lower limits R_0' are determined, the producer's risk α and user's risk β can be described as followed:

$$\left\{ \begin{array}{l} \alpha = \frac{\int_{R_0}^1 P(\text{Re} | R) \cdot \pi(R|D) dR}{\int_0^1 \pi(R|D) dR} \\ \beta = \frac{\int_0^{R_1} P(\text{Ac} | R) \cdot \pi(R|D) dR}{\int_0^1 \pi(R|D) dR} \end{array} \right. \quad (7)$$

Once α and β are determined, the qualification scheme (n, c) can be calculated.

The result of Equ.(7) is the optimal theoretical scheme of qualification test, but not the optimal scheme for small sample test.

3 Comparison and Analysis of Bayesian Plan and Evaluation Scheme

In engineering practice, both the producer and the user concern about how to achieve the most satisfying results from minimum sample size. Therefore, the Bayesian plan of qualification test is more suitable for small sample case. In order to illuminate the characteristics of Bayes qualification test, we compare this method with the evaluation scheme, and analyze their differences.

Suppose that there is a binomial system, if the confidence level $\gamma = 0.8$, then the system reliability is 0.75. In practice, design and technics of this product has been improved twice. The initial sample has been tested 100 times, and the passing rate is 88%; the evaluation model has been tested 17 times, and the passing rate is 71%; the final production should be tested under environments I and II. Suppose that 10 tests have been done under environment I, and the fault number maybe 0 or 1, then how many tests should be done under environment II?

3.1 Implementation of Bayesian Plan

Step1. Estimating the historical data with classical reliability evaluation method [6]. If the system reliability is more than 0.75, the reliability qualification test will be carried out.

Step2. Calculating system reliability with classical reliability qualification scheme [4]. The result is shown in table 1, if $D = 2.0, \alpha = \beta = 0.2$, the required reliability qualification scheme is $(31, 5)$. We can see that, the result is not acceptable for both the producer and the user. Thus, reliability must be calculated with given sample size.

Table 1. Classical qualification scheme (0.75)

R_1	D	$\alpha = \beta = 0.3$	$\alpha = \beta = 0.2$	$\alpha = \beta = 0.1$
		(N C)	(N C)	(N C)
0.75	2.0	(14, 2)	(31, 5)	(64, 11)

Step 3. Using the environmental factors to convert the system reliability[7]. Equ.(8) is the converting equation. Assume that environment II represents the real environment (rigor), the environmental factor K is known, and the good environment is converted to the rigor one. K is divided with the sample size of good environment, if the fault number f is invariable, that means the test information is turning worse. Assume that the unreliability is P_1 and P_2 under environment I and II, respectively. When the environment data is converted from I to II, then:

$$K = 1 - R_2 / 1 - R_1 \Rightarrow R_2 = 1 - K(1 - R_1) \tag{8}$$

From Equ. (8), we can get different environment conversion.

Step 4. Using prior information we can get the value of a, b . Suppose that 10 tests have been carried out under environment I, and the fault number is 0, 1 and 2. By using the Bayesian plan, the sample quantity and real risk under environment II can be calculated by Equ.(7). Table 2 shows the sample size and real risk under environment II.

Table 2. Bayesian plan of qualification scheme (D=2.0)

Environment I		Environment II		Decision Risk			
Num. of tests	Fault times	Num. of tests	Fault number	nominal value		Real value	
		N	C	α_{π_0}	β_{π_0}	α'_{π_0}	β'_{π_0}
10	0	3	0	0.20	0.20	0.4910	0.1841
		6	1	0.20	0.20	0.3259	0.1935
		11	2	0.20	0.20	0.3016	0.1914
		16	3	0.20	0.20	0.2479	0.1855
		20	4	0.20	0.20	0.1850	0.2059
	1	5	0	0.20	0.20	0.1733	0.1658
		8	1	0.20	0.20	0.1182	0.2043
		13	2	0.20	0.20	0.1013	0.1930

In practice, both the producer and user concern more about the lowest test cost. For example, if the confidence interval $\gamma = 0.8$, there will be several test schemes that can satisfy the system reliability $R_L \geq 0.75$. According to GB4087.3-1985[8](first line in

Table 3), if the fault number is 0, 1, 2, 3, 4, 5, 6, then the producer and user may tend to choose the scheme with minimum tests, such as (6,0). However, GB4087.3 doesn't calculate the exact duplex risk, so it is difficult to decide whether the minimum test scheme is optimal.

Table 3. One-sided lower confidence limit of reliability for binomial distribution ($\gamma = 0.8$)

(N, C)	(6, 0)	(11, 1)	(16, 2)	(21, 3)	(26, 4)	(31, 5)	(35, 6)
R_L	0.7647	0.7514	0.7512	0.7529	0.7549	0.7569	0.7521
α'	0.5512	0.4081	0.3229	0.2634	0.2187	0.1837	0.1395
β'	0.1779	0.1971	0.1971	0.1917	0.1844	0.1764	0.1919

3.2 Comparison and Analysis of Bayesian Plan and GB4087.3-1985

For some determined (n, c) , the probability of batch production is:

$$L(p) = \sum_0^c P(n, r|p) \tag{9}$$

Thus, (n, c) satisfies:

$$\begin{cases} 1 - L(p_0) = 1 - \sum_0^c P(n, r|p_0) = \alpha \\ L(p_1) = \sum_0^c P(n, r|p_1) = \beta \end{cases} \tag{10}$$

Since n, c are positive integers, Equ.(7) can be approximate satisfied. Using Equ.(10), we can get α' and β' which represent the real risk of qualification scheme (n, c) . The value of α' and β' is shown in Table 3.

From table 3, we can see that if the test judgment fault number is 0,1,2,3 and 4, in order to guarantee the user's risk is approximate to 20%, the producer's risk is more than 20%. GB 4087.3-1985 only gives the confidence interval, but in reliability qualification test, both the producer's risk and the user's risk should be considered. Furthermore, sample size and test judgment fault number should be provided.

After determine the risk, the reliability qualification test scheme should be chosen according to actual situation. In line 2 of table 3, we can see that there are several groups satisfying the reliability of 0.75. In practice, we need to select the groups that can satisfy both the reliability and the risk. Therefore, test scheme (26,4),(31,5) in table 3 are selected, because they can satisfy both the reliability of 0.75 and the duplex risk of 20%. Hence, the reliability evaluation and classical qualification are coincident. Whereas the classical qualification scheme does not use the prior

information, so the test risk is higher. After the reliability qualification test, the test data can be estimated and compared with GB 4087.3-1985 to check whether the product is acceptable.

Moreover, from table 2, we can see that if the test scheme is (10,0) under environment I, then it should be (16,3) under environment II. Compared with the classical qualification scheme, the sample quantity of Bayesian plan is 48.38% less. If the test scheme is (11,2) under environment II, then the sample quantity of Bayesian plan is 64.5% less than the classical qualification scheme. From all above, we can get the conclusion that the Bayesian plan of qualification test can save the sample size efficiently.

4 Conclusion

This paper discusses the test scheme for a Bayesian plan of qualification test in binomial case. For small sample reliability test, a new Bayesian plan is proposed for qualification test. The comparison between our method and the evaluation scheme shows that our Bayesian plan can guarantee the real risk. Therefore, in engineering practice, the Bayesian plan is more rational and efficient.

Acknowledgement. This paper is supported by the Technology Foundation Project of General Equipment Headquarters of PLA.

References

1. Zhang, J.H., Liu, Q., Feng, J.: Analysis methods of Bayesian test. University of National Defence Technology Press, Changsha (2007) (in Chinese)
2. Cai, H., Zhang, S.F.: Analysis and evaluation of Bayesian test. University of National Defence Technology Press, Changsha (2004) (in Chinese)
3. Jiang, L.P., Zhang, Z.H.: A Bayesian Method of Reliability Qualification Test in Binomial Case. *Journal of Engineering Mathematics* 17(4), 25–29 (2000) (in Chinese)
4. He, G.W.: Techniques of reliability test. National Defense Industry Press, Beijing (1995) (in Chinese)
5. Martz, H.F., Waller, R.A.: The basics of Bayesian reliability estimation from attribute test data. Los Alamos Scientific Laboratory: Report UC-79p (February 1976)
6. Zhou, Y.Q.: Qualitative reliability growth and evaluation method. Beijing University of Aeronautics and Astronautics Press, Beijing (1999) (in Chinese)
7. Hu, M.X., Song, B.W., Shao, C.: Reliability Growth of Torpedo in Various Environments. *Journal of Projectiles, Rockets, Missiles and Guidance* 26(1), 437–439 (in Chinese)
8. GB 4087. 3-85. Statistic disposal and interpretation for data, one-sided lower confidence limit of reliability for binomial distribution. China Standard Press, Beijing

Wireless Sensor Networks Fusion Based on the Hybrid of Ordering Filter and Adaptive Weight

Changjiang Wang^{1,2}, Liqin Fu², and Yongmei Zhang³

¹ School of Economics and Management, North University of China, Taiyuan, China

² National Key Laboratory for Electronic Measurement Technology, North University of China, Taiyuan, China

{changj.w, liqin_fu}@126.com

³ School of Information Engineering, North China University of Technology, Beijing, China

Abstract. Wireless sensor network (WSN) has developed with recent advancement in wireless communication and electronics. Wireless sensor network consists of very large number of tiny sensing devices, also called motes, distributed over physical space. However, a system of sensor nodes is highly energy constrained. Fusion algorithms provide a mechanism to extract pertinent information from massively sensed data and identify incipient sensor failures. Aiming at save energy in wireless sensor network, this work presents a novel fusion approach which is based on a hybrid of order filter and adaptive weighted fusion. The sensor fusion system includes three steps: aggregation, abnormal data eliminating and adaptive weighted fusion. The aggregation made the task of dividing sensors into several groups. The abnormal data eliminating is to eliminate pulse noise which can bring bad influence to fusion results. The experimental results have verified the effectiveness of the algorithm.

Keywords: Wireless Sensor Network, Data Redundancy, Data Fusion, Fusion Cost.

1 Introduction

In the past couple of years, wireless sensor networks have found a wide range of applications such as environmental monitoring [1], robotic and object localization and tracking [2]. In wireless sensor networks, sensor nodes can self organize to form a network and can communicate with each other in a wireless manner. Each node has transmit power control and an omni-directional antenna, and therefore can adjust the area of coverage with its wireless transmission.

In practical applications, WSN has the following physical resource constraints: Firstly, The bandwidth of wireless links connecting sensor nodes is usually limited, on the order of a few hundred Kbps. Secondly, sensor nodes have limited supply of energy, and energy conservation is one of the main system design considerations. Current small batteries provide about 3000mAh of capacity, powering the MICA Mote for approximately one year in the idle state and for one week under full load. Note that future sensor nodes will have sophisticated power management features; current nodes already have three different sleep modes with several orders of

magnitude different power usages [3]. Thirdly, sensor nodes have limited computing power and memory sizes that restrict the types of data processing algorithms. Finally, signals detected at physical sensors have uncertainty due to limitations of the sensor, and they may contain environmental noise. Sensor malfunctions might generate inaccurate data, and unfortunate sensor placement might bias individual readings.

With consideration of power consumption, each particular scheme will depend significantly on the relative power consumption of different approaches. Specifically, if communication energy is significantly higher than the computation energy, then it is important to develop localized algorithms that will require only a limited amount of communication. Therefore, with respect to fault tolerance, it is important to consider schemes that conduct error detection using only local information. Or, if one wants to ensure fault tolerance during the sensor fusion, the goal is to design fault tolerant techniques that do not significantly increase the communication overhead. On the other hand if the computation energy is significantly higher than the communication requirements, it is a good idea to support communication resources at one node with the computation resources at other nodes. It is preferable to develop fault tolerant sensor fusion approaches that require little additional computation regardless of any additional communication requirements.

Besides, data fusion algorithm should be designed in consideration of application background. Each device is capable of limited computing, radio communication and sensing.

2 Sensor Fusion Algorithm

2.1 General

While a mote sensor network promises to provide more useful information about the monitored environment than other sensing alternatives, the scope of each single miniature sensor node is much smaller and easily affected by local disturbances. Sensor information is always corrupted to some degree by noise and sensor degradation varying with operating conditions, environmental conditions, and other factors. The inexpensive, less reliable and massively distributed MEMS mote sensors will be even more prone to individual failure. Moreover, massively transmitted data packages in the sensor network increase the chance of data collision and interference, and thus are more likely to cause loss of data or incorrect data.

The accuracy and efficiency of the mote sensor network depends on how the pertinent global information is extracted from the local information of each sensor node, highlighting the need for an effective sensor fusion algorithm. The sensor fusion system includes three steps: aggregation, abnormal data eliminating and adaptive weighted fusion, shown in Fig.1. The aggregation made the task of dividing sensors into several groups. The abnormal data eliminating is to eliminate pulse noise which can bring bad influence to fusion results.

2.2 Aggregation

Aggregation refers to delivering data from distributed source sensor nodes to a central node for computation. It is one of the most popular computation and communication

pattern for sensor network. A join of sensor readings from different groups could be another example; all records need to be collected at a central node where the join takes place. Aggregation involves two important issues. First, from a computational point of view, the aggregation has to complete at a "leader" node (unless the final computation of the aggregate is delegated to a gateway node or happens outside of the network). Second, the data records have to be delivered from source sensor nodes to the designated leader. Note that the system has to be designed to tolerate the volatility of the underlying communication layer: messages could get lost, nodes could die, and the network could be partitioned for a while. Let us shortly contemplate leader selection and data delivery in the following two paragraphs.

If the computation is designated to a leader node, such a leader needs to be selected among the sensor nodes. There are several basic requirements for the leader selection policy. First, the leader should be dynamically maintained in case of sensor or link failure. (We imagine backup leaders to reduce the cost in case of leader failure.) Here we can draw upon a large literature about algorithms in distributed systems and the leader election problem. Second, we would like to select a leader with physically advantageous location: The cost of data delivery from source sensors to the leader and the delivery cost from the leader to the gateway node need to be taken into account to save communication energy.

In this paper, the algorithm discussed in reference [4] is adopted. When all the nodes are grouped, nodes within each group can be fused respectively.

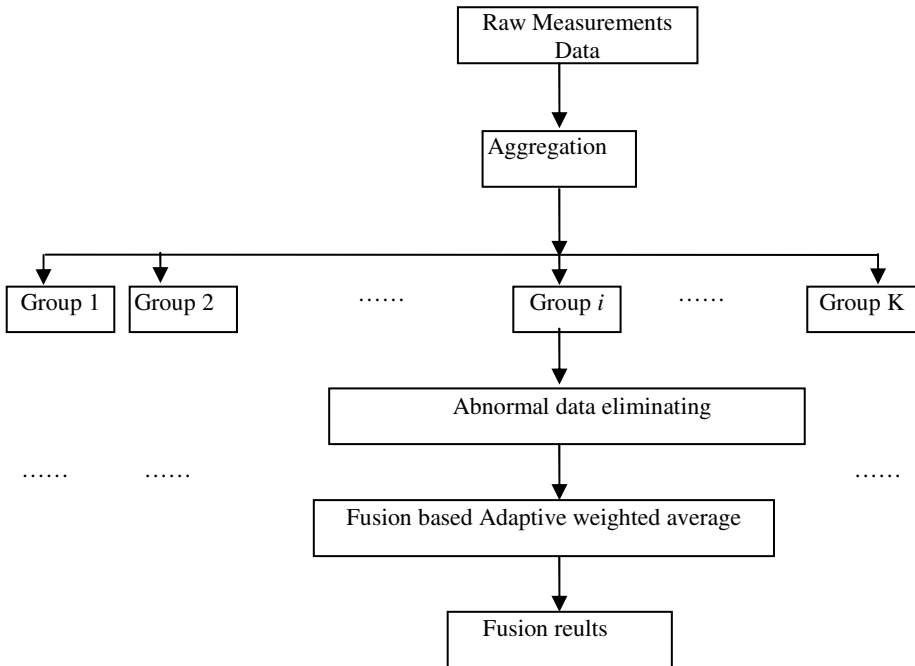


Fig. 1. The diagram of sensor fusion

2.3 Abnormal Value Elimination

The abnormal data eliminating is based on order filtering. This method of analyzing the behavior of a cluster of sensor readings first filters out the obvious false readings based on the physical limitation of the sensors. It identifies the abnormal data by evaluating the difference between a value and other values in its neighbor region. Generally, the median value should reveal a reasonable estimation of the majority of sensor readings. If the difference between a sensor reading and the median is larger than the threshold, we think it is abnormal and eliminate it in order to reduce the influence from unreasonable readings which stand for pulse noise.

The algorithm orders all data in a group firstly and finds the median Med . Next, sets a threshold and compare each value before the median in ordered array with the median. If the difference is larger than the threshold, eliminate the corresponding sensor reading, till all small abnormal values are eliminated. Then compare each value after the median in ordered array with the median. If the difference is larger than the threshold, eliminate the corresponding sensor reading, till all large abnormal values are eliminated.

2.4 Adaptive Weighted Average

Because of the difference in accuracy and environment among sensors, the conventional simple average fusion can not attain good result. Fusion based adaptive weighted average is performed by taking the average of measurements in a group weighted by corresponding confidence values plus the predicted value weighted by α , an adaptive parameter representing the sensor reliability. If the sensor is reliable and in steady state, α is set to a large value in order to weight its measurement more; on the other hand, if the sensor is not reliable relatively or in a transient state, α is set to a small value in order to weight its measurement value less so as to reduce its influence.

In this paper, we implement fusion based adaptive weighted average for each group. The determination of weight is the key step. The basic principle is to make the evaluation value \hat{X} attain optimal. The weight α is searched adaptively according to measurement values of all sensors in a group on the premise of minimal average square error.

3 System Performance Evaluation via Simulations

OMNeT++ simulation platform is used to evaluate fusion algorithm.

3.1 Energy Consumption Calculation

As shown in Fig. 2, a total of N sensors are evenly deployed in the region of interest (ROI), which is a square with area $50m \times 50m$.

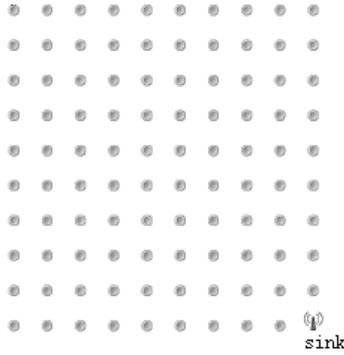


Fig. 2. 100-node topology for a 50m x 50m network. Sink is located at (10,10)

We assume that all sensor nodes have ability of data collection and retransmission, each node produce a data package of 200 bytes, and data will be sent to sink node. To evaluate the performance of, we also simulate adaptive weighted fusion (AW) which will be compared with the algorithm we proposed (OFAW) .

Let r_c stand for maximal communication radius of sensor node. When the distance between two nodes is smaller than r_c , communication chain can be set up. Let d stand for Euclidean distance between two nodes, e stands for communication route between two nodes. e can be calculated as Equation (1) when $d < r_c$.

$$c(e) = \beta d^n + \varepsilon \quad (1)$$

There stands for the energy consumption when 1bit data be sent or received. Its value is within the range of 20~200nJ /bit. β and n are communication parameters which are adjustable. In this paper, $n=2$, $\beta = 100$, $\varepsilon = 100$.

To validate the ability to save energy consumption of the fusion algorithm, node increasing gradually is used and it is assumed that there are 20% abnormal data (according to abnormal nodes). The energy consumption of adaptive weighted fusion (AW) and hybrid of order filter and adaptive weighted fusion (OFAW) is shown in Fig.3. X-coordinate represents the number of nodes, Y-coordinate represents energy consumption ratio which is the ratio of energy consumption of after-fusion and before-fusion.

Low-energy-consumption is the basic requirement to wireless sensor network. From Fig.3, we can see that the energy consumption ratio of both algorithms will decrease with the increasing of the number of nodes and OFAW is more obvious.

3.2 Average Communication Load

Average communication load is used to measure the performance of OFAW. We can calculate it through counting the average jump number from source node to sink node. In this section, we select five kind of network size which have the same node density. In the first networks, 60 nodes distributes in the region of 40m*30m. There are 100

nodes distributed in the region of 50m*40m in network 2. There are 150 nodes in the region of 60m*50m in network 3. There are 200 nodes in the region of 80m*50m in network 4. There are 300 nodes in the region of 100m*60m in network 5. It is assumed that there is a sink node in each network and sink node is in the central of each network. Moreover, there are 20% abnormal data.

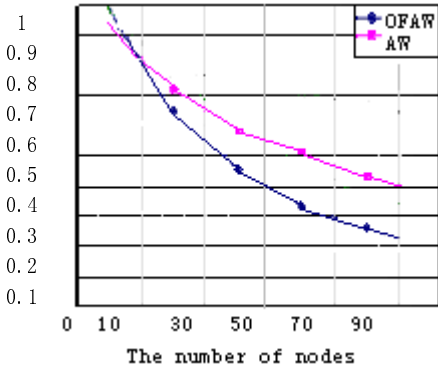


Fig. 3. The relation between ECR and node number $N = 100, n = 2, \epsilon = 100$

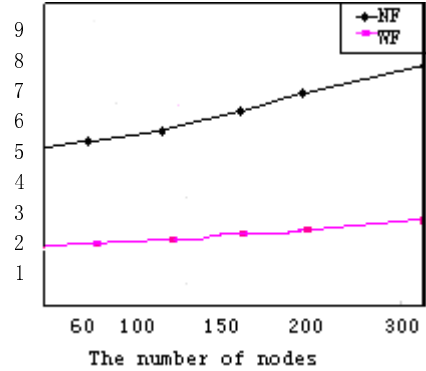


Fig. 4. The relation between average communication load and network size

From Fig.4, we can obtain the conclusion: average communication load is lower using the algorithm with fusion.

4 Conclusion

We have proposed a fusion algorithm based on the order filter and adaptive weighted average for a WSN with a large number of sensors. Our analysis shows that this fusion rule can achieve very good performance.

In this paper, the total number of sensors N is assumed known. In the future, this assumption will be relaxed and the case where N is a random variable will be investigated.

References

1. Elson, J., Estrin, D.: Sensor Networks: A bridge to the Physical World, pp. 3–20. Kluwer Academic Publishers, Norweill (2004)
2. Pottie, G., Kaiser, W.: Wireless Integrated Network Sensor. Communication of the ACM 43, 551–558 (2000)
3. Hill, J., Szewczyk, I., Woo, A., Culler, D., Hollar, S., Pister, K.: System architecture directions for networked sensors. ACM SIGPLAN Notices 35, 93–104 (2009)
4. Fan, K.-W., Liu, S., Sinha, P.: On the potential of structure-free data aggregation in sensor networks. IEEE INFOCOM (2006)

Process Control and Intelligent Algorithms of Network Optimization Problems under Uncertain System

Jing Zhang

Basic Courses Department, Beijing Union University
Beijing, China
zhang1jing4@sina.com

Abstract. The establishment of network optimization model under uncertain systems and intelligent algorithms mainly based on the theory of uncertainty with random variables or fuzzy variables of basic network optimization problems, such as the shortest path problem, minimum spanning tree problems, and transportation problems were studied. For the intelligent optimization algorithms, such as artificial neural networks, evolutionary computation, swarm intelligence (e.g. ant colony optimization, particle swarm optimization, etc.), artificial immune systems, making network optimization based on uncertain system model of the theory, including parameter selection, hybrid intelligent algorithm design. Intelligent design has been calculated in principle on the basis of efforts to put forward new intelligent optimization algorithms. Overcome the current difficulties encountered in intelligent computing problems, and make a reasonable explanation based on mathematical theory, algorithm in the full network optimization model of effectiveness uncertain.

Keywords: Intelligent algorithms, Fuzzy control, Uncertain theory, Network optimization.

1 Introduction

Network optimization problem is the engineering, economic management and scientific research and other problems often encountered in many fields, is an important branch of operations research, mainly research how to effectively plan, manage and control network systems to maximize efficiency. And there are some combinatorial optimization problems, such as scheduling, packing problem, knapsack problem, assignment problem, traveling salesman problem, Steiner minimum tree problems, often can be transformed into a network optimization to deal with. Since many practical problems can be transformed into these fundamental issues are resolved, so on these issues, theory, algorithms and applications in a number of documents are described in detail in [1-3].

In many specific applications, we have encountered information, there is subjective or objective uncertainty, these uncertainty manifestations are diverse, such as randomness, fuzziness, roughness, etc. There is uncertainty both objective factors but also non-objective factors, but its existence is widespread. Such as transportation

network, weather, traffic uncertainty, logistic network in the uncertainty of customer demand, or the lack of historical data and the data is not perfect, some data can only experience by experts according to their own subjective estimates given the formation of the uncertainty. On the other hand, the key to network optimization problem is to study combinatorial optimization problems for the network model and solution techniques, which are usually non-linear optimization problems, multi-objective and multi-dimensional characteristics, and many problems are NP-hard problem. For the above two considerations, how to deal with uncertainty, how to achieve effective optimization, we solve the problem and focus of the two basic points.

2 Theory and Algorithms

Around these uncertain systems, uncertainty theory came into being. Probability is a number of laws of random phenomena research branch of mathematics. Today, the idea of probability theory has been infiltrated into the various disciplines and is widely applied to engineering, management, military, aerospace and other fields. We can say that probability theory first revealed one aspect of uncertain systems. The concept of fuzzy sets was first proposed by Zadeh [4] and subsequently Kaufmann [5] proposed the concept of fuzzy variables, Zadeh [6] introduced the possibility theory, and Nahmias [7] the possibility of using three axioms define the measure, and re-defines the concept of fuzzy variables, fuzzy variables is defined as the possibility of space to the function set of real numbers.

Liu [8] gives the product of the fourth axiom defines the possibility measure, probability theory to the study of a broader platform. The basis of these four axioms, Liu successfully established credibility of the axiomatic system. Thus, Liu system described the contents of the theoretical study of uncertainty; the uncertainty of the basic theory, the theoretical part is divided into three categories: probability theory, credibility theory, opportunity theory. Development and improvement of the theory of uncertainty for our network optimization problems encountered in dealing with uncertainty has laid a theoretical foundation.

Traditional optimization algorithms can not effectively solve this problem, so find a suitable for massively parallel algorithms and intelligent features of the subject has become a major research objective and compelling research. Since the 1980s, a number of novel optimization algorithm, such as artificial neural networks, evolutionary computation, swarm intelligence (e.g. ant colony optimization, particle swarm optimization, etc.), artificial immune systems, through analog or reveal certain natural phenomena or processes to develop their ideas and content related to mathematics, physics, biological evolution, neuroscience, and statistical mechanics, etc., to solve complex problems, provides new ideas and methods. Unique advantages of these algorithms and mechanisms, causing extensive attention of scholars home and abroad, and set off a wave of research in this field, and in many fields have been successfully applied. Since these algorithms construct the intuitive and natural mechanism, which is often called the modern intelligent optimization algorithm or optimization algorithms. Therefore, we are dealing with uncertain systems under network optimization problems, specific ways to achieve intelligent computing is natural for a special issue of how to design efficient algorithms that not only is a smart application of technology, but also a basic theory. This is the application of this project as one of the key elements of basic research.

3 Status of Research

Martin [32] first studied in 1965, subject to multi-dimensional normal distribution side of a random network of independent random variables. In 1969 Frank [10] is the first study and published results of randomized studies shortest path scholar, he made a similar distribution of the optimal path calculation. Mirchandani [11] proposed a random network shortest path to get the distribution function method, Martins [12] use dynamic programming to find the shortest path problem of double standards valid path set; Hall [13] studied the time-dependent stochastic shortest path problem on the network. 1980, Dubois and Prade [14-15] the first analysis for the fuzzy shortest discuss their ideas mainly on the typical shortest path algorithm, such as Ford and Floyd algorithm changes. Based on possibility theory, they proposed a way of criticality concept. Klein's concept based on fuzzy utility function, we propose a dynamic programming method [33]. Okada et al [17-18] on the fuzzy shortest research more, the larger the network, I get a lot of non-dominated paths, which policy-makers, it is difficult to choose a satisfactory solution in this the direction of the study is how to reduce non-dominated paths are set for the number of decision-makers choose the path. Calculate the exact distribution of possible flow method in the literature is given. 2004, Katagiri et al [19] discussed the fuzzy random edge weights under a bottleneck spanning tree problem of chance-constrained programming model. Chanas et al [21] will be extended to the transport of the classic fuzzy environment, Mohammad and Chalam [22] transport of random fuzzy goal programming method proposed. Luhandjula [23] studied the single-objective transportation problem solving interval number uncertain planning problems.

Bit [25] using a very small operator and a linear membership function with fuzzy coefficients proposed multi-objective transportation problem algorithm. Datar and Ranade [27] studied the bus network problems, imagine traveling by bus in the city, the goal is to minimize the total travel time. In this case, a natural goal is to minimize the total expected travel time of planning. For network flow problems, Fishman [28, 29] studied the capacity of a continuous arc of the situation, and gives the simulation obtained using the solution on the boundary.

Uncertain system network optimization problem solving, is to focus on uncertainty theory, network optimization model, and the intersection of intelligent algorithms, a comprehensive study of its engineering, economic management and scientific research and other fields to play a role, its importance is not self-evident. Therefore, we propose a research project with the current international research trends, and have important theoretical and practical significance.

4 Contents and Purposes

Focus on network optimization problem under uncertain system models and intelligent algorithms, mainly based on credibility theory and the theory of opportunistic random variables or fuzzy variables with the shortest path problem, minimum spanning tree problems, transportation problems and network location problems for research. Strive to research and production practice models and algorithms have great similarity, can be directly used in some specific practical decision-making.

For the intelligent optimization algorithms, such as artificial neural networks, evolutionary computation, swarm intelligence (e.g. ant colony optimization, particle swarm optimization, etc.), artificial immune systems, making network optimization based on uncertain system model of the theory, including parameter selection, hybrid algorithm design. Intelligent design has been calculated in principle on the basis of efforts to put forward new intelligent optimization algorithms.

But the practical applications there are many factors constraining its functioning. Therefore, it is necessary to address the practical application of intelligent computing in urgent need exists to overcome the key issues, such as: artificial immune system antibodies and other high performance and complexity of the algorithm; genetic algorithm to solve multi-objective optimization problem of coding issues and "premature," and so on; PSO local convergence problems. These problems also restrict the intelligent algorithms in dealing with uncertain systems based on network performance optimization, how to make targeted treatment and response is one of the elements of this research.

Research objectives of this project is uncertain system for basic network optimization problems, for example, the shortest path problem, minimum spanning tree problem, minimum cost flow problem, transportation problems, the establishment of a package with a higher production practice the fit model to an effective intelligent algorithm implementations. Establish a framework for both the theoretical level, but also the establishment of application-level operation of the program.

The establishment of network optimization model under uncertain systems, making the model for the solution of practical problems have general applicability, rather than not determine the network optimization model theory and give the reference, so that from a typical practical application, or a model with more great specificity, affecting the promotion of its theoretical value and practical application.

In the intelligent computing, the need to seek for a class of uncertain systems model of effective network optimization algorithms, and thus able to propose a new intelligent algorithm; or a hybrid intelligent strategy, configuration algorithm for an effective combination of program issues, to analyze the combined performance of the internal mechanism and then seek one of the laws of science.

In this research project the application level, how to use the established network optimization models to solve practical with a strong background in combinatorial optimization problems, such as the scheduling problem, bin-packing problem, knapsack problem, they are also technical difficulties, then do not make a deal network optimization system to determine the practical problems the package design.

5 Conclusions

The first one from the existing practical problems starting, such as ship scheduling problem (little research on this problem at home and abroad), the abstract analysis of the required network optimization model, given the general uncertainty of the system and then the network optimization model under . Thus, our proposed model is more realistic network optimization and application value. Mathematical and computer simulation experiments as a smart algorithm to establish an efficient way, the actual efficiency of the algorithm to determine the algorithm as the real merits of the

measure as a basis, the mechanism by mathematical analysis of its causes, trying to create newer and better intelligence algorithm. Also, note that a mixed strategy for intelligent computing solutions. Focus on technology integration.

Acknowledgments. The work is supported by Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (PHR (IHLB)) (THR201108407). Thanks for the help.

References

1. Murty, K.G.: Network Programming. Prentice-Hall, Englewood Cliffs (1992)
2. Abuja, R.K., Magnanti, T.L., Orlin, J.B.: Network Flows: Theory, Algorithm, and Applications. Prentice-Hall, Englewood Cliffs (1993)
3. Ravindran, A., Phillips, D.T., Solberg, J.J.: Operation Research: Principles and Practice. Wiley, New York (1987)
4. Zadeh, L.A.: Fuzzy sets. *Information and Control* 8, 338–353 (1965)
5. Kaufmann, A.: Introduction to the Theory of Fuzzy Subsets. Academic Press, New York (1975)
6. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 1, 3–28 (1978)
7. Nahmias, S.: Fuzzy variables. *Fuzzy Sets and Systems* 1, 97–110 (1978)
8. Liu, B.: Uncertainty Theory: An Introduction to its Axiomatic Foundations. Springer, Berlin (2004)
9. Zhang, J.: Intelligence Computing Methods in Electronic Commerce and Security. In: 2nd International Symposium on Electronic Commerce and Security, ISECS 2009, May 22–24, vol. 1, pp. 398–402. IEEE Computer Society, Los Alamitos (2009)
10. Frank, H.: Shortest path in probabilistic graphs. *Operations Research* 17, 583–599 (1969)
11. Mirchandani, P., Soroush, H.: Optimal Paths in Probabilistic Networks: A Case with Temporary Preferences. *Computer and Operations Research* 12, 365–381 (1985)
12. Martins, E.: On a multi-criteria shortest path problem. *European Journal of Operational Research* 16, 236–245 (1984)
13. Hall, R.: The fastest path through a network with random time-dependent travel time. *Transportation Science* 20(3), 182–188 (1986)
14. Dubois, D., Prade, H.: Systems of linear fuzzy constraints. *Fuzzy Sets and Systems* 3, 37–48 (1980)
15. Dubois, D., Prade, H.: *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York (1980)
16. Zhang, J.: Interweave Neural Networks with Evolutionary Algorithms, Cellular Computing, Bayesian Learning and Ensemble Learning. In: Wang, S., Yu, L., Wen, F., He, S., Fang, Y., Lai, K.K. (eds.) *Business Intelligence: Artificial Intelligence in Business, Industry and Engineering*. 2009 International Conference on Business Intelligence and Financial Engineering, BIFE 2009, Beijing, China, July 24–26, pp. 65–68. IEEE Computer Society, Los Alamitos (2009)
17. Osaka, S., Gen, M.: Fuzzy shortest paths problem. *Computers and Industrial Engineering* 27(4), 465–468 (1994)
18. Osaka, S., Gen, M.: Order relation between intervals and its application to shortest path problem. *Computers and Industrial Engineering* 25(1), 147–150 (1994)

19. Katagiri, H., Sakawa, M., Ishii, H.: Fuzzy random bottleneck spanning tree problems using possibility and necessity measures. *European Journal of Operational Research* 152, 88–95 (2004)
20. Zhang, J.: Advances in fuzzy method for natural computing. In: Pan, J.-S., Liu, J., Abraham, A. (eds.) *Proceedings - 2009 9th International Conference on Hybrid Intelligent Systems, HIS 2009, Shenyang, China, August 12-14, vol. 2*, pp. 18–23. IEEE Computer Society, Los Alamitos (2009)
21. Chanas, S., Delgado, M., Verdegay, J., Vila, M.: Interval and fuzzy extension of classical transportation problems. *Transportation Planning and Technology* 17, 203–218 (1993)
22. Hussein, M.L.: Complete solutions of multiple objective transportation problems with possibilistic coefficients. *Fuzzy Sets and Systems* 93, 293–299 (1998)
23. Luhandjula, M.K.: Mathematical programming in the presence of fuzzy quantities and random variables. *The Journal of Fuzzy Mathematics* 11(1), 27–38 (2003)
24. Zhang, J.: The applications of fuzzy neural network. In: Yu, F., Liu, Y., Yue, G. (eds.) *2009 International Symposium on Intelligent Information Systems and Applications, Proceedings, IISA, Qingdao, China, October 28-30*. Academy Publisher, Finland (2009)
25. Bit, A.K.: Multi-objective fuzzy transportation problems. *The Journal of Fuzzy Mathematics* 6, 43–49 (1998)
26. Zhang, J.: Natural Computation for the Traveling Salesman Problem. In: *2009 2nd International Conference on Intelligent Computing Technology and Automation, ICICTA 2009, Changsha, Hunan, China, October 10-11, vol. 1*, pp. 366–369. IEEE Computer Society, Los Alamitos (2009)
27. Datar, M., Ranade, A.: Commuting with delay prone buses. In: *Proceeding of the Eleventh Annual ACMSSM Symposium on Discrete Algorithms*, pp. 22–29 (2000)
28. Fishman, G.S.: The distribution of maximum flow with application to multi-state reliability systems. *Operations Research* 35, 607–618 (1987)
29. Fishman, G.S.: Maximum flow and critical cutset as descriptors of multi-state systems with randomly capacitated components. *Computer Operations Research* 14, 507–520 (1987)
30. Zhang, J.: Applications of intelligent computation. In: Yan, W., Su, R. (eds.) *Proceedings - 2nd IEEE International Conference on Advanced Computer Control, ICACC 2010, March 27 - 29, vol. 4*, pp. 481–485. Institute of Electronics Engineers, Inc., China (2010)
31. Zhang, J.: On the non-monotone Armijo-type line search algorithm. In: *ICIC 2010 - 3rd International Conference on Information and Computing, vol. 3*, pp. 308–311 (2010)
32. Martin, J.: Distribution of time through a directed acyclic network. *Operations Research* 13, 46–66 (1965)
33. Klein, C.M.: Fuzzy shortest paths. *Fuzzy Sets and Systems* 39, 27–41 (1991)

A Hybrid Clustering Based on ACO and Single-Pass

Wen Xiong

Institute of Chinese Information Processing, Beijing Normal University,
Beijing, China

Stevens7979@sina.com

Abstract. Clustering is an important unsupervised machine learning (ML) method, and single-pass (SP) clustering is a fast and low-cost method used in event detection and topic tracing. To obtain a better result, SP is always needed to run multi-times in previous literatures due to it is dependent on the order of the instances appearance, which is not the global optimum. To tackle this situation, this paper proposed a hybrid clustering based on ant colony optimization (ACO) and SP, which exploits the global optimization ability of ACO and the superiority of SP, and takes the results of SP as a positive feedback implemented in ACO to improve the quality of the clustering. Experiments on two UCI datasets show that it is better than the multi-times running of the SP according to two different evaluation measures, which verifies its effectiveness and availability.

Keywords: Data mining (DM), swarm intelligence, ant colony optimization (ACO), single-pass (SP) clustering, topic detection and tracing (TDT).

1 Introduction

Clustering is an important task in data mining (DM), which is a sort of unsupervised learning method. When running on a set of instances, clustering methods are normally unknown about the category of each instance in the set and the number of categories. It completes automatically classification by minimizing the similarity of instances among different groups and maximizing that in the same group.

There are many clustering methods in the previous literatures, which can be divided into six categories according to their characteristics such as clustering based on partitioning, based on density, based on subspace, based on hierarchy, based on grid, and based on hybrid. The hybrid method can possess of the superiorities of different methods since it utilizes the mechanisms derived from several methods, so it is very important for us to investigate the hybrid method using multi-method.

Single-pass (SP) clustering is a sort of method based on partitioning. On the one hand, the SP has a very high efficiency since that it needs to scan once a set of instances to build the final clusters, so it is usually used for very large data set to attain quickly the result of clustering. On the other hand, the result of the SP is normally not a global optimum since that the SP is dependent on an order of the instances scanned, so in practice, researchers run often SP multi-times and take the best one as a final result.

Each clustering method has its corresponding evaluation function which is used to evaluate results of clustering, or to guide the generation of clusters as a heuristic info. Considering the use of evaluation function, the problem of clustering can be regarded as a problem of combination optimization which is a NP hard problem, and can be resolved with relative small computational cost to obtain a global or nearly global optimal.

There are some intelligence optimization methods in previous literatures. Genetic algorithm (GA) handles the problem of combination optimization with simulating the evolution procedure of biology, in which the chromosomes cross and mutate to create the individual with a better adaptive degree. Simulate anneal (SA) solves efficiently the problem by emulating the cooling procedure of metal from high temperature to low, in which the inner structure of metal has gradually been optimized.

Swarm based optimization is a sort of popular method recently, including particle swarm based, ant colony based, bee colony based, and fish school based, etc. Ant colony based methods utilize the mechanisms of the shortest path for ants searching food and the pheromone interacting with environment to complete the task of optimization. At the same time, there are some clustering methods by simulating the behavior of ants clearing their nest.

Considering that SP clustering needs to run multi-times to obtain a better result in practice, this paper proposes a novel hybrid clustering method based on ants colony and SP (CASP), which integrates the superiorities both high efficiency of SP and small computational cost of ant colony optimization (ACO) in the problem of combination optimization by utilizing the middle result of multi-running of SP feeding back to the method controlled by the ACO.

The remainder of the paper is organized as: in section 2, we introduce the related research; in section 3, we describe the proposed hybrid clustering method; in section 4, we report the experiments and the results using the method on two datasets of UCI, and compare the results with that of multi-times running of SP; finally, we conclude the work of the paper in section 5.

2 Related Work

To cluster large amount of distributed data in a real-time with low-cost clustering algorithms for streaming data, Gupta and Grossman proposed a single-pass generalized incremental algorithm (GenIc) [1]. Klampanos, et al. experimented on two large web-based testbeds showing that the order of document does not practically matter for document clustering within the information retrieval (IR) [2]. Allah et al. presented a new SP clustering algorithm [3], which maps the data into low-dimensional feature space. Experiments show its availability.

Lukasz Machnik proposed a two-phase document clustering based on ACO [4]. The experimental results show that the quality of ACO clustering is very high. Asharaf and Murty proposed an adaptive rough fuzzy SP algorithm (ARFL) [5]. Experiments show that the ARFL algorithm gives the best performance.

Papka and Allan proposed an on-line new event detection using SP clustering [6]. Experiments show its effectiveness. For web name disambiguation, Iosif proposed an unsupervised method [7]. The evaluation score was found to be equal to the highest official score.

3 Method

Since the number of clusters created by SP can not be decided beforehand, the method creates a certain number of clusters by appointing a maximum threshold of distance. In this paper, we use Euclid distance and centroid as the prototypes of clusters, which are expressed as follows:

$$distance(x, y) = (\sum_{i=1}^M (x_i - y_i)^2)^{1/2} \quad (1)$$

Where x and y are instances, whose dimension is M .

$$c_i = \sum_{j=1}^{|C_i|} x_j / |C_i|, \quad i = 1 \dots K \quad (2)$$

Where c_i is the centroid of i -th cluster, and K is the number of the clusters. When we add a new instance into a previous cluster, we need to update the center of the cluster, which is expressed as follows:

$$c_{ij}^* = (c_{ij} \cdot |C_i| + x_{ij}) / (|C_i| + 1), \quad i = 1 \dots K, \quad j = 1 \dots M \quad (3)$$

3.1 Ant Colony Optimization

When natural ants are searching food, they can always find a shortest path between food and their nest. The mechanism obtained from biology evolution inspires the biologist and experts in the artificial intelligence fields. By simulating the behavior of ant colony, Dorigo M. et al. proposed ant colony optimization in 1992 [8], in which artificial ants consider not only the distance as a heuristic info, but also the pheromone as a accumulative effect when they select the next node which is a part of entire solution. The probability of the node selection can be expressed as follows:

$$P_{ij} = (\tau_{ij})^\alpha \cdot (\eta_{ij})^\beta / \sum_{l \in \text{Node_left}, l \neq j} (\tau_{il})^\alpha \cdot (\eta_{il})^\beta \quad (4)$$

Where P_{ij} is the transfer probability from node i to j , and τ_{ij} is the pheromone at position (i, j) and η_{ij} is the heuristic info. α and β are the weighty parameters. After ants select all nodes, ACO evaluates the solution using an evaluation function, and updates the pheromone matrix according to the evaluation value. In the meantime, as the situation in a real environment, the pheromone will volatilize along with the time passing. The mode of pheromone update can be expressed as follows:

$$\tau_{ij} = (1 - \rho) \cdot \tau_{ij} + \rho \cdot \tau_0 / length(path) \quad (5)$$

Where τ_0 is an initial pheromone, and ρ is a factor for pheromone volatilization and $length(path)$ is the evaluation value of the solution.

3.2 A New Hybrid Clustering Method Based on ACO and SP

Combining the high efficiency of SP and global optimization ability of ACO, this paper proposes a hybrid clustering method based on ACO and SP. When deciding the

cluster which the current instance will join, the method uses not only the distance between the current instance and the center of clusters, but also the accumulative average value of pheromone between the current instance and all instances in the cluster, which can be expressed as follows:

$$P(x_i, c_j) = \frac{(\sum_{l=1}^{|C_j|} \tau_{il} / |C_j|)^\alpha \cdot (\text{distance}(x_i, c_j))^\beta}{\sum_{l \in \text{Node_left}, l \neq j} (\sum_{l=1}^{|C_{l1}|} \tau_{il} / |C_{l1}|)^\alpha \cdot (\text{distance}(x_i, c_{l1}))^\beta} . \quad (6)$$

Thus the proposed method utilizes the positive feedback of ACO compared to the previous SP clustering. The pheromone update in the method according to (5), where the *length (path)* can be expressed as follows:

$$\text{length}(\text{path}) = \begin{cases} 1 - f_0 / 100, & \text{if } (Eval_l \geq Eval_g) \\ 1 - |Eval_l| / |Eval_g|, & \text{if } (|Eval_l| \leq |Eval_g|) \\ 1 - |Eval_g| / |Eval_l|, & \text{otherwise} \end{cases} . \quad (7)$$

Where $Eval_l$ is the current evaluation value and $Eval_g$ is the global optima, and f_0 is a positive constant value less than 100.

The pseudo code of CASP is expressed as follows:

CASP (*max-distance*, τ_0 , α , β , ρ , *max-iteration*)

A. Initialize the pheromone matrix of ants using the τ_0 (an initial value of pheromone).

B. Loop *max-iteration* times

B.1 Select arbitrarily an instance from the data set, and take the instance as a new cluster and the center of the cluster.

B.2 Select arbitrarily an instance from the instances left in the data set, and calculate the distances between the instance and all centers of clusters. If the minimum distance is greater than the *max-distance*, take the instance as a new cluster and the center of the cluster, then go to step B.5.

B.3 Take the distance as heuristic info, and the accumulative average value of pheromone between the current instance and all instances in the cluster as another factor to calculate the selection probability of the selected center of cluster according to (6).

B.4 Decide which cluster the current instance is belong to using roulette method, and insert the instance into the cluster and update the center of the cluster according to (3).

B.5 If there are instances left, go to step B.2.

B.6 Evaluate the clustering result using an evaluation function, and handle the volatilization of pheromone and update the pheromone matrix according to (5) and (7).

B.7 If the current evaluation value is better than the global, update the global value and solution.

4 Experiments

To examine the validity of the proposed CASP method, we used the previous SP clustering method as a baseline method, and ran the baseline method multi-times and took the best result as a reference value. The experimental data sets were taken from UCI datasets [9] including *wine* and *iris* whose characteristics are summarized as follows:

Iris: class = 3; feature = 4; sample = 150;

wine: class = 3; feature = 13; sample = 178.

We used an evaluation measure in [10], which is determined as follows:

$$\begin{aligned} Eval &= BGD - WGAD \\ BGD &= \sum_{i=1}^K dist(c_i, c) \\ WGAD &= \sum_{i=1}^K \sum_{j=1}^{|C_i|} dist(x_j, c_i) / |C_i| . \end{aligned} \quad (8)$$

Where c is a center of all clusters, we also used another evaluation measure for comparing, which is listed as follows:

$$DIFF = \frac{2 \cdot \sum_{i=1}^K dist(c_i, c)}{k \cdot (k-1)} - \frac{\sum_{i=1}^K \sum_{j=1}^{|C_i|} dist(x_j, c_i)}{k} . \quad (9)$$

We tried CASP method several times, and selected the experimental parameters as follows:

Iris_Eval: $\tau_0 = 2.5$; $f_0 = 95$; $\rho = 0.2$; $\alpha = 20$; $\beta = 2$;

Iris_DIFF: $\tau_0 = 5.5$; $f_0 = 95$; $\rho = 0.7$; $\alpha = 2$; $\beta = 30$;

Wine_Eval: $\tau_0 = 1.5$; $f_0 = 95$; $\rho = 0.1$; $\alpha = 2$; $\beta = 30$;

Wine_DIFF: $\tau_0 = 1.5$; $f_0 = 95$; $\rho = 0.1$; $\alpha = 2$; $\beta = 30$;

For the *iris* data set, we ran SP and CASP methods from the threshold 0.05 to 4.2 with step 0.2, and 324 times for each threshold. As for the *wine* data set, we ran them from the threshold 10 to 180 with step 3.5, and 100 times for each threshold. To obviate the random factor, we ran CASP five times for each data set, and averaged the appearing times which listed as follows.

Table 1. The average appearing times of maximum according to the number of clusters.

Dataset	Type	CASP>SP	CASP<SP	CASP=SP	!CASP	!SP
iris	<i>Eval</i>	30	14.6	2	2.4	3.8
	<i>DIFF</i>	23.2	19.6	3	1.2	4.2
wine	<i>Eval</i>	48.2	34.4	0	2.4	3.4
	<i>DIFF</i>	43.8	37.8	0.2	2.2	4

Where “*CASP>SP*” indicated that maximum values of *CASP* is greater than that of *SP* in a certain number of the clusters (e.g. clusters = 10, 13, etc.), and “*!CASP*” indicated that the results of clusters of *CASP* are not formed in a certain number of the clusters compared to that of *SP* (and “*!SP*” correspondingly).

From table 1, we found that the *CASP* method has more maximum values according to the number of clusters than *SP* method, which has a significant time.

5 Conclusions

The experimental results indicates that *CASP* is superior to the multi-times running of the *SP* since the *CASP* utilizes the clustering results of the single-pass as a feedback value and the global optimization capability of *ACO*, which shows its value. How to apply the *CASP* to text clustering and textual TDT would be our future work.

Acknowledgments. The paper is supported by "the Fundamental Research Funds for the Central Universities". We would like to thank our anonymous reviewers for their comments.

References

1. Gupta, C., Grossman, R.L.: GenIc: A Single-Pass Generalized Incremental Algorithm for Clustering. In: *SDM* (2004)
2. Klampanos, I.A., Jose, J.M., Rijsbergen, K.V.: Single-Pass Clustering for Peer-to-Peer Information Retrieval: the Effect of Document Ordering. In: *Infoscale* (2006)
3. Ataa- Allah, F., Grosky, W.I., Aboutajdine, D.: On-Line Single-Pass Clustering Based on Diffusion Maps. In: Kedad, Z., Lammari, N., Métais, E., Meziane, F., Rezgui, Y. (eds.) *NLDB 2007*. LNCS, vol. 4592, pp. 107–118. Springer, Heidelberg (2007)
4. Machnik, L.: Documents Clustering Method Based on Ants Algorithm. In: *International Multiconference on Computer Science and Information Technology*, pp. 123–130 (2006)
5. Asharaf, S., Murty, M.N.: An Adaptive Rough Fuzzy Single Pass Algorithm for Clustering Large Data Sets. *Pattern Recognition* 36, 3015–3018 (2003)
6. Papka, R., Allan, J.: On-line New Event Detection Using Single Pass Clustering. *UMASS Computer Science Technical Report*, pp. 98–21 (1998)
7. Iosif, E.: Unsupervised Web Name Disambiguation Using Semantic Similarity and Single-Pass Clustering. In: Konstantopoulos, S., Perantonis, S., Karkaletsis, V., Spyropoulos, C.D., Vouros, G. (eds.) *SETN 2010*. LNCS, vol. 6040, pp. 133–141. Springer, Heidelberg (2010)
8. Dorigo, M.: Optimization, Learning and Natural Algorithms. Ph.D. Thesis, Politecnico di Milano, Italy (1992) (in Italian)
9. Frank, A., Asuncion, A.: UCI machine learning repository. University of California, School of Information and Computer Science, Irvine, CA (2010), <http://archive.ics.uci.edu/ml>
10. Chaimontree, S., Atkinson, K., Coenen, F.: Best Clustering Configuration Metrics: Towards Multiagent Based Clustering. In: Cao, L., Feng, Y., Zhong, J. (eds.) *ADMA 2010, Part I*. LNCS(LNAI), vol. 6440, pp. 48–59. Springer, Heidelberg (2010)

Advanced Sensing Device for Gesture Recognition*

Yong Mu Jeong¹, Ki-Taek Lim¹, and Seung Eun Lee^{2,**}

¹ SoC Research Center, Korea Electronics Technology Institute Sungnam-si,
Gyeonggi-do, Korea

{ymjeong, limkt}@keti.re.kr

² Dept. of Electronic & Information Technology
Seoul National University of Science and Technology, Seoul, Korea
seung.lee@seoultech.ac.kr

Abstract. In this paper, we introduce a gesture recognition platform for a realistic game. The realist game experience is realized by utilizing four acceleration sensors interacting with a game server. We demonstrate the efficiency of the propose system by controlling an avatar in the game with gestures only.

Keywords: Gesture Recognition, Game System, User Experience.

1 Introduction

Over the last decade, user experience becomes a major issue that is how to support more natural and immersive user interface. The user interface has evolved from simple devices such as a keyboard and a mouse, to smart equipments employing gesture recognition, motion tracking, and speech recognition. In this paper, we introduce a gesture recognition platform for a realistic game. The realist game experience is realized by utilizing four acceleration sensors interacting with a game server. We demonstrate the efficiency of the propose system by controlling an avatar in the game with gestures only.

The rest of paper is organized as following: Section 2 describes the architecture of proposed platform including description of ASD module and experimental result. We conclude in section 3 by outlining the direction for the future work on this topic.

2 The Proposed System

The proposed system for the realist game environment utilizes four advanced sensing devices (ASDs), and communication module (see Fig. 1). Zigbee module establishes the communication channels between the game server and ASD. The gesture recognition and management of interaction between recognized gesture and game server are realized in software. The gesture recognition module consists of aggregator, signal classifier, sequence tracker and gesture recognizer. The interaction manager controls the game server' I/O.

* This research was supported by R&D Program of MKE(Ministry of Knowledge Economy) in korea, [10032108, Development of Multiverse Game Platform Technology with Multimodal Interaction].

** Corresponding author.



Fig. 1. Overview of the proposed system

2.1 Advanced Sensing Device (ASD)

In the proposed system, motion recognition is made by obtaining the accelerator values of user’s motion with the advanced sensing device. We use Em357 [1] as a main processor of ASD. The Em357 combines a 2.4GHz IEEE 802.15.4 radio transceiver with 32-bit microprocessor, flash memory and RAM. In order to obtain acceleration value, we embedded the MMA7260QT [2] in ASD. The MMA7260QT is a low cost capacitive micromachined accelerometer that features signal conditioning, a 1-pole low pass filter, and temperature compensation. PCB antenna is used for cost-effectiveness (see Fig. 2). It operates with 160mAh Li-polymer battery.

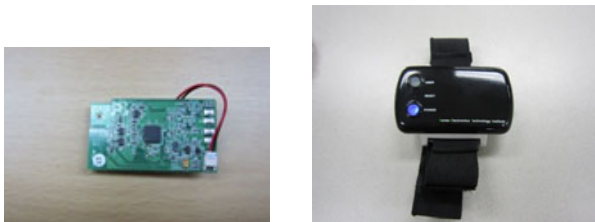


Fig. 2. PCB board and mockup case of the proposed ASD

2.2 Gesture Recognition Module

Gesture Recognition module classifies user’s gesture based on collected sensor data and transmits the result to the interaction manager. The classification is completed by comparing feature vector with predefined DB. In order to match input feature vectors, we use a brute force match algorithm that exhaustively compares a pair of feature vectors from each gesture based on the Manhattan (a.k.a. L1) distance. It should be noted that several other match algorithms are available, but the brute force match is simple to implement and has sufficient accuracy for the usage model of interest.

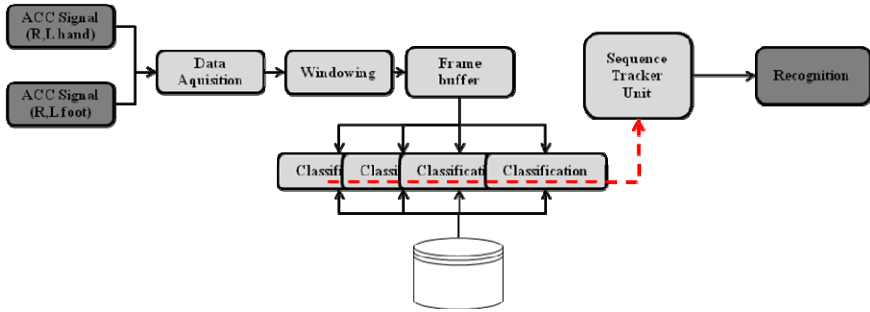








Fig. 3. Diagram of gesture recognition module

2.3 Experimental Results

We set up the metaverse server by using opensim [3] and used hippo viewer [4] as a client. Interaction manager provide the way to control the character in the server based on the result from the gesture recognition module. We defined the six gestures as shown in Table 2.

The recognized gesture information is translated to the command for the client, hippo viewer, through the interaction manger, realizing the control of the avatar in the realist game only using the gestures. The accuracy of the gesture recognition is

Table 2. Gestures definition for the nine commands to metaverse game

Gesture	Command	Gesture	Command	Gesture	Command
	Left Hand up		Right Hand up		forward
Gesture	Command	Gesture	command	Gesture	Command
	Backward		Jump		Fly

suitable to experience the realistic game. The sensor data is sampled in 25 msec interval and transmitted to the host PC after packing 4 consecutive data into a packet. The response time of the gesture recognition time was less than 150msec which is sufficient latency to enjoy the game. The maximum power consumption including the LED light for debug purpose was 35mAh.

3 Conclusions

In this paper, we proposed the gesture recognition platform for the realistic game. We accomplished to gaining full control of the avatar in the realistic game by using the body gestures, enhancing the user experience in the game platform. The future work in this area is as follows. We plan to improve the gesture recognition time for the real-time synchronization between a user and an avatar. We also plan to establish bi-direction communication channel to feedback the experience of the avatar to the game player. We expect that enhancing the user experience on virtual world such as realist game will bring forth a new spectrum of novel usage models for smart devices, bio-medical equipments, entertainment system for automobile and robotics.

Acknowledgement. This research was supported by R&D Program of MKE (Ministry of Knowledge Economy) in Korea, [10032108, Development of Multiverse Game Platform Technology with Multimodal Interaction].

References

1. Ember em357, http://www.ember.com/products_zigbee_chips_e300series.html
2. Freescale MMA7260QT, http://www.freescale.com/webapp/sps/site/prod_summary.jsp?co-de=MMA7260QT
3. Opensim, http://opensimulator.org/wiki/Main_Page
4. Hippo viewer, <http://mjm-labs.com/viewer/>

Research on the Reactive Power Control of Distributed Photovoltaic System*

Zhiqiang Jia and Weiping Luo**

Wuhan Textile University, Wuhan, China

Abstract. In this paper, distributed power generation systems with multiple reactive power distribution network optimizations. Through the grid of the inverter output voltage in d, q axis component of decoupling, to achieve the active and reactive independent regulator, to take full advantage of the distributed generation of reactive power compensation capacity, improve the operating performance of distribution network. Established based on genetic algorithm optimization of distributed generation system reactive power control model, give full consideration to the net loss minimum and the node voltage constraints on the grid of the inverter reactive power to be optimized, with a strong convergence and fast.

Keywords: Distributed generation, reactive power control, genetic algorithms.

1 Introduction

Distributed generation generally refers to a few kilowatts of generating power to hundreds of megawatts of small, modular, distributed, arranged in the vicinity of the users' efficient and reliable power generation unit^[1]. And here we talk about is the need for the use of photovoltaic technology in general and with the power inverter as an interface, a distributed power supply system with grid connection and energy exchange between.

This paper studies the grid inverter active and reactive decoupling control method, and then net loss minimization objective, the use of genetic algorithms and grid-connected inverter reactive power to be optimized to achieve a distributed power system reactive power optimization of control.

2 Distributed Power Generation System and Grid-Connected Inverter Reactive Power Control Strategy

Three-phase voltage inverter with the mathematical model is based on three-phase voltage and inverter topology, the three-phase stationary coordinate system (a, b, c) the use of the basic law of the circuit (Kirchhoff's voltage current law) of the voltage inverter established general mathematical description of three-phase voltage and inverter topology shown in Figure 1.

* The Project is Supported by Hubei Provincial Department of Education (D20081705).

** Corresponding author.

But in this mathematical model, the AC side is time-varying exchange capacity, and therefore not conducive to the control system design. Therefore, through coordinate transformation the three phase stationary coordinate system into the power of the fundamental frequency synchronous d-q rotating coordinate system.

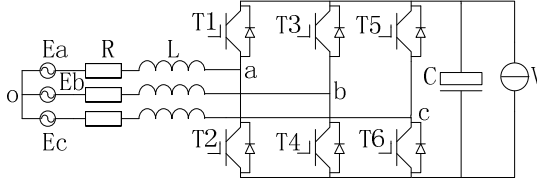


Fig. 1. Topology diagram of grid-connected inverter

Based on three-phase voltage-type PWM inverter d-q model, control system design generally use double-loop control. Because we want to achieve independent control of reactive components, typically selected q-axis EMF vector E coincides with the grid, so that the reactive q-axis component of the reference value, d-axis active component reference. As a result of synchronous rotating coordinate system control structure, symmetrical three-phase AC voltage and AC current transformation to the synchronous rotating coordinate system into direct traffic, so the current inner loop with PI regulator can adjust the static error-free. In order to achieve independent control of reactive components, the grid electromotive force space vector and q-axis orientation, the

$$e_q = E_m \quad e_d = 0 \tag{1}$$

E_m is a power force peak.

Three-phase voltage inverter in the synchronous rotating coordinate system the mathematical model:

$$\begin{aligned} L \frac{di_q}{dt} &= E_m - \omega L i_d - R i_q - v_{dc} s_q \\ L \frac{di_d}{dt} &= \omega L i_q - R i_d - v_{dc} s_d \end{aligned} \tag{2}$$

As a result, the availability of grid-connected inverter control shown in Figure 2.

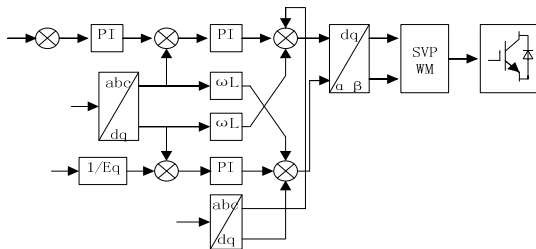


Fig. 2. Control diagram of grid-connected inverter

Using MATLAB software above control strategy is simulated. The use of three-phase voltage inverter with active and reactive power can be independently controlled to regulate the system of reactive power, the DC voltage, AC current waveform of active power and less impact, and can fast recovery steady-state.

3 Distribution Network with Distributed Generation of Reactive Power Optimization

With distributed generation of reactive power optimization of distribution network that distributed power to issue non-reactive power control to reduce system loss and ensure active network node voltage within the specified range. The mathematical model includes the objective function, the trend equation equality constraints and inequality constraints in three parts.

3.1 The Objective Function

Select the smallest objective function for the net loss, that is

$$\min F = P_{loss} = \sum_{i=1, j \in H}^n G_{ij} (U_i^2 + U_j^2 - 2U_i U_j \cos \theta_{ij}) \quad (3)$$

3.2 Current Equation Equality Constraints

Each node active and reactive power balance constraints are as follows:

$$\begin{aligned} P_i &= U_i \sum_{j \in H} U_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \\ Q_i &= U_i \sum_{j \in H} U_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \end{aligned} \quad (4)$$

In the formula (3),(4): P_{loss} is a net loss f; P_i, Q_i to node i into the active and reactive; U_i, U_j to i, j node voltage amplitude; $G_{ij}, B_{ij}, \theta_{ij}$ respectively i, j between the nodes of the conductance, susceptance and voltage phase angle difference; H that all the nodes i directly connected nodes.

3.3 Inequality Constraints

With reactive power optimization problem in variables can be divided into control variables and state variables. They must meet the conditions:

Control variables need to be met:

$$Q_{gi} \min \leq Q_{gi} \leq Q_{gi} \max \quad (i=1, 2, \dots, n) \quad (5)$$

Including $Q_{gi} \max, Q_{gi} \min$ for the reactive power compensation capacity of the upper and lower limits; n to access the distributed power of the number of units; State variables need to be met:

$$U_i \min \leq U_i \leq U_i \max \quad (i=1,2,\dots,n) \quad (6)$$

Including $U_i \max, U_i \min$ is the node voltage amplitude of the upper and lower limits; m is the number of load nodes; s on behalf of balancing node.

3.4 Optimization

Distributed generation system reactive power optimization is a multi-variable, complex nonlinear optimization problem, so this paper genetic algorithm can be better optimization results. Genetic algorithm to solve the problem of reactive power optimization steps: First, a set of randomly generated initial current solution, and then encode, through selection, crossover and mutation operations to regroup, and finally to become the best solution. For the "Select" operation, this paper is the roulette wheel method. Each individual probability of being selected for the

$$F(x_i) = \frac{f(x_i)}{\sum_{i=1}^N f(x_i)} \quad (7)$$

$x_i, f(x_i)$ is the individual fitness, the higher the fitness value, the greater the likelihood of being selected.

"Cross" operation is the population of the two individuals according to equation (8) shows a linear formula.

$$\begin{aligned} C_1 &= p_1 \bullet a + p_2 \bullet (1 - a) \\ C_2 &= p_1 \bullet (1 - a) + p_2 \bullet a \end{aligned} \quad (8)$$

Type (8) p_1, p_2 for the two populations of the Central Plains and some individuals, a is 0 to 1 in the random variable c_1, c_2 to produce two new individuals.

"Mutation" operation is the population of individuals randomly assigned to the locus of variation points to pre-set mutation probability values for these loci to change in order to achieve individual variation. Specific flow chart is shown in Figure 3.

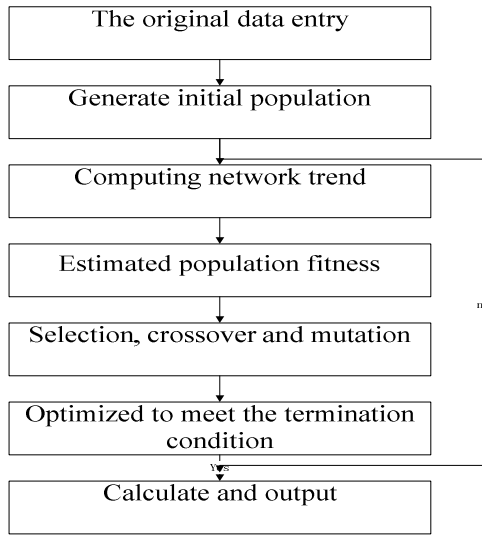


Fig. 3. Flow chart of genetic algorithm

4 Analysis of Cases

Example 1: 20-bus case

The examples of pure chain network, the network structure shown in Figure 4.

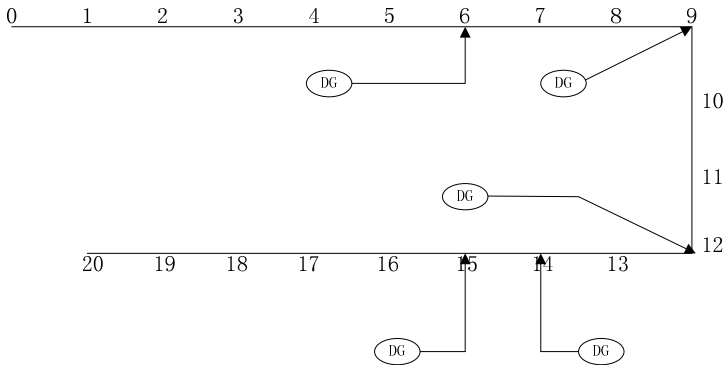


Fig. 4. 20 node network structure with distributed generation

The distribution network includes 20 nodes^[4], where node 0 to balance the node, other nodes to PQ node, five distributed generation access location (node 6,9,12,14,15). This study is the use of genetic algorithm parameters: crossover probability 0.85, mutation probability of 0.001, the size of the population each generation of 20 individuals, the number of iterations to 120.

Table 1. Control variables' value after optimization

Variable Name	Initial state	Optimization results
Q6	0	2.6457
Q9	0	1.6299
Q12	0	2.9528
Q14	0	1.7953
Q15	0	1.9606

5 Conclusion

This chapter examines the distribution of power with more reactive power distribution network optimization model and algorithm research, to identify a network of loss minimization objective function, the flow equations as equality constraints bound to distributed generation of reactive power to control variables, and to construct the penalty function to solve the genetic algorithm optimization. And through several examples of simulation of the genetic algorithm is feasible. It has a strong and fast convergence, for reducing network power losses and improving voltage quality to a certain extent.

References

1. Liang, C.-H., Duan, X.-Z.: Distributed Generation and Its Impact on Power System. *Automation of Electric Power Systems* 25(12), 53–56 (2001)
2. Su, L.: Reactive Power Optimization of Power System Based on Improved Genetic Algorithm, Thesis. Southwest Jiaotong University, Chengdu (2006)
3. Ning, Y., Qian, K.-J., Yuan, Y.: Impacts of Distributed Generation on Distribution System. *Automation of Electric Power Systems* (2005)
4. Wang, Z.-Q., Zhu, S.-Z., Zhou, S.-X., et al.: Impacts of Distributed Generation on Distribution System Voltage Profile. *Automation of Electric Power Systems* 28(16), 56–60 (2004)

Application of Gateway in Fault Location System of Belt Conveyor

Shaojie Yin and Jinsheng Sun

College of Information Engineering, Hebei United University,
Tangshan 063009, China
Shaojieyin@163.com, sunjinsheng@heuu.edu.cn

Abstract. Fieldbuses are more widely used in industry control system, but there does not have the standardized unique protocol among them. Thus, the different fieldbuses can not communicate directly in the same network. To solve the problem, we proposed an implement method of gateway between Profibus-DP and CAN based on STM32, and applied it in fault location system of belt conveyor. This paper emphatically describes the operation principle of the fault location system, the hardware and software design method of the gateway. The practical result shows that the gateway works stably and reliably in this system and has a high application value.

Keywords: Belt conveyor, fault location, gateway, Profibus-DP, CAN, STM32.

1 Introduction

Belt conveyor is a major and common logistic equipment. It is widely used in coal mines, metallurgy and other industrial fields as the important means of material transportation. Nevertheless, Owing to poor working conditions, belt conveyors often have some faults during operation, such as tearing, aberrancy, skid, etc. With the increase of production capacity, belt conveyors tend to be in large scale. As a result, it is a time-consuming task to locate the fault position, which may bring great economic losses to enterprises. In order to solve the problem, the fault location system of belt conveyor is proposed. The system supervises conveyor's work condition. Once the conveyor works abnormally, the system will send out a warning and locate the fault position immediately so as to ensure the maintenance of conveyor in time.

With the rapid development of technologies of computer, data communication and control automation, FCS (field bus control system) technology is more and more applied in industry control system. At present there are dozens of different kinds of fieldbuses worldwide. Among those, CAN bus and Profibus are the most widely applied fieldbuses. The CAN communication protocol efficiently supports distributed real-time control with a very high level of security and its communication distance could reach 10km. It may satisfy most industrial control field's request. Profibus is recognized as the fastest fieldbus and is an open standard, based on the European E50170 Electrical Specification. It is convenient for applying in industry control system. Nevertheless, owing to the two communication protocols are incompatible, it is

difficult for them to coexist in a system. To solve the above problem, an implement method of gateway between Profibus-DP and CAN based on STM32 is proposed, and it is applied in the fault location system of belt conveyor.

2 Fault Location System of Belt Conveyor

The fault location system has been used CAN bus and Profibus-DP. In order to combine the two different fieldbuses, designing a gateway between them is the major task of the paper. The gateway serves as an intelligent node in CAN bus network, while it serves as an intelligent slave station in Profibus-DP network. Figure 1 is the fault location system structure.

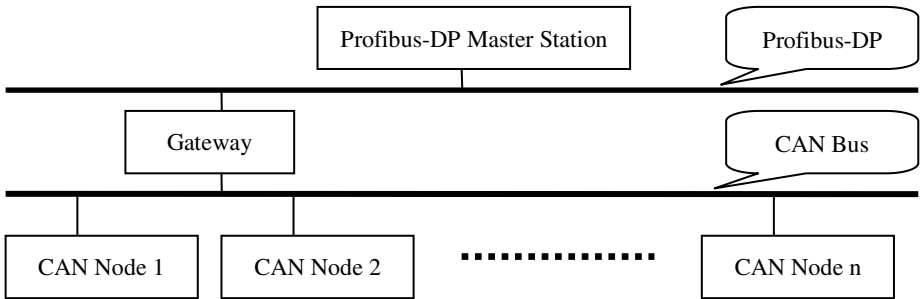


Fig. 1. Fault location system structure

In the system, several CAN bus intelligent measurement nodes locate along the belt conveyor. Each of them could detect 4 kinds of transducers' data. Because any node on CAN bus may start to transmit a message, if any of those nodes detects a fault, it will transmit the fault location message to the gateway through CAN bus without delay. The gateway receives and stores the fault information. When Profibus-DP master station establishes communication with the gateway, which serves as a slave station, the gateway will transmit the information to the master station so as to locate the fault position.

3 Hardware Design of the Gateway

The gateway include CAN bus interface module and Profibus-DP interface module. The overall block diagram as illustrated in figure 2. The main chip of the gateway is STM32F103VE, which is produced by STMicroelectronics. STM32F103VE incorporates the high-performance ARM Cortex™-M3 32-bit RISC core operating at a 72 MHz frequency, high-speed embedded memories (Flash memory up to 512 Kbytes and SRAM up to 64 Kbytes), and meets completely the gateway's requirement for data memory and program memory. It has standard and advanced communication interfaces: two I2Cs and SPIs, three USARTs, an USB and a CAN.

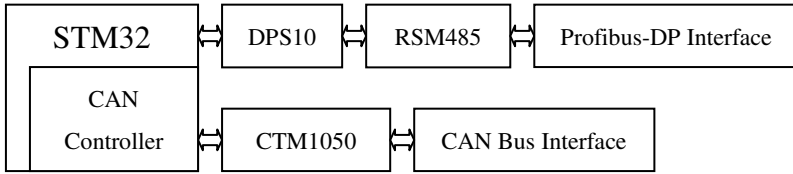


Fig. 2. The overall block diagram of hardware

3.1 CAN Bus Interface Module

The CAN bus interface module includes CAN controller (the basic extended peripheral of STM32) and the high-speed CAN transceiver CTM1050. It is used to receive the fault information that intelligent measurement nodes transmit.

The CAN controller supports the CAN protocols version 2.0A and B. As a result, the hardware design of CAN interface is simplified. Bit rate of it is up to 1 Mbps. It has three transmit mailboxes, two receive FIFOs, 14 scalable filter banks and has been designed to manage a high number of incoming messages efficiently with a minimum CPU load. There have a CAN transceiver CTM1050 between CAN controller and the physical bus. It is not only an interface between them but also has other functions, such as power isolation, electrical isolation, CAN bus protection, etc.

3.2 Profibus-DP Interface Module

The Profibus-DP interface module includes DPS10 (the Profibus-DP slave station interface) and Profibus-DP transceiver RSM485P. It is used to transmit the fault information from gateway to Profibus-DP master station. Figure 3 is the Hardware design of Profibus-DP interface.

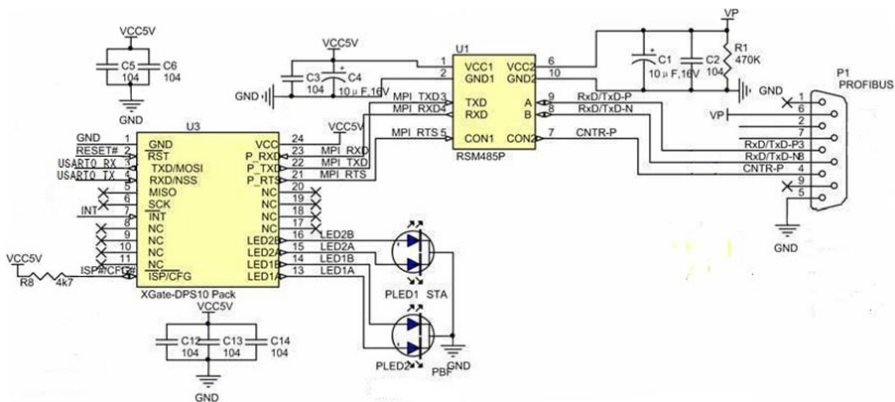


Fig.3. Hardware design of Profibus-DP interface

DPS10 is used to assist users to design Profibus-DP intelligent slave stations. It uses DIP24 package, and is small in size, low in cost and high in reliability. Bit rate of it could be up to 12 Mbps. Because DPS10 supports the Profibus-DP protocol and the communication with master station is controlled by the integrated state machine, STM32 need not to process the state machine. STM32 only need to initialize DPS10 and exchange data with DPS10, and then the gateway could communicate with the Profibus-DP master station.

STM32 and DPS10 communicate with half-duplex mode through USART0 interface and support at most 244Byte input/output data. RSM485P is the Profibus transceiver. It is the interface between DPS10 and the physical bus, and also has other functions, such as power isolation, electrical isolation, Profibus protection, etc.

4 Software Design of the Gateway

The software design of CAN-Profibus gateway uses the STM32 firmware library in the MDK environment for debugging and simulation. It is used to receive the fault information that CAN bus intelligent measurement node transmitted, transform frame's format and transmit it to Profibus-DP master station. It includes CAN protocol handling module and main program.

4.1 CAN Protocol Handling Module

The CAN protocol handling module is used to receive the fault information and store it to CAN receiving buffer. Because the CAN protocol only uses the lower two layers (Physical Layer and Data Link Layer) of the Open Systems Interconnection (OSI) Net Work Layering Reference Model, when we apply CAN protocol, an Application Layer protocol should be established by ourselves. In this system, the CAN protocol frame uses the standard format. Figure 4 is the standard format of data frame.

SO F	Arbitration Field		Control Field			Data Field	CRC Field	ACK Field			EOF	
	Identifier		R T R	I D E	R 0	DLC	CRC Sequence					
1 Bit	11 Bits		1 Bit	1 Bit	1 Bit	4 Bits	0-8 Bytes	15 Bits	1 Bit	1 Bit	1 Bit	7 Bits

Fig. 4. Standard format of data frame

The most important thing in Establishing Application Layer protocol is defining Arbitration Field, Control Field and Data Field of the Data frame. This paper defines the three fields based on the requirements of the fault location system.

The Arbitration Field consists of Identifier and Remote Transmission Request (RTR) Bit. The Identifier's length is 11 bits, from ID10 to ID0. In the system, ID10-ID4

represents the source address of the fault information. Every CAN bus intelligent measurement node has a different 7-bits address. ID3-ID0 represents the target address, which is the gateway's address. Its address is 1010. In Data Frames, the RTR Bit has to be "dominant". In Remote Frames, the RTR Bit has to be "recessive".

The Control Field consists of six bits. It includes the 4-bits Data Length Code (DLC) and two bits reserved for future expansion. The length of the Data Field is indicated by the DLC.

The Data Field consists of the fault information. Because each of CAN bus node could detect 4 kinds of transducers' data, the Data Field contains 4 bytes. Each byte represents a kind of fault information.

The CAN controller of STM32's initialization can be done while the hardware is in Initialization mode. The initialization includes set up operating mode, baud rate, filters, interrupts, etc. When an intelligent measurement node detects fault information, it will transmit the message to the gateway in the data frame. CAN controller of STM32 will detect a data frame in CAN bus. Then it will filter identifier of the frame. If the frame passes through the identifier filtering, CAN controller will store it in FIFO. Once a message has been stored in the FIFO, the FMP [1:0] bits are updated and an interrupt request is generated if the FMPIE bit in the CAN_IER register is set. The module accesses the message and stores it in CAN receiving data buffer in the interrupt. Then the software must release the FIFO by setting the RFOM bit, so that it is free to store the next message. Figure 5 is the Flow chart of CAN bus reception handling.

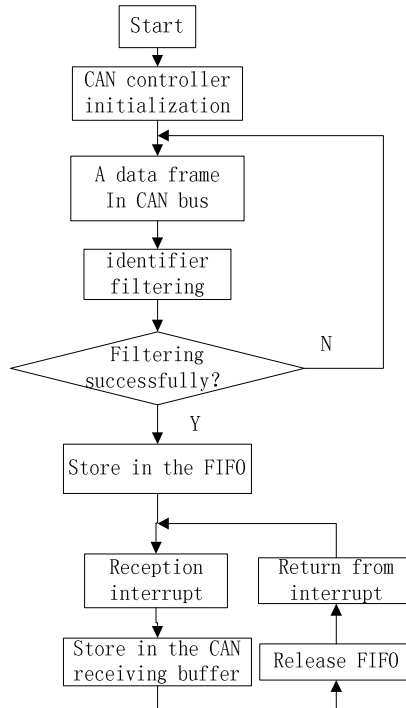


Fig. 5. Flow chart of CAN bus reception handling

4.2 Main Program

Main program is used to initialize the system, transform data frame’s format and transmit it to Profibus-DP master station. System initialization includes setting up clocks, general-purpose I/O (GPIO) ports, nested vectored interrupt controller (NVIC), universal synchronous asynchronous receiver transmitter (USART), CAN controller, DPS10,etc.

STM32 and DPS10 communicate with half-duplex mode through USART0 interface. STM32 accesses DPS10 by the command frame, DPS10 asks by the response frame. Table 1 is the command frame format. The Starting Byte is 0x7E. CMD represents the types of command frame, including initialization command frame, data exchanging command frame, etc. CMDinfo represents the length of DATA. SpeByte is segmentation message. DATA could not be longer than 244Byte in a frame. If the DATA is longer than 244 Byte, we need transmit it segmented.

Table 1. Command frame format

Starting Byte	CMD	CMDinfo	SpeByte	DATA	CRC
1Byte	1Byte	1Byte	1Byte	n bytes	1Byte

In order to communicate with Profibus-DP master station, STM32 should send the initialization command frame to DPS10 through UART0 at first. The DATA of the frame includes I/O Configuration Length, I/O Configurations, Input and Output Data Length. The I/O Configuration Length represents the number of the I/O module. The I/O Configurations represent the data types of each I/O module. In the system, every I/O module correspond a CAN mode, so the I/O Configuration Length is equal to the number of CAN bus measurement nodes, and all of the I/O Configuration are 4 Byte input. These data must be the same as the Generic Station Description (GSD), including the ranking and value. GSD file is used to describe Profibus-DP devices. It can make different Profibus-DP devices be identified by configuration software tools and connect to Profibus-DP.

After initialize DPS10, It begins to exchange data with Profibus-DP master station through data exchanging command frame. The DATA of the frame consist of all the I/O modules’ Input data in due order. Every I/O module corresponds to 4 Byte in the DATA of the frame. In other words, Every CAN bus mode corresponds to 4 Byte in the DATA. Table 2 is the correspondence relation of the data.

Table 2. The correspondence relation of data

DATA	I/O modules	CAN Nodes	Source address
0~3 Byte	I/O module 1	CAN Node 1	0000001
4~7 Byte	I/O module 2	CAN Node 2	0000010
8~11 Byte	I/O module 3	CAN Node 3	0000011
.....

When the CAN receiving data buffer received a data frame, STM32 will pick up the source address (ID10-ID4) and the fault information (the Data Field) of the frame, and then transmit them to Profibus-DP data buffer. The main program will place the fault information into the corresponding place of the data exchanging command frame according to the source address, and transmit the frame to DPS10 through USART0. After DPS10 received the command frame, it will ask by the response frame, indicating that the communication is successful and STM32 could transmit the next frame. Then DPS10 begins to communicate with the master station by the integrated state machine in order to transmit the fault information to Profibus-DP master station. Figure 6 is the Flow chart of main program.

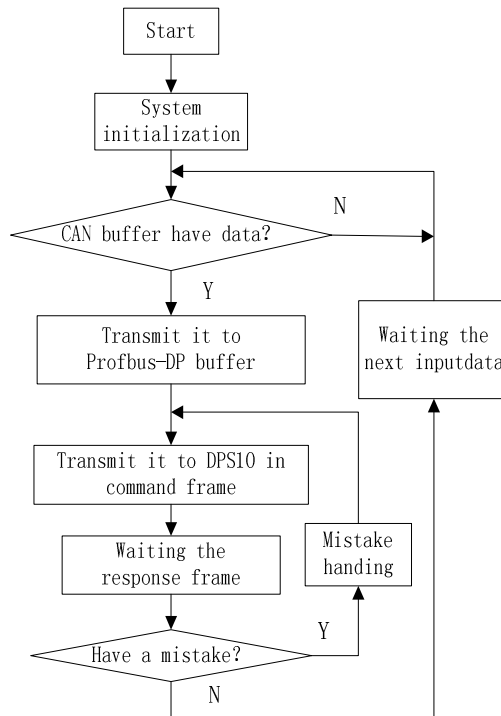


Fig. 6. Flow chart of main program

5 Conclusion

The practical result shows that the gateway between Profibus-DP and CAN based on STM32 works stably and reliably. It has solved the problem that the two different protocols are incompatible in the fault location system. It not only extends the communication distance and reduces the wiring of the system, but also makes it have the feature of high security, reliability and real-time. This fault location system is suitable for the application in long distance belt conveyor.

References

1. Giraud, L., Masse, S.: Belt Conveyor Safety. *Professional Safety* (11), 20–27 (2004)
2. Zong, M., Sun, J.Y.: Fault location system of adhesive tape transport machine based on CAN bus. *Journal of Shenyang University of Technology* 30(6), 667–671 (2008)
3. Cenesiz, N.: Controller Area Network (CAN) for Computer Integrated Manufacturing Systems. *Journal of Intelligent Manufacturing* (8), 30–49 (2004)
4. Ren, S.J., Niu, Q.F.: Design of Gateway between Profibus and CAN Based on Freescale S12. *Automation & Instrumentation* 24(7), 26–30 (2009)
5. Shen, Y., Li, X.Q.: Design of electric vehicle control system based on CAN bus. *Electronic Design Engineering* 18(11), 143–145 (2010)
6. Pan, T., Zhang, N.: Design and Application of CAN Bus Repeater Based on STM32. *Microcontrollers & Embedded Systems* 11(1), 46–48 (2011)

A Contract Performing Scenario Model for Web Services Collaboration

Denghui Zhang¹ and Ji Gao²

¹ College of Information and Technology, Zhejiang Shuren University,
310015 Hangzhou, China
zzddhh@263.net

² College of Computer Science, Zhejiang University, 310027 Hangzhou, China
gaoji1@zju.edu.cn

Abstract. Self-organized service chain formed by service agent can't complete the scheduled collaborative target because of exceptions of service provider. This paper presents a service chain evolution model based on collaborative state. In the model, service collaborative relationship is modeled as a contract set, whose performance results are regarded as the result of service coordination state. And then the monitoring and evolution policies of collaborative process are established based on collaborative state. On this basis, a collaborative scenario-driven adaptive and self-evolutionary mechanism for service chain is established. Based on planning policies, different evolution plans are formulated to obtain a better flexibility of service chain.

Keywords: Virtual organization, service exception, collaborative scenario, self-evolution.

1 Introduction

Currently, using service-oriented architecture(SOA)[1] to build collaborative service based virtual organization(VO) is becoming an effective method to solve resource sharing and collaborative problems[2-3]. But, service collaboration can't effectively, flexibly cope with dynamic changes of the network environment and application requirements because of the non-autonomous of software services. For this purpose, researchers have studied methods of service workflow adaptive evolution, dynamic formation of virtual organizations, dynamic collaboration of multi-Agent and etc. for the service collaborative change [4-9].

However, these collaborative evolution models are difficult to apply to the open, dynamic, heterogeneous network environment. On the one hand, the operational environment of VO is lack of clear semantic description. On the other hand, the evolution of VO members is lack of guide policy. Fortunately, policy-based self-management[10] and norm-based monitoring[11] have separately been well studied in different areas. Combining with contract and policy, this paper proposes a contract performing scenario-driven VO self-evolution model, which considers the VO operational process as VO members performing collaborative activities in accordance with contractual agreements.

2 Overview of the Model

Contract performing circumstance driven adaptive and flexible evolution model defined as the 5-tuple is shown in Fig. 1.

$$CCA E=(AVMP, VC, CPCM, JCCM, VAEM)$$

AVMP: Adaptive and eVolutionary Manage Policies.

VC: VO Contract set, which includes service providing/demanding contract and role assumption contract.

CPCM: Contract Performing Circumstance Mechanism, which reflects the implementation status of service performing contract, is used to compose service collaborative scene of VO runtime.

JCCM: Joint Contract Compliance Mechanism. It respects contract performing agreement by signing service contracts, and monitors the implementation process of collaborative services in VO through the self-examination and mutual checking of contract performing status.

VAEM: VO Adaptive and self-Evolutionary Mechanism, which includes four phases: scenario monitoring, scenario variation analysis, collaborative evolution planning, evolution plan executing, is used to support adaptive and flexible evolution for services collaboration.

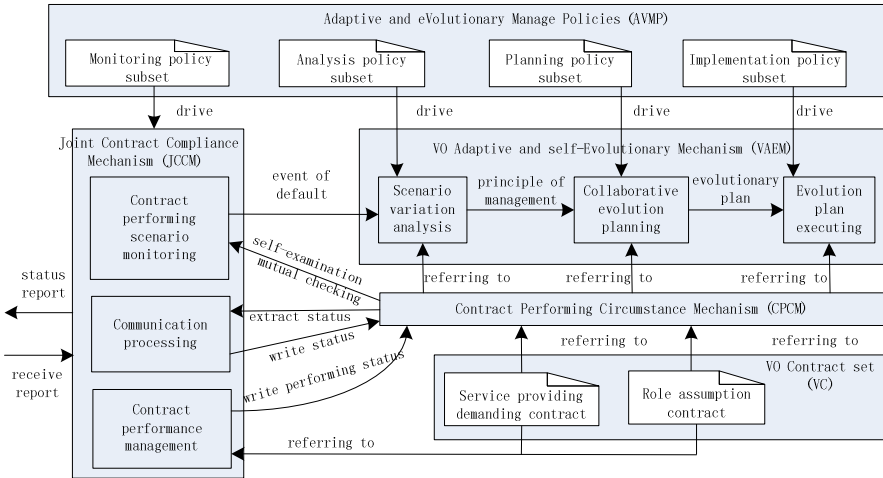


Fig. 1. VO adaptive and flexible self-evolution model

3 Formal Specifications

3.1 Adaptive and eVolutionary Manage Policies (AVMP)

Policy is a set of rules, in which every rule is composed of an evaluation condition ρ and an action (or an action sequence) α . The action α is executed by agent while agent

believes that the evaluation condition is true. Formally, a policy can be defined in the form as follows.

Definition 1. Policy

$policy = \{(\rho_i, \alpha_i) \mid i \in \{1, \dots, m\}\}$, In this definition, ρ_i represents an evaluation condition of this policy, and α_i represents the action sequence executed after that ρ_i is satisfied, α_i can be defined as $\{a_{i1}, \dots, a_{in}\}$.

The execution process of a policy is an action sequence, formally defined as follows.

Definition 2. Action Sequence of Policy Execution

$$PolicyExecute(policy) = \left(\bigwedge_{i=1}^m (\rho_i \Rightarrow Does(a_{in})) \right) ? , a_{in} \in \alpha_i$$

Wherein, ‘;’ and ‘?’ is the primitive symbols of dynamic description logic, ‘;’ denotes the action sequence, ‘?’ denotes action test.

Therefore, $PolicyExecute(policy)$ denotes the action sequence of policy execution, that is, the execution process of policy is represented by the execution of a action sequence. For every rule in the policy, agent judges whether the evaluation condition of the rule is satisfied, and execute the action of the rule if satisfied.

3.2 VO Contract set (VC)

Contract is the conditions and basis of collaboration between VO members. It is divided into two categories: role assumption contract and service providing/demanding contract. Role assumption contract is the credentials of the service to join VO, and service providing/demanding contract represents rights and obligations between cooperation partners.

A. Role assumption contract: This contract is defined as the 4-tuple: $BPR = (PRR, PIR, Right, Obligation)$. PRR denotes the response role set of role assumption. PIR is the starting role set of role assumption. $Right$ expressed by a group of norm sets denotes the right of the role. Here, each norm provides the trigger concept, the effective time period and the specific disposal operations. $Obligation$ denotes the obligations of the role, which is also expressed by a group of norm sets.

B. service providing/demanding contract: Contract of service(SC) is defined as a 3-tuple: $SC = (CBM, CSI, PN)$, where CBM denotes contract’s basic information such as contract number, credit card number, service provider code and so on. CSI denotes the detailed service items provided by contract. PN is the half-ordered set of performing norms of SC.

3.3 Contract Performing Circumstance Mechanism (CPCM)

Definition 3. Performing Norm of Service Contract

$PN = OB_a^{SC}(\rho \leq \delta \mid \sigma) \mid FB_a^{SC}(\rho \leq \delta \mid \sigma) \mid PB_a^{SC}(\rho \leq \delta \mid \sigma)$, indicating respectively that, when σ holds true, the role a (contractors of SC) is obligated to, forbidden to, or authorized to make ρ true before deadline δ (here ρ , δ , and σ are all the propositions describing service cooperation status).

Definition 4. *Executing Status of Contract Performing Norm*

$ES = (PN\text{-number}, Status\text{-type}, Status\text{-description})$, where PN-number denotes the number of current performing PN that belongs to SC. Status-type denotes the type of performing states, such as success, failure and exception etc. Status-description gives the description of performing scenario.

When performing norm (PN) of obligation type (OB_a^{SC}) or authorization type (PB_a^{SC}) executes successfully, the Status-description is the example of proposition ρ or ρ , which is specified by PN that should be transformed into true, otherwise, failure or exception description including description type and content will be given. Performing norm (PN) of forbidden type (FB_a^{SC}) would not perform or record implementation status under normal circumstances, until encountering the breach of contract. Here, the type of implementation status is indicated by “abnormal” and exception description is given.

Based on the above definitions of SC and PN, Contract performing circumstance (CPCsc) of each SC is modeled as the sequence (execution orders are specified by contact performing agreement) of executing status (ESi) of contract performing norm.

$CPCsc = \{ES1, ES2, \dots, ESm\}$, $ESi = \text{Executing-state } (PNj)$, $PNj (\in PN\text{-sets}c)$, where, $PN\text{-sets}c$ denotes performing norm set constituted for service contract. PNj represents one of the performing norms. Executing-state (PNj) denotes implementation status of performing norm.

3.4 Joint Contract Compliance Mechanism (JCCM)

Joint Contract Compliance Mechanism (JCCM), which is used to implement joint compliance for single contract, is expressed as the following 7-tuple:

$CCM = \{VM\text{-set}, Contract\text{-set}, PN\text{-set}, Self\text{-executing}, Self\text{-examining}, Inter\text{-reporting}, Inter\text{-examining}\}$

VM-set: the set of members in the VO;

Contract-set: the set of service contracts in VO;

PN-set: Joint performing norm set of all service contracts, $pns(sc)$, $\cup pns(sc)$, ..., $\cup pns(sc)$, where, $pns(sc)$ denotes the performing norm set of $sc (\in Contract\text{-set})$.

Self-executing: $VM\text{-set} \times Contract\text{-set} \rightarrow Ppn\text{-set}$. Based on the signed $sc (\in Contract\text{-set})$, each $vm (\in VM\text{-set})$ executes service contract performing norm set which belongs to its obligations and authorities (can be empty set).

Self-examining: $VM\text{-set} \times Contract\text{-set} \rightarrow Ppn\text{-set}$. Based on the signed $sc (\in Contract\text{-set})$, each $vm (\in VM\text{-set})$ examines the performance status of its performing norm, and $self\text{-examining}(vm, sc) = self\text{-executing}(vm, sc)$.

Inter-reporting: $VM\text{-set} \times Contract\text{-set} \rightarrow Ppn\text{-set}$. Based on the signed $sc (\in Contract\text{-set})$, each $vm (\in VM\text{-set})$ reports the execution status of performing norm to collaboration partners, and $Inter\text{-reporting}(vm, sc) = Self\text{-executing}(vm, sc)$.

Inter-examining: $VM\text{-set} \times Contract\text{-set} \rightarrow Ppn\text{-set}$. Based on the signed $sc (\in Contract\text{-set})$, each $vm (\in VM\text{-set})$ examines the performing norm of service provider/demander, and $Inter\text{-examining}(vm, sc) \cup Self\text{-examining}(vm, sc) = pns(sc)$, $Inter\text{-examining}(vm, sc) \cap Self\text{-examining}(vm, sc) = \emptyset$.

3.5 VO Adaptive and self-evolutionary Mechanism (VAEM)

Although the service contract has set up remedial performing norm to deal with foreseeable exceptions of contract performance, it will still encounter some unforeseen exceptions which lead to abnormal termination of the contract. The maintenance activities for these exceptions will be undertaken by the self-evolutionary mechanism.

VO Adaptive and self-Evolutionary Mechanism (VAEM) is defined as the following 10-tuple:

$VAEM=(CPC,AEP,CV-events,CVT-principles,CE-plans,CE-actions,Monitoring,Analyzing,Planning,Executing)$.

CPC : Service contract performing circumstance set $\{CPC^{sc_1}, CPC^{sc_2}, \dots, CPC^{sc_n}\}$, CPC^{sc_i} denotes performing circumstance sc_i of service contract i .

AEP : Adaptive and evolutionary policy set, $AEP = mo-policies \cup an-policies \cup pl-policies \cup ex-policies$, where *mo-policies*, *an-policies*, *pl-policies*, *ex-policies* respectively denote monitoring, analysis, planning, implementation policy subset.

4 Adaptive and Flexible Evolvment Process

The flexible evolvment process is divided into four phrases: collaborative scenario monitoring, scenario variation analysis, collaborative evolution planning and evolution plan executing.

4.1 Collaborative Scenario Monitoring

$Monitoring(x, CPC^{SC}) \Rightarrow \exists mop(\in mo-policies) \wedge Exception(x, mop, CPC^{SC}) \wedge Create(x, CVE)$. Monitoring activities are composed by self-examination and mutual examination of executing status of service contract performing norm. Common or specific scenario monitoring policies $mop (\in mo-policies)$ are used to discovery exceptions of $CPC (\in CPC)$, and then establish corresponding event of default $CVE(\in CV-events)$.

4.2 Scenario Variation Analysis

$Analyzing(x, CPC^{SC}, CVE) \Rightarrow \exists anp(\in an-policies) \wedge Compl(x, anp, CPC^{SC}, CVE) \wedge Create(x, cvtp)$. Specific scenario variation analysis policy $anp(\in an-policies)$, which is activated by $CVE(\in CV-events)$, is used to analyze $CPC (\in CPC)$, and provide analysis results $cvtp(\in cvt-principles)$ (disposal principles of event of default).

4.3 Collaborative Evolution Planning

$Planning(x, cvtp) \Rightarrow plp(\in pl-policies) \wedge enabled(cvtp, plp) \wedge Create(x, cep)$. Specific collaborative evolution planning policy $plp (\in pl-policies)$, which is activated by disposal principles of event of default $cvtp (\in cvt-principles)$, is used to program and generate evolution plan $cep (\in ce-plannes)$.

4.4 Evolution Plan Executing

$Executing(x, cep) \Rightarrow exp (\in ex-policies) \wedge enabled(cep, exp) \wedge starting(x, ceas)$. Specific evolution plan executing policy $exp (\in ex-policies)$, which is activated by evolution plan $cep (\in ce-plannes)$, is used to start evolution action $ceas (\subseteq ce-actions)$ specified by exp .

5 Case Study

This section applies a case of geospatial information services flow to verify effectivity of the proposed VO self-evolution model. The case is used to aid decision-making for flood emergency decision, in which, the VO established by flood emergency decision service provider FDS needs to integrate multiple geospatial information services (including WCS, WFS, WPS). The collaborative process of VO is shown in Fig. 2.

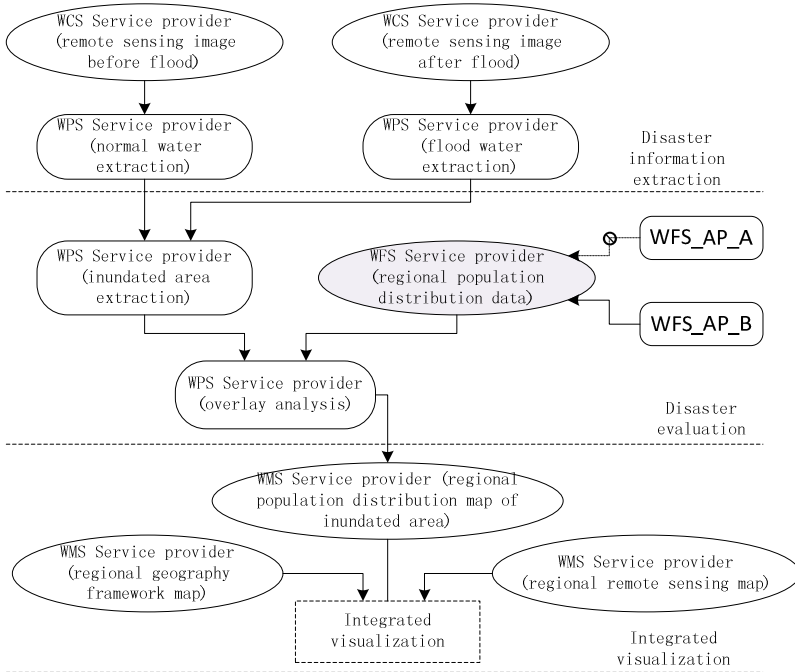


Fig. 2. Collaborative process of VO for flood emergency decision

Suppose $APC:(Contract_info, Data_items, pn)$ has been established through consultation between FDS and regional population distribution data service provider WFS_AP. It means that data service provider signs commitment to provide datum service specified by Data_items to FDS in accordance with pn (contract performing norm set).

$pn = \{(Norm001, \text{contract inuring, FDS, pay } 50\$, \text{ in } 3 \text{ days}),$
 $(Norm002, \text{contract Performance, Norm001, Success, WFS_AP, open download}$
 $\text{Service, in } 8 \text{ hours}),$
 \dots
 $(Norm070, \text{contract Performance, Norm002, Fail or Abnormity, FDS,}$
 $\text{CancelContract, Now}),$
 $(Norm071, \text{contract Performance, Norm002, Fail or Abnormity, FDS,}$
 $\text{PostponDepositProvide, Now})\}$

These norms represent the following meanings:

Norm001, FDS is duty-bound to pay \$50 deposit to *WFS_AP* within 3 days after execution of *APC contract*.

Norm002, FDS is duty-bound to open the network service within 8 hours after successful execution of *Norm001*.

Norm070, FDS has the right to immediately cancel *APC contract* when an execution exception is thrown by *Norm002*.

Norm071, FDS has the right to delay margin payment when an execution exception is thrown by *Norm002*.

Self-evolution reasoning polices for *FDS* are as follows:

- 1) *Scene Monitor Policy(SMP):{(first, time), (second, reputation)}*
- 2) *Scene Analysis Policy(SAP):{(time limited, penalty), (time limited and reputation of WFS_AP <7, cancellation), (time unlimited and reputation of WFS_AP <6, contract cancellation), (time unlimited and reputation of WFS_AP >6, renegotiation)}*
- 3) *Cooperation Evolvement Policy(CEP): {(penalty, penalty starting), (renegotiation, renegotiation starting), (contract cancellation and having person, changing person), (contract cancellation and having path, changing path), (contract cancellation and having no path, process thing left)}*
- 4) *Evolvement Plan Deployment Policy(EPDP): {(penalty starting, notify counter partner to pay penalty), (renegotiation starting, pause every related sub operation and start negotiation module), (changing person, pause every related operation and start changing person operation), (changing path, pause every related operation and start changing path operation), (process things left, pause every related operation and notify up operation)}*.

We assume that *FDS* pays 50\$ in 3 days after the execution of contract, but *WFS_AP* does not provide data service in 8 hours, thus a *WFS_AP* happens, which *FDS* must handle.

Firstly, *FDS* captures an exception in performing *Norm002* with *SMP(first, time)*, and finds that *WFS_AP* does not provide data service in 8 hours, so *FDS* concludes that it must cancel the *APC* by *SAP(time limited and reputation of WFS_AP <7,cancellation)*. According to path generation policy, *FDS* find that there exists another *WFS_AP* provider, thus produces the evolvement plan: changing provider. Finally, according to plan deployment policy, *FDS* replaces the *WFS_AP_A* with *WFS_AP_B* and continues the cooperation. Thus the whole process of VO evolvement is over.

6 Conclusions

This paper presents a model of flexible virtual organization evolving facilitated by contract performing scene facilitated by policies and contracts, to handle exception event happened in VO. The model introduces the concept of policy, contract, and norm. Policy is the dynamic constraint to agent behavior, which can regulate the deducing behavior of autonomous agent. Contract is the constraint to agents involve in VO. The introduction of policy separates the management logic from application logic, which advances the applicability of the system. This paper defines policy attempt operator as an extension of attempt operator, and describes the flexible evolvement process after the happening of exception.

However, there is a great deal of further work required to make the policies and contract facilitated exception handling model more comprehensive, including the autonomic configure of policy, the importing of norm to regulate the behavior of individual agent, the improvement of contract description, etc.

Acknowledgements. This work is supported by the Priority Theme Emphases Project of Zhejiang Province, China (Grant 2010C11045).

References

1. Stal, M.: Using architectural patterns and blueprints for service-oriented architecture. *IEEE Software* 23(2), 54–61 (2006)
2. Erl, T.: *Service-Oriented Architecture: Concept, technology, and design*. Professional Technical Reference. Prentice Hall, Upper Saddle River (2005)
3. Erl, T.: *Service-Oriented Architecture: A Field Guide to Integrating XML and Web Services*. Pearson Education, London (2004)
4. Afsarmanesh, H., Camarinha, M.: A Framework for Management of Virtual Organization Breeding Environments. In: *Proceedings PRO-VE 2005*, pp. 35–48 (September 2005)
5. Yu, L., Herman, L., Stanley, Y.W.S.: Adaptive Grid Service Flow Management: Framework and Model. In: *ICWS 2004* (2004)
6. Yathiraj, B.U., Munindar, P.S.: Contract Enactment in Virtual Organizations: A Commitment-Based Approach. In: *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI)*, pp. 722–727. AAAI Press, Boston (2006)
7. Abdel-illah, M., Laurent, J.: Dynamic Coalition of Resource-Bounded Autonomous Agents. In: *ICTAI*, vol. (1), pp. 117–123 (2007)
8. Mati, G., Avigdor, G.: Optimizing Exception Handling in Workflows Using Process Restructuring. *Business Process Management*, 407–413 (2006)
9. Liao, B., Gao, J., Hu, J., Chen, J.: Ontology-Based Conceptual Modeling of Policy-Driven Control Framework: Oriented to Multi-agent System for Web Services Management. In: Chi, C.-H., Lam, K.-Y. (eds.) *AWCC 2004*. LNCS, vol. 3309, pp. 346–356. Springer, Heidelberg (2004)
10. Barrett, R.: People and Policies: Transforming the Human-Computer Partnership. In: *Proceedings of the Fifth IEEE International Workshop on Policies for Distributed Systems and Networks, POLICY 2004* (2004)
11. Grossi, D., Aldewereld, H.: Ubi Lex Ibi Poena: Designing norm enforcement in E-institutions. In: *Proceedings of the AAMAS 2006*, Hakodate, Japan (May 2006)

A SVM Approach for MCs Detection by Embedding GTDA Subspace Learning

Xinsheng Zhang, Minghu Wang, and Fan Yu

School of Management, Xi'an University of Architecture and Technology
Xi'an 710055, Shaan Xi China
xinsheng.zh@gmail.com, minghuwang@126.com,
yufan@xauat.edu.cn

Abstract. This paper presents a SVM based approach to microcalcification clusters (MCs) detection in mammograms by embedding general tensor discriminant Analysis (GTDA) subspace learning. In the approach GTDA and other subspace learning methods are employed to extract subspace features. In extracted feature domain, the MCs detection procedure is formulated as a supervised learning and classification problem, and SVM is used as a classifier to make decision for the presence of MCs or not. A large number of experiments are carried out to evaluate and compare the performance of the proposed MCs detection algorithms. The experiment result suggests that the proposed method is a promising technique for MCs detection.

Keywords: Microcalcification clusters, support vector machines, subspace learning, general tensor discriminant analysis.

1 Introduction

The presence of microcalcifications (MCs) can be one of the early signs of breast cancer. They are small calcium deposits that appear as bright spots in a mammogram, and they may appear as single spots or they can group to form clusters. Individual MCs are sometimes difficult to detect due to their variation in shape, orientation, brightness and size (typically, 0.05mm-1mm), and because of the surrounding breast tissue. Moreover, both the spatial distribution of the MCs into the cluster and the shape of the MCs are strictly correlated to the likelihood of the presence of a malignant tumor.

Because of its importance in early breast cancer diagnosis, accurate detection of MCs has become a key problem. Recently, a number of different approaches, which could assist radiologists in diagnosis of breast cancer, have been applied to the detection of MCs. A thorough review of various methods for MCs detection was made in [1]. Considered MCs detection as a classification problem, various classification methodologies have been proposed for the characterization of MCs, such as, fuzzy rule-based systems [2], and support vector machines [3]. In the past decade, the most work reported in the literature employed neural networks for MCs detection [6-8]. With the rise and development of SVM, various SVMs are employed to classify the ROIs.

To improve the performance of the available MCs detection algorithms, this paper employs SVM as a classifier in the GTDA subspace domain to distinguish the MCs from the other ROIs. The classifier is first trained by the real mammograms from DDSM. Then it is used to detect other ROIs. In the experiments, the proposed approach achieved the best average sensitivity as high as 94.31% with respect to 9.31% false positive rate.

2 Discriminant Subspace Learning Background

Subspace learning algorithms [4] such as principal component analysis (PCA) and linear discriminant analysis (LDA) traditionally transform the input data as vectors and often in a high dimensional feature space to a low-dimensional subspace. Because of the under sample problems (USP): the dimensionality of the feature space is much higher than the number of training samples (the training number is much smaller than the input feature dimension in our task, $m \ll 115 \times 115$), PCA and LDA do not work well. Recently, multilinear algebra, the algebra of higher-order tensors, was applied for analysing the multifactor structure of image ensembles, such as TSA and GTDA [5]. Tensor space analysis (TSA) considers an image as the second order tensor in $R^{n_1} \otimes R^{n_2}$, where R^{n_1} and R^{n_2} is two vector spaces. And the relationship between the row vectors can be naturally characterized by TSA and GTDA.

2.1 PCA and LDA

PCA is a typical linear dimensional reduction algorithm. The basic idea of PCA is to project the data along the directions of maximal variances so that the reconstruction error can be minimized. Given a set of data points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, let \mathbf{w} be the transformation vector and $y_i = \mathbf{w}^T \mathbf{x}_i$. The objective function of PCA is

$$\mathbf{w}_{opt} = \arg \left\{ \max_{\mathbf{w}} \sum_{i=1}^n (y_i - \bar{y})^2 \right\} = \arg \left\{ \max_{\mathbf{w}} \mathbf{w}^T \mathbf{C} \mathbf{w} \right\} \quad (1)$$

where $\bar{y} = 1/n \sum y_i$ and \mathbf{C} is the data covariance matrix. Compared with PCA seeking directions that are efficient for representation, linear discriminant analysis (LDA) seeks directions that are efficient for discrimination. Suppose we have a set of n samples $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, belonging to k classes. The objective function of LDA is as follows,

$$\mathbf{w}_{opt} = \arg \left\{ \max_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \right\} = \arg \left\{ \max_{\mathbf{w}} \frac{tr(\mathbf{w}^T \mathbf{S}_B \mathbf{w})}{tr(\mathbf{w}^T \mathbf{S}_W \mathbf{w})} \right\} \quad (2)$$

$$\mathbf{S}_B = \frac{1}{n} \sum_{i=1}^k n_i (\mathbf{m}^{(i)} - \mathbf{m})(\mathbf{m}^{(i)} - \mathbf{m})^T \quad (3)$$

$$\mathbf{S}_W = \frac{1}{n} \sum_{i=1}^k \left(\sum_{j=1}^{n_i} (\mathbf{x}_j^{(i)} - \mathbf{m}^{(i)})(\mathbf{x}_j^{(i)} - \mathbf{m}^{(i)})^T \right) \quad (4)$$

where \mathbf{m} is the total sample mean vector, n_i is the number of samples in the i th class, $\mathbf{m}^{(i)}$ is the average vector of the i th class. We call \mathbf{S}_W the within-class scatter matrix and \mathbf{S}_B the between-class scatter matrix.

2.2 General Tensor Discriminant Analysis

GTDA is a tensor extension of the differential scatter discriminant criterion (DSDC). The differential scatter discriminant criterion(DSDC) is defined by

$$\mathbf{U}^* = \arg \left\{ \max_{\mathbf{U}^T \mathbf{U} = \mathbf{I}} \left(\text{tr}(\mathbf{U}^T \mathbf{S}_B \mathbf{U}) - \xi \cdot \text{tr}(\mathbf{U}^T \mathbf{S}_W \mathbf{U}) \right) \right\}, \quad (5)$$

where ξ is a tuning parameter; $\mathbf{U} \in R^{N \times N}$ ($N \ll N$), constrained by $\mathbf{U}^T \mathbf{U} = \mathbf{I}$, is the projection matrix; \mathbf{S}_B and \mathbf{S}_W are defined in (3) and (4). The tensor DSDC can be get as

$$\begin{aligned} \mathbf{U}^* &= \arg \left\{ \max_{\mathbf{U}^T \mathbf{U} = \mathbf{I}} \left(\text{tr}(\mathbf{U}^T \mathbf{S}_B \mathbf{U}) - \xi \cdot \text{tr}(\mathbf{U}^T \mathbf{S}_W \mathbf{U}) \right) \right\} \\ &= \arg \left\{ \max_{\mathbf{U}^T \mathbf{U} = \mathbf{I}} \left(\begin{aligned} &\sum_{i=1}^k n_i \left\| \left((\mathbf{m}^{(i)} - \mathbf{m}) \times_1 \mathbf{U}^T \right) \otimes \left((\mathbf{m}^{(i)} - \mathbf{m}) \times_1 \mathbf{U}^T \right) \right\| \\ &- \xi \sum_{i=1}^k \sum_{j=1}^{n_i} \left\| \left((\mathbf{x}_j^{(i)} - \mathbf{m}^{(i)}) \times_1 \mathbf{U}^T \right) \otimes \left((\mathbf{x}_j^{(i)} - \mathbf{m}^{(i)}) \times_1 \mathbf{U}^T \right) \right\| \end{aligned} \right) \right\}, \quad (6) \end{aligned}$$

where $\|\cdot\|$ is the Frobenius norm and the projection matrix \mathbf{U} is constrained by $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Let $\mathcal{X}_j^{(i)}$ denote the j th training sample (with tensor representation) in the i th individual class k , $\mathcal{M}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathcal{X}_j^{(i)}$ is the class mean tensor of the i th class, $\mathcal{M} = \frac{1}{n} \sum_{i=1}^k \mathcal{M}_i$ is the total mean tensor of all training tensors, and \mathbf{U}_l denotes the l th projection matrix obtained during the training procedure. Moreover, $\mathcal{X}_j^{(i)} \Big|_{1 \leq j \leq n_i}$, $\mathcal{M}_i \Big|_{i=1}^k$, and \mathcal{M} are all N th order tensors that lie in $R^{N_1 \times N_2 \times \dots \times N_N}$. Based on the analogy in (7), GTDA can be defined by replacing $\mathbf{x}_j^{(i)}$, $\mathbf{m}^{(i)}$ and \mathbf{m} with $\mathcal{X}_j^{(i)}$, \mathcal{M}_i and \mathcal{M} , respectively, as

$$\arg \max_{\mathbf{U}^T \mathbf{U} = \mathbf{I}} \left(\begin{aligned} &\sum_{i=1}^k n_i \left\| \left((\mathcal{M}^{(i)} - \mathcal{M}) \prod_{l=1}^M \times_l \mathbf{U}_l^T \right) \otimes \left((\mathcal{M}^{(i)} - \mathcal{M}) \prod_{l=1}^M \times_l \mathbf{U}_l^T \right) \right\| \\ &- \xi \sum_{i=1}^k \sum_{j=1}^{n_i} \left\| \left((\mathcal{X}_j^{(i)} - \mathcal{M}^{(i)}) \prod_{l=1}^M \times_l \mathbf{U}_l^T \right) \otimes \left((\mathcal{X}_j^{(i)} - \mathcal{M}^{(i)}) \prod_{l=1}^M \times_l \mathbf{U}_l^T \right) \right\| \end{aligned} \right). \quad (7)$$

The optimal problem does not have a closed form solution, but Tao[5] used the alternating projection method, which is an iterative algorithm, to obtain a numerical solution. In GTDA, the original general tensor \mathcal{X} can be transformed by using the projected tensor $\mathcal{Y} = \mathcal{X} \prod_{l=1}^M \times_l \mathbf{U}_l$ to a tensor subspace.

3 SVM Classifier for MCs Detection

For a given digital mammography image, we consider the MCs detection process as the following steps: 1) Preprocess the mammography image by removing the artifacts, suppressing the inhomogeneity of the background and enhancing the microcalcifications; 2) At each pixel location in the image, extract an $A_{m \times m}$ small window to describe its surrounding image feature; 3) Apply the subspace learning algorithm to get the subspace feature vector \mathbf{x} ; 4) Use the SVM classifier to make decision whether \mathbf{x} belongs to MCs class or not.

3.1 Input Patterns for MCs Detection

After the subspace learning stage, we transform $A_{m \times m}$ into a feature vector \mathbf{x} . $A_{m \times m}$ is a small window of $m \times m$ pixels centered at a location that we concerned in a mammogram image. The choice of m should be large enough to include the MCs (in our experiment, we take $m=115$). The task of the SVM classifier is to decide whether the input window $A_{m \times m}$ at each location is a MCs pattern ($y = +1$) or not ($y = -1$).

3.2 SVM Training and Optimization Parameter Selection

After the positive and negative training examples are gathered, some parameters should be initialized first, such as the type of kernel function, its associated parameter, and the regularization parameter c_1 and c_2 in the structural risk function. To optimize these parameters, we applied m -fold cross validation method to get the optimal parameters.

Once the best parametric setting (i.e., the type of the kernel function and its associated parameter) is determined, the SVM classifier is retrained using all the available samples in the training set to obtain the final form of the decision function. The resulting classifier will be used to detect the MCs in mammogram.

4 Experimental Results

Up till now, we have shown our approach to MCs detection. In this section we evaluate the performance of our method by using the real mammogram data from DDSM. The data in the training, test, and validation sets were randomly selected from the preprocessed dataset. Each selected sample was covered by a 115×115 window whose center coincided with the center of mass of the suspected MCs. The blocks included 2231 with true MCs and 8364 with normal tissue. 75% of the blocks were assigned to the training set, 25% to the test set.

4.1 Subspace Feature Extraction

Before training the classifier, we first use the subspace learning algorithm to extract MCs feature in subspace domain. For PCA and TSA we should find the projection matrix \mathbf{w}_{opt} , \mathbf{U}^T and \mathbf{V} . A random subset X_i is taken from the training data set, and the rest of database are used to learn the subspace to get \mathbf{w}_{opt} , \mathbf{U}^T and \mathbf{V} . The testing samples X_i are then projected into the low dimensional representation subspace. Note that, for LDA and GTDA, X_i is taken with labels Y_i , and the rest with labels are also used to learn the subspace to get \mathbf{w}_{opt} and $\prod_{l=1}^M X_i U$. Similar to PCA and TSA, we can get the low dimensional subspace feature by the projection formula discussed in section 2.

4.2 Performance Evaluation Results

We perform MCs detection to compare PCA and LDA with TSA and GTDA based subspace learning method with SVM on the real dataset. To simplify our task, we randomly select 1000 MCs and 1000 normal samples as dataset pool for the evaluation task. 75% of all samples(1500) in the dataset pool were selected for training task, the others (500) for test. To evaluate the stability of each method, we repeat the sampling 20 times so that we can compute the mean and standard deviation of the detection accuracy, Sensitivity and Specificity. In the each sampling of 20 times, we fixed the test samples(500), and we reduced the fraction of all the training samples from 95% to 5%. That is we will perform the detection task 20 rounds, and in each round we randomly select training samples from 95% of all training samples (1500) to 5% to train classifiers. The trained classifiers are evaluated using the 500 test samples.

Average experimental results of SVM classifier based on different subspace learning algorithms are shown in Table 1. It can be shown that, the SVM classifier with GTDA preprocessing has a higher detection accuracy rate compared to the other subspace learning methods. It also can be seen that when tensor representation method is used in subspace learning before MCs detection, more discriminant information, which will feed to the next stage classifier, will be preserved in the tensor subspace.

Table 1. Experimental Results of SVM approach Based on Different Subspace Learning Algorithm

Method	Average Sensitivity	Average Specificity	Average Error-rate	Average Az
PCA+SVM	0.9115	0.8962	0.0962	0.9496
LDA+SVM	0.9077	0.8885	0.1019	0.9380
TSA+SVM	0.9077	0.9154	0.0885	0.9419
GTDA+SVM	0.9431	0.9069	0.0714	0.9482

5 Conclusions

In this paper, a SVM based approach is proposed for detection of MCs in mammograms. In this method, GTDA and other subspace learning methods, are employed to extract subspace features and SVM is trained through supervised learning with the extracted features to test at every location in a mammogram whether an MCs is present or not. Experimental results show that the proposed approach is effective and efficient.

Acknowledgments. This work is partially supported by the Scientific Research Plan of Shaanxi Education Department (No.09JK563 and No.09JK542) and the Startup Fund for Scholars in Xi'an University of Architecture and Technology.

References

1. Cheng, H.D., Cai, X., Chen, X., Hu, L., Lou, X.: Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern Recognition* 36(12), 2967–2991 (2003)
2. Riyahi-Alam, N., Ahmadian, A., Tehrani, J.N., Guiti, M., Oghabian, M.A., Deldari, A.: Segmentation of suspicious clustered microcalcifications on digital mammograms: using fuzzy logic and wavelet coefficients. In: *Proc. IEEE Int'l Conf. Signal Processing (ICSP 2004)* (2004)
3. Sukhwinder, S., Vinod, K., Verma, H.K., Dilbag, S.: SVM Based System for classification of Microcalcifications in Digital Mammograms. In: *Proc. IEEE Int'l Conf. Engineering in Medicine and Biology Society (EMBS 2006)* (2006)
4. Brand, M.: Continuous nonlinear dimensionality reduction by kernel eigenmaps. In: *International Joint Conference on Artificial Intelligence (IJCAI), Acapulco Mexico* (2003)
5. Tao, D., Li, X., Wu, X., Maybank, S.J.: General Tensor Discriminant Analysis and Gabor Features for Gait Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(10), 1700–1715 (2007)

Differential Evolution Using Local Search for Multi-objective Optimization

Youyun Ao

School of Computer & Information, Anqing Teachers College,
246011 Anqing, Anhui, China
youyun.ao@gmail.com

Abstract. Differential evolution has the characteristics of fast convergence, less parameters, and ease of implementation. This paper proposes an enhanced DE using the local search for multi-objective optimization, which is called DEMOLS. In DEMOLS, two candidate mutation variants are randomly chosen to enhance the search ability by taking their advantages and strengths and two local search mechanisms are designed to improve the ability of local adjustment. Numerical experiments are performed on a set of multi-objective optimization problems, and the experimental results show that DEMOLS has the ability to solve multi-objective optimization problems.

Keywords: Differential evolution, evolutionary algorithm, multi-objective optimization, local search, mutation

1 Introduction

Evolutionary algorithms have been demonstrated to be very useful methods for single and multi-objective optimization and widely applied to various fields over the past several decades [1], [2]. Evolutionary algorithms have the potential superiority to find multiple Pareto-optimal solutions in a single simulation run when applied to handle multi-objective optimization problems. The aim of evolutionary algorithms for multi-objective optimization is to produce a well-converged and well-distributed approximation set near to the true Pareto front [3].

As a branch of evolutionary algorithms, differential evolution (DE) [4] has increasingly received research interests mainly due to its simplicity and efficiency. Recently, some studies have shown that DE can be successfully applied to single and multi-objective optimization [4], [5]. Unfortunately, the search performance of DE can drop to a certain extent when applied to handle hard multi-objective optimization problems which are scalable with respect to objectives and decision variables, and complex with respect to constraints and the shape of the Pareto front [6]. Therefore, it requires researchers to develop and design good evolutionary algorithms to meet this challenge [7]. In this study, we develop two candidate local search strategies to improve the search precision, and propose an enhanced DE using two candidate mutation variants for hard multi-objective optimization problems.

2 Proposed Algorithm

2.1 Mutation Strategy

Conventional DE employs single mutation to generate the new point. In order to overcome drawbacks of single mutation and take cooperative advantages of multiple mutation variants, the proposed algorithm DEMOLS employs two candidate mutation variants “DE/rand/1/bin” and “DE/best/1/bin” [4] to generate the new point $v_i(t)$ according to the following formula:

$$v_i(t) = \begin{cases} x_{r_1}(t) + F \times (x_{r_2}(t) - x_{r_3}(t)), & \text{if } \xi \leq 0.5 \\ x_b(t) + F \times (x_{r_1}(t) - x_{r_2}(t)), & \text{otherwise} \end{cases} \quad (1)$$

where ξ is the uniformly distributed random number from the range [0,1]. “DE/rand/1/bin” and “DE/best/1/bin” can explore the search space randomly and exploit the search space effectively through learning from the best point, respectively.

2.2 Local Search Strategy LSA

Non-uniform mutation [8] can improve the search precision. In order to maintain the diversity of population, inspired by non-uniform mutation, we develop the first local search strategy LSA. In LSA, the following formulas are used:

$$x_{ij}(t)' = x_{ij}(t) + \Delta(t, U_j - x_{ij}(t)) \quad (2)$$

and

$$x_{ij}(t)' = x_{ij}(t) - \Delta(t, x_{ij}(t) - L_j) \quad (3)$$

where the function $\Delta(t, y)$ is defined as

$$\Delta(t, y) = y \times (1 - r^{(1-t/T)^b}) \quad (4)$$

For any given point $x_i(t) = (x_{i1}(t), \dots, x_{ij}(t), \dots, x_{in}(t))$, according to the candidate equations (2) and (3) above, a new point $x_i(t)' = (x_{i1}(t)', \dots, x_{ij}(t)', \dots, x_{in}(t)')$ is generated at a time. Therefore, $2 \times n$ new points are generated according to the following steps:

Step 1. For each index j , according to the equation (2), calculate and save the new point. Therefore, n new points are generated.

Step 2. For each index j , according to the equation (3), calculate and save the new point. Therefore, n new points are generated.

Step 3. Select the best point from the union of the $2 \times n$ new points and the target point $x_i(t)$ as the resulting point.

2.3 Local Search Strategy LSB

In order to improve the search grain, we develop the second local search strategy LSB, which has the ability of local adjustment. In LSB, the following formulas are used:

$$x_{ij}(t)' = x_{ij}(t) + \delta(t, d_j, \xi, b) \quad (5)$$

and

$$x_{ij}(t)' = x_{ij}(t) - \delta(t, d_j, \xi, b) \quad (6)$$

where the function $\delta(t, d_j, \xi, b)$ is defined as

$$\delta(t, d_j, \xi, b) = d_j \times \xi \times (1 - t/T)^b \quad (7)$$

here, ξ is the uniformly distributed random number from the range [0,1], and $d_j = U_j - L_j$ is the width of the j th decision variable.

For any given point $x_i(t) = (x_{i1}(t), \dots, x_{ij}(t), \dots, x_{in}(t))$, according to the candidate equations (5) and (6) above, a new point $x_i(t)' = (x_{i1}(t)', \dots, x_{ij}(t)', \dots, x_{in}(t)')$ is generated at a time. Therefore, $2 \times n$ new points are generated according to the following steps:

Step 1. For each index j , according to the equation (5), calculate and save the new point. Therefore, n new points are generated.

Step 2. For each index j , according to the equation (6), calculate and save the new point. Therefore, n new points are generated.

Step 3. Select the best point from the union of the $2 \times n$ new points and the target point $x_i(t)$ as the resulting point.

2.4 The Framework of DEMOLS

According to the details above, the framework of DEMOLS is outlined as follows:

Step 1. Randomly generate the initial population $P(0)$ with size NP .

Step 2. Set $t = 0$ (t the generation number).

Step 3. Select $\alpha \times NP$ better points from the population $P(t)$ to generate the new population $Q(t)$ according to local search strategies LSA or LSB randomly.

Step 4. The population $P(t)$ performs DE operations and generates the offspring population $R(t)$.

Step 5. Select NP points from the union set $P(t) \cup Q(t) \cup R(t)$ into the next generation population $P(t+1)$.

Step 6. If (the termination criterion is achieved) then stop, go to **Step 7**; else set $t = t + 1$, go to **Step 3**.

Step 7. Output the optimal points of the population $P(t+1)$.

DEMOLS employs the crowding-distance and constraint-handling techniques to fast estimate the crowded density to preserve the diversity of population and effectively handle the constrained multi-objective optimization problems, respectively [3].

3 Experimental Studies

3.1 Benchmark Functions and Parameter Setup

In the experimental studies, we employ a test set of 10 constrained multi-objective optimization problems (CF1-CF10) to investigate the performance of DEMOLS, where CF1-CF7 and CF8-CF10 are two and three-objective optimization problems, respectively. The details of these problems are described in [7].

The parameter settings are summarized as follows: the scaling factor $F = 0.5$, the shape parameter $b = 3$, for CF3, CF5, and CF9, the crossover probability $CR = 0.2$; for CF1, CF2, CF4, CF6, CF7, and CF10, the crossover probability $CR = 0.6$, the maximum generation number $T = 500$, for CF1-CF7, the population size $NP = 200$, the scaling coefficient used for selecting better individuals to perform the local search $\alpha = 0.1$; for CF8-CF10, the population size $NP = 300$, the scaling coefficient used for selecting better individuals to perform the local search $\alpha = 0.05$.

3.2 Performance Metric (IGD)

Let PF^* be a set of uniformly distributed points along the Pareto front PF and APF be an approximate set near to the PF , then the average distance from PF^* to APF is defined as [7]:

$$IGD(APF, PF^*) = \frac{\sum_{p \in PF^*} d(p, APF)}{|PF^*|} \quad (8)$$

where $d(p, APF)$ is the minimum Euclidean distance between the point p in PF^* and the points in APF . If $|PF^*|$ is large enough to represent the PF very well, $IGD(APF, PF^*)$ can measure both the diversity and convergence of APF in a sense. A smaller value of $IGD(APF, PF^*)$ indicates a better quality of APF .

3.3 Experimental Results and Discussion

To intuitively observe the performance of DEMOLS, for CF1-CF10, **Figures 1-10** illustrate Pareto fronts produced by DEMOLS. As shown in **Figures 1-10**, DEMOLS can find multiple Pareto-optimal solutions in a single simulation run, which clearly indicate that the proposed approaches are effective. Furthermore, for CF1, DEMOLS can produce an approximation set with good distribution and convergence. Most problems of CF1-CF10 have multiple disconnected Pareto fronts, which can cause the performance of many evolutionary algorithms for multi-objective optimization.

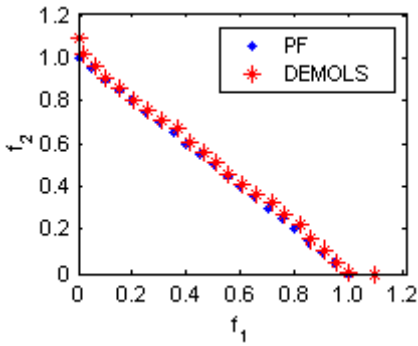


Fig. 1. Pareto front of CF1

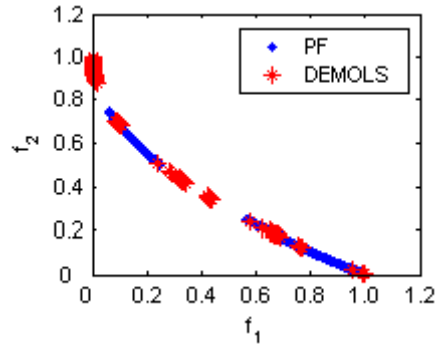


Fig. 2. Pareto front of CF2

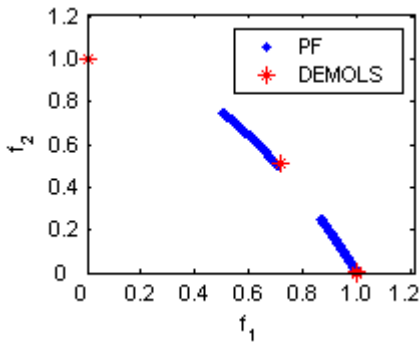


Fig. 3. Pareto front of CF3

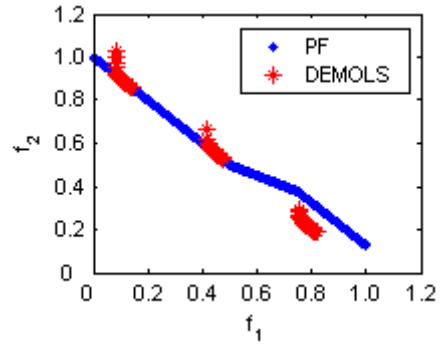


Fig. 4. Pareto front of CF4.

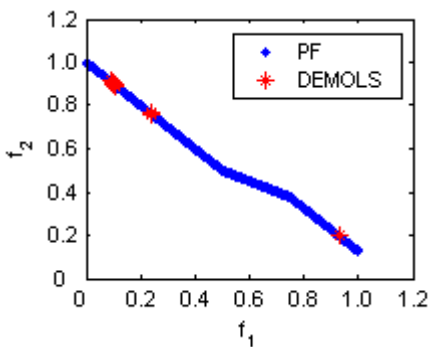


Fig. 5. Pareto front CF5

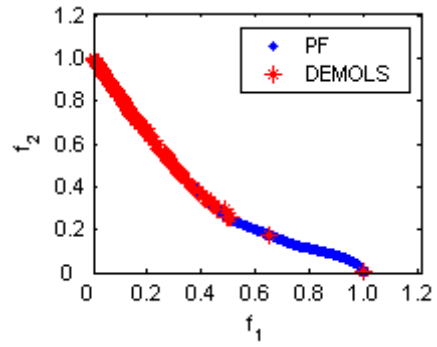


Fig. 6. Pareto front of CF6

3.4 Comparison with Other Evolutionary Algorithms

In order to further verify the relative performance of DEMOLS in terms of solution precision, the experimental results of IGD values (over 20 and 30 runs with respect to DEMOLS and other compared algorithms DECMOSA-SQP [5] and MOEADGM [9], respectively) are summarized in **Table 1**, where the best statistical results with respect to IGD values are highlighted in bold. For all CF1-CF10 with respect to all algorithms, the maximum number of function evaluations is set to $3.0e+05$.

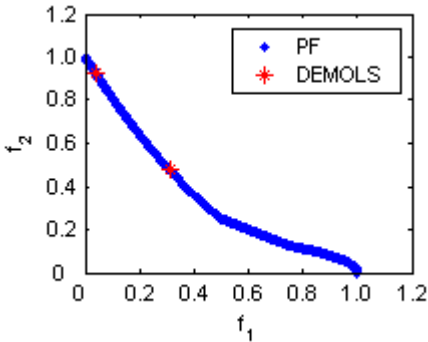


Fig. 7. Pareto front of CF7

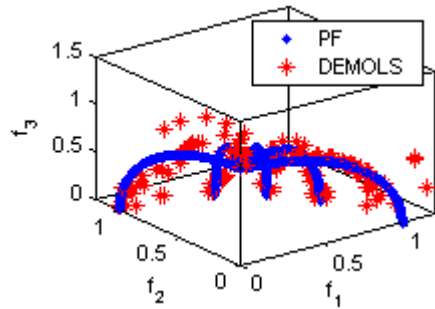


Fig. 8. Pareto front of CF8

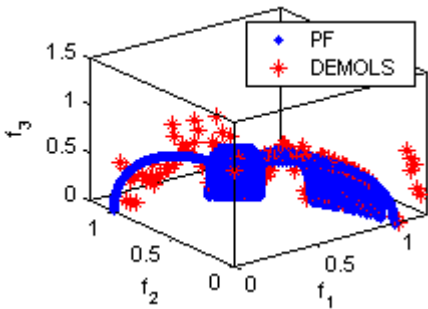


Fig. 9. Pareto front of CF9

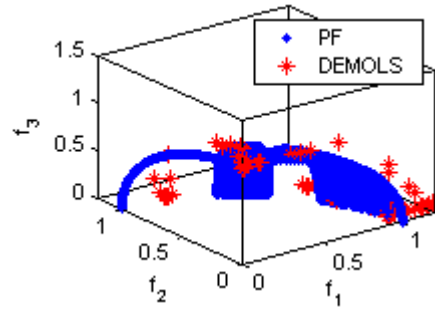


Fig. 10. Pareto front of CF10

According to **Table 1**, we can find that for CF1, CF2, and CF10, the mean results obtained by MOEADGM are better than those obtained by DECMOSA-SQP and DEMOLS, while the best result obtained by DEMOLS is obviously superior to those obtained by MOEADGM and DECMOSA-SQP. For CF3-CF9, the mean, std, best and worst results obtained by DEMOLS are mostly superior to or not obviously worse than those obtained by DECMOSA-SQP and MOEADGM. In sum, for most of CF1-CF10, DEMOLS can perform better than or not worse than other algorithms MOEADGM and DECMOSA-SQP. Therefore, we can conclude that DEMOLS outperforms DECMOSA-SQP and MOEADGM, and that the proposed techniques are beneficial to the search performance of DEMOLS.

Table 1. Experimental results of IGD Metric on 10 constrained test instances

method	status	CF1	CF2	CF3	CF4	CF5	CF6	CF7	CF8	CF9	CF10
DEMOLS	mean	0.012930	0.032201	0.213790	0.072685	0.219964	0.083740	0.261291	0.125541	0.125685	0.485768
	std	0.004173	0.011186	0.053536	0.015146	0.075272	0.030323	0.087388	0.022487	0.016575	0.233828
	best	0.006366	0.013741	0.097358	0.041690	0.116049	0.041919	0.119098	0.088500	0.105950	0.188022
	worst	0.022744	0.050677	0.288021	0.099795	0.412364	0.140139	0.404132	0.182344	0.178696	0.811085
DECMOSA-SQP	mean	0.107736	0.094608	1000000	0.152659	0.412756	0.147824	0.26049	0.176344	0.127132	0.507051
	std	0.195921	0.294282	0	0.466596	0.590779	0.124723	0.259938	0.625776	0.145787	1.19892
	best	0.058125	0.037197	1000000	0.053345	0.096384	0.082309	0.170495	0.097747	0.108369	0.284772
	worst	0.186248	0.257011	1000000	0.301675	0.537761	0.186999	0.373809	0.420665	0.265246	0.811085
MOEADGM	mean	0.0108	0.0080	0.5134	0.0707	0.5446	0.2071	0.5356	0.4056	0.1519	0.3139
	std	0.00250	0.00999	0.07143	0.10144	0.17231	0.00010	0.10030	0.12824	0.04125	0.10384
	best	0.0076	0.0025	0.4138	0.0166	0.0472	0.2070	0.3956	0.2220	0.1034	0.1849
	worst	0.0175	0.0343	0.6525	0.5287	0.7017	0.2074	0.7533	0.7168	0.2754	0.6330

4 Conclusions and Future Work

This paper gives an enhanced DE using two candidate local search strategies and two candidate mutation variants for hard multi-objective optimization, which is called DEMOLS. The local search strategies employ the fine-grain search and can improve the local search ability. Numerical results indicate that DEMOLS has the ability to find the optimal solutions in a single simulation run. In order to further improve the search performance of DEMOLS, in the near future, we can continue to adjust the parameters of DEMOLS and incorporate adaptive and heuristic search mechanisms into DEMOLS during the evolutionary process.

References

1. Sarker, R., Mohammadian, M., Yao, X.: Evolutionary Optimization. Kluwer Academic Publishers, Norwell (2002)
2. Coello Coello, C.A.: Evolutionary Multiobjective Optimization: A Historical View of the Field. IEEE Computational Intelligence Magazine 1(1), 28–36 (2006)
3. Deb, K., Pratap, A., Agarwal, S., et al.: A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6(2), 182–197 (2002)
4. Storn, R., Price, K.: Differential Evolution- A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces. Technical Report TR-95-012, International Computer Science Institute, Berkeley, CA (1995)
5. Zamuda, A., Brest, J., Bošković, B., Žumer, V.: Differential Evolution with Self-adaptation and Local Search for Constrained Multiobjective Optimization. In: 2009 IEEE Congress on Evolutionary Computation (CEC 2009), pp. 195–202. IEEE Press, Trondheim (2009)
6. Zhou, A., Zhang, Q., Jin, Y.: Approximating the Set of Pareto Optimal Solutions in Both the Decision and Objective Spaces by An Estimation of Distribution Algorithm. Working Report CES-485, Dept of CES, University of Essex (June 2008)

7. Zhang, Q., Zhou, A., Zhao, S.Z., Suganthan, P.N., Liu, W., Tiwari, S.: Multiobjective optimization test instances for the CEC 2009 special session and competition. Technical Report CES-487, University of Essex and Nanyang Technological University (2008), <http://dces.essex.ac.uk/staff/qzhang/moeacompetition09.htm>
8. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs, 3rd edn. Springer, Heidelberg (1996)
9. Chen, C.-M., Chen, Y.-P., Zhang, Q.: Enhancing MOEA/D with Guided Mutation and Priority Update for Multi-Objective Optimization. In: IEEE Congress on Evolutionary Computation (CEC 2009), pp. 209–216. IEEE Press, Trondheim (2009)

The Research about Fuzzy Control of Electric Generator Stator Bar-Ends Cure-Heating

Jie Bai, GaoJun Li, and XiaoDong Tan

School of Mechanical Engineering Dalian Jiaotong University, Dalian 116028, China

Abstract. The process of electric generator stator bar-ends cure-heating is a key step during the generator production. The paper makes a thermal chemical analysis of the cure-heating system, then builds a thermodynamics model. To control the non-linear and one-step inertial system, it adopts the method of Fuzzy self-tuning PID parameters to adjust the PID parameters including, and the output according to the real-time heating temperature feedback, thus, controlling the heaters action to achieve the objective of heating temperature control and solving the strong-coupling between the multi-loops of heating and improve the automatic control degree of the cure- heating control system. By simulating in MATLAB results show that the fuzzy PID control is better than traditional PID control. As it can not only exploit the advantages of Fuzzy control, but also possesses the high quality dynamic tracking function and better static precision of PID control.

Keywords: Stator bar, Fuzzy PID control, MATLAB.

1 Introduction

The process of electric generator stator bar cure-heating is a key step during the generator production. The system of electric generator stator bar cure-heating in the past traditional is control instrument majority, and most rely on manual operation. Development of the times and simply can not meet the requirements of process control. At present, the computer control the increasing roles, and use of computers to enhance the operation of the heater control, which can help to reduce maintenance costs and facilitate the selection of control methods. It also can improve the system automatically adjust and the level of the control, to ensure that the thermal efficiency of the heater, and improve the cured production quality.

Bar of the heating process is typical in process control. It is general with the properties of inertia, nonlinearity and Strong coupling. Especially for both ends of the coil over multiple heating, and its coupling is Strong. It is often large overshoot, be difficult to restore balance after disturbance and difficult to match with the set temperature.

If adopting the traditional PID control of a large overshoot, the low steady-state controls accuracy, and especially the poor dynamic control accuracy, it can't meet the contemporary generators on the quality requirements of bar curing. In this paper, the projects of Jinan Power Plant carried out advanced thermal control technology in the end application of curing system, and achieved satisfactory control results.

2. Stator End of Curing Thermodynamic Modeling

2.1 Introduction of Process Flows

The inner insulation of the electric generator stator bar are the glass fabric, the mica paper and the composite layer of epoxy resin which been pressed. The stator bar cure-heating is that make the stator bar dressed by the mica tape with less gelatins, make the tank vacuum and impregnation epoxy resin, heating to made it insulation curing. In the process of electric generator stator bar cure-heating, the whole curve of the insulation curing is been divided into three sections, One group of the curve of the insulation curing is the following:

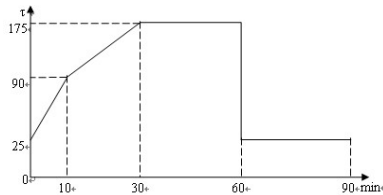


Fig. 1. The curve of the insulation curing

2.2 The Detection of Heat Curing Temperature to Stator Bar and Its Control System Constitution

The electric heating plate actuator of the system, power $P=500W$; the temperature sensing element of the system adopt Pt100thermister, Measuring temperature range - $60^{\circ}C \sim 400^{\circ}C$, Accuracy of measurement $0.5^{\circ}C$. This control system mainly consists of 4 parts: Temperature sensing transmitter constitute by thermoelectric couple and AD595, control device constitute by Siemens PLC S7-200CPU and Expansion Module EM235, the triac actuator and the heating apparatus. System controls the conduction time by control the thyristor in given period, and adjust the power of the electric heating plate to achieve temperature control.

2.3 Analysis of the Heating Process

Using 5 of 500W electrical heating plates which in a straight line reheating the end of the bar in the system, as a big electrical heating plate with 2.5kw、 3.2kg reheating the bar.

When reheating, the electrical heating plates stick to the surface of the bar. Because the mica layer is very thin. So we could consider that the electrical heating plates are direct contact with the copper bar. At the initial stage, the temperature of the heating plates is gradually rising, that is to say that the temperature of the bar's surface is time-dependent. And its adaxial surface is contact with the insulation asbestos board, so assumed the bar's unsteady heat conduction process is one side insulation, another side is the problem of the one dimension big planomural which is the first kind of boundary condition and thickness is δ . Make the surface heat transfer coefficient

$h \rightarrow \infty$, Picasso prisms number $B_i \rightarrow \infty$, so $\beta_1 \delta = \pi/2$. In this formal condition stage, the bar's unsteady heat conduction temperature distribution which under the first kind of condition is

$$\frac{\theta_x}{\theta_0} = \frac{t_x - t_\infty}{t_0 - t_\infty} = \frac{4}{\pi} \cos \left(\frac{\pi}{2} \frac{\delta - x}{\delta} \right) \exp \left[- \left(\frac{\pi}{2} \right)^2 F_0 \right] \quad (2.1)$$

x is the distance between the point which on the surface to the inner of the bar. t_x is the temperature of the point. $\delta=10\text{mm}$, Take bar initial temperature $t_0=20^\circ\text{C}$, the given temperature t_∞ is 175°C .

Fourier number $F_0 = a^{\tau/\delta^2}$, temperature diffusivity $a = \lambda/(\rho c)$,

λ is the bar coefficient of thermal conductivity. ρ is the density the bar. C is the specific heat capacity of the copper bar.

Bar as on the single heating surface, the lower surface is reaction conditions slowly to the surface. The lower surface is the most important place of Non-steady-state response. Assuming the temperature is t_d . When $F_0=0.3$, can be obtained from the above equation: $t_d=81^\circ\text{C}$; when $F_0=1$, $t_d=158^\circ\text{C}$; when $F_0=2$, or to say $\tau=1.7\text{s}$, $t_d=174^\circ\text{C}$. With setting temperature difference is about 1°C . As time goes on, eventually lower surface temperature and set temperature of the bar close unlimitedly, and bar inside temperature distribution from initial temperature distribution influence. Each point temperature in the same rule changes with time.

In the heating process of the bar, heat transfer mode included heat conduction, heat transfer, and heat radiation. Radiation heat is less, for convenient calculation, ignored radiation. According to the heat analysis, assuming bar unsteady heat conduction for side adiabatic, and the another side is the first type of boundary condition Ohira wall one-dimensional unsteady heat conduction problems. Therefore, in this only consider electrical heating plate with air between the convection, and the thermal conductivity between electric heating plate and copper bars. Assuming uniform distribution of the quality bar, $q=7.144\text{kg/m}$, considered that heat quantity in the end of the bar to central conduction L , and L department quality through calculation converted to bar ends quality. According to production practice experience, Bar L department for temperature distribution roughly is $t(x) = 20 + 155 e^{-13736 x} (0 \leq x \leq 4.5)$, by the energy conservation, $M_0 \cdot t_\infty = \int_0^{4.5} c \cdot q \cdot t(x) dx$ given the reduced mass

$$M_0 = 8.27\text{kg}.$$

Set the end of the bar's curing heating controlled parameter for heating temperature, control quantity for electrical heating plate power P . Set the end of the bar quality for M , heat transfer coefficient between bar and air for h , heat exchanger area for A , the temperature of the bar before heat is T_0 , after heating temperature for T . According to thermodynamics,

$$c \cdot (M_0 + M) \cdot \frac{d(T - T_0)}{dt} + A \cdot h \cdot (T - T_0) = P \tag{2.2}$$

Relative to the period of time, T_0 is a constant, and set $T = T - T_0$, Then available with (2.2) corresponding Incremental differential equation,

$$c \cdot (M_0 + M) \cdot \frac{dT}{dt} + A \cdot h \cdot T = P \tag{2.3}$$

Set $\tau = (M_0 + M) c / Ah$, $K=1/Ah$, equation (2) can be written into

$$\tau \frac{dT}{dt} + T = KQ \tag{2.4}$$

Can get the transfer function of the bar's temperature changing between the temperature variation and heating variation:

$$G(s) = \frac{T(s)}{Q(s)} = \frac{K}{\tau s + 1} \tag{2.5}$$

After calculated $K \approx 0.05$, $\tau \approx 297$

$$G(s) = \frac{0.05}{297s + 1} \tag{2.6}$$

The heating system adopt the duty cycle modulated square wave to Control the power's on and down. Its A/D conversion time, computing time, zero-Order Holder and so on caused Control delay. But these delay is ms rank, compared with the entire system terms can be neglected. But for electrical heating board, which heated to set temperature has certain delay. System using five 500W electric heating plates which been straight into a line to heat the end of the bar. That can be regarded as a big electric heating plate which power is 2.5kW, quality is 3.2kg. Electric heating plate materials for chromium nickel alloy, Specific Heat Capacity is 460 ($kg \cdot ^\circ C$), temperature incremental ΔT is $155^\circ C$, So on the whole we use formula 2.6 as a system with transfer function.

3. Fuzzy PID Controller Curing Systems Design and Simulation

3.1 Fuzzy Self-tuning PID Parameters of Controller Structure

Fuzzy self-tuning PID controller is fuzzy controller combined with the traditional PID controller. Using the idea of fuzzy reasoning to determine, depending on the different bias and the deviation change rate on the parameters of PID to on-line self-tuning. After the traditional PID controller gain new, the amount of control object output control. The structure diagram of fuzzy PID controller block diagram shown in Figure 2.

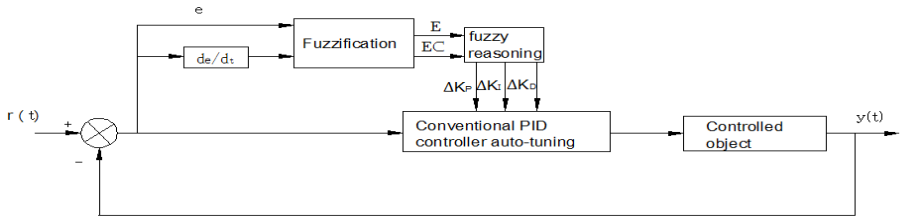


Fig. 2. Fuzzy self-tuning PID controller structure

3.2 Curing System Fuzzy Self-tuning PID Parameters the Controller Design

Base on the researching for fuzzy PID controller theory and bar heat cure control system model, and according to the fuzzy PID controller design steps, aiming at bar heat cure control design fuzzy PID controller. According to the fuzzy controller design steps, derive fuzzy control table. Form.

Every time when sampling error and bias changes, we quantify. And then we can from the control table to find the corresponding amount of control, be multiplied by the scaling factor, at last we can input to the PID controller parameter adjustment.

3.3 PID Control and Fuzzy Self-tuning PID Control Comparison of MATLAB Simulation

To the mathematical model of $G(s) = \frac{0.05}{297s + 1}$ respectively, using the traditional

PID and fuzzy self-tuning PID controlled, we constructed the block diagram of the whole control system simulation in Matlab simulink which shown in Figure 3.

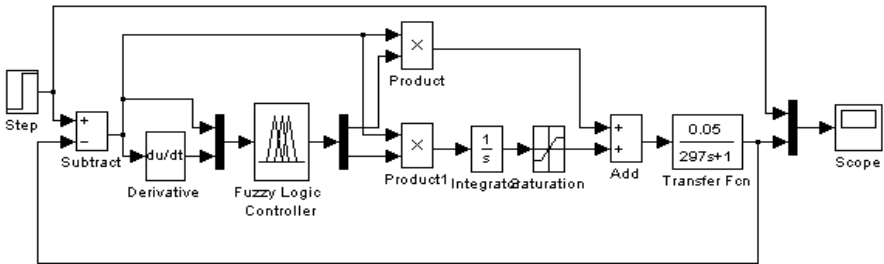


Fig. 3. Fuzzy self-tuning PID control system imitation structure

Because the step signal to the system is the most severe working conditions, if the system under the action of the step signal could meet the requirements of the performance index, then under the action of the input single which in other singles, the system performance can meet the requirements of the general. Therefore, when simulation we using the step signal $rin(k) = 1$ for the given input, observe the input-output response curve.

1) Fuzzy PID control has a smaller overshoot and shorter settling time than the conventional PID control, also has better dynamic response and steady state characteristics.

2) Because the fuzzy PID control of the heating process could according to the rate of change in the bias and deviation automatically adjust the PID parameters, so it has good adaptive ability.

4 Conclusion

1、 In this paper the fuzzy controller is applied to the temperature control of the generator end of the bar which is hot pressing, then achieved good control effect.

2、 Analysis on the heating system is thermal, thermodynamic model is correct.

3、 The design parameters of fuzzy self-tuning PID control method and calculation is correct.

4、 After the Matlab simulated to the temperature of the heating system and fuzzy PID control performance, then comparative analysis with the traditional of the PID control quality, the results show that put the conventional PID control and fuzzy control combined, play features and advantages of both, we can reached the better heating temperature control effect of the end of the bar.

References

1. Su, D., Chen, R., et al.: The Heat-Insulated Analysis and Temperature Controlling on the Hot Pressing Equipment of Motor Stator Coil. *Explosion-proof Electric Machine* 41(03), 42–45 (2006)
2. Liu, H., Dong, H.: Manufacturing Technology for Large Air-cooled Turbogenerator Stator Bar. *Large Electric Machine and Hydraulic Turbine* (2), 10–14 (2006)
3. Yu, J., Ying, A.: The Analysis of Stator Strip Type Bar Ends in Hydro-generator. *Electrical Machinery Technology*
4. Zhang, J., Zheng, Y., et al.: Remedial Method of Defect in Generator Stator Winding Insulation. *Electrical Machinery Technology* 11, 38–41 (2002)
5. Zhang, W.: *Advanced control the theory and method of introduction*. Northwestern Polytechnical University Press, Xian (2000)
6. Xiong, X., Zhang, F.: Intelligent Temperature Control System for Aging Process of Stator Coil. In: *The 7th Industrial Instrumentation and Automation Conference* (2006)
7. Lopez, A.M., Murrill, P.W., Smith, C.L.: Tuning PI and PID Digital Controllers. *Instruments and Control Systems* 429(2) (1969)

The Study on Phase-Control of Vacuum Circuit Breaker for Electric Railway Substation

Jie Bai, KeLiang Shu, and XiaoDong Tan

School of Mechanical Engineering Dalian Jiaotong University, Dalian 116028, China

Abstract. Through analyzing the points of the vacuum circuit breaker with permanent magnet mechanism, which is high accuracy in the switching time (that means the phase accuracy when switching on or switching off the circuit breaker), and small dispersivity in the action time, the method to realize the phase-control circuit breaker is determined. As different kinds of loads characters are analyzed, it is discussed that when we use the phase-control circuit breaker and switching at the proper phases, and we can effectively reduce the inrush and cut down the switching over voltage in the electrical system, and also enhance the stability of the power unit.

Keywords: Electric Railway, Phase-Control, Permanent Magnet Mechanism, Vacuum Circuit Breaker.

1 Introduction

Electric railway intelligent phase-selection control permanent magnet mechanism of vacuum circuit breaker, is designed by the development of the circuit breaker and the demand of the electric railway. The equipment adopts vacuum arcing technology, permanent magnet driven machinery technology, phase-control technology and intelligent integration technology, rated voltage 27.5 KV, single-phase power frequency control of phase-control permanent mechanism vacuum circuit breaker, applicable to electric railway traction power supply system of electric power receiving, control, distribution and protection.

The device can achieve the situation that vacuum switch can be switched accurately at the reserved phase angle by the intelligent phase-selection system and is especially suitable in those where need frequent operation. As the feeder switch, that can effectively decrease of transformer Perceptual load inrush current, reduce power grid operating over-voltage. Its single-phase ac high-pressure move off switch complete device is applicable to the rated voltage 275 KV, frequency 50HZ indoor trolley type switch device, used for distribution electricity open wide type distribution intervals.

2 The Theory of Phase-Selection Control

The phase selection is that the main contact closes at the voltage zeropoint and separates at the current zeropoint. The synchronous closing of breaker can greatly

reduce and even wipe out the inrush current of the condenser bank and the transformer. As a result, it can enhance the breaking capacity of the disconnecter and decrease the switching overvoltage. At first, phase selection detects the zeropoint of the current (or voltage), and then calculate the time that the current zeropoint (or voltage zeropoint) is about to appear, at last it controls the main contact to separate (close) at the current zeropoint (voltage zeropoint) accurately. However, the control must be very accurate. Hence, the volume dispersion of the actuation time in control mechanism should be as small as possible. It can satisfy the demand of the synchronous switch in the precision time of the Permanent magnetic close break-brake. After considering the time compensation of the ambient temperature and the condenser voltage to the permanent magnet, the decentrality of Permanent magnetic close break-brake of the action can be controlled at the set time. Additionally, the accuracy of the F will increase gradually by applying the adaptive algorithm and the aim of the control phase selection will come true. Figure 1 is the theory of phase-selection control.

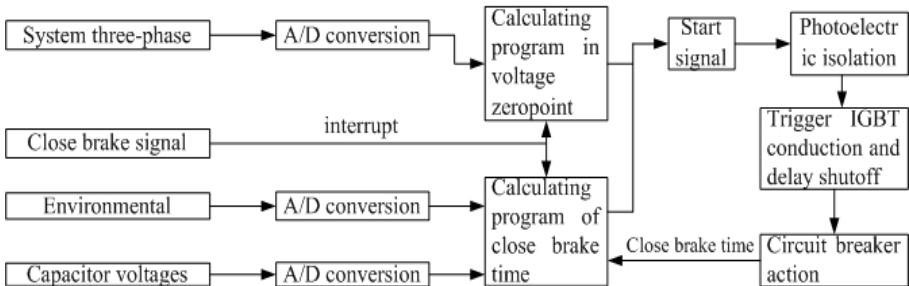


Fig. 1. Drawing of the phase-control theory

3 Requirements of Phase-Selection Control

The first requirements of Phase-selection control detected current or voltage zeropoint, then calculated the current or voltage will be zeropoint moment, then precise control of main contactor in current (or voltage) zeropoint separation (or closed).

Phase-selection control of circuit breaker function analysis: Adopt breaker for phase-selection control can be greatly reduced disconnecter operating over-voltage and inrush current. The task of circuit breaker is to close load current and open circuit short-circuit fault current protection back road electric equipment from damage, but the circuit breaker is in making these close brake, break-brake operations are produced when overvoltage and inrush current phenomenon, will endanger equipment of insulation and power system voltage stability, also can interfere with the circuit or the nearby circuit high sensitivity to the electric equipment working normally.

Based on the no-load transformer transient process when close brake analysis results show that: (1) if close brake, close brake angle $\alpha = 00$ (i.e. $u_1 = 0$ instant close brake), can produce 8 times ~ 15 times rated current inrush current. (2) if close brake, close brake $\alpha = 90^\circ$ (i.e. $u_1 = u_{1m}$), circuit no transient components, because the

contactor close after, magnetic flux immediately enter the stable state. So in no-load transformer off, can use control device to make the breaker contacts between voltage U_m when contacts close, can eliminate impact of inrush current.

Through of single-phase capacitor set open circuit of transient overvoltage of analysis results show that: the contactor open, current phase Φ_0 more small, when current at zeropoint distance of contactor medium strength the greater, also do not appear to recrudescence and Restrike, the gap of arc $\Phi = 0o-90o$ breakdown after current past zeropoint, won't produce overvoltage. So when capacitor set break-brake, can use control device to make the breaker contactor separated produce overvoltage minimum.

(2) The key technology of phase of Circuit breaker close brake, break-brake

Phase-selection control circuit breaker is by phase control devices and high voltage circuit breaker composition. Phase-selection control circuit breaker can achieve the effect of suppressing overvoltage and impact inrush current, its key is operating time accuracy (i.e. the accuracy of phase when close-brake, break-brake).

(3)Circuit breaker property requirements

Circuit breaker close brake, break-brake operations dispersion wants small.

In order to realize phase-selection operation, breaker operator performance should be stable, so that can reduce every operation dispersion, obtain accurate close brake or break-brake phase. Every time close-brake, break-brake time error must be within ± 0.5 ms.

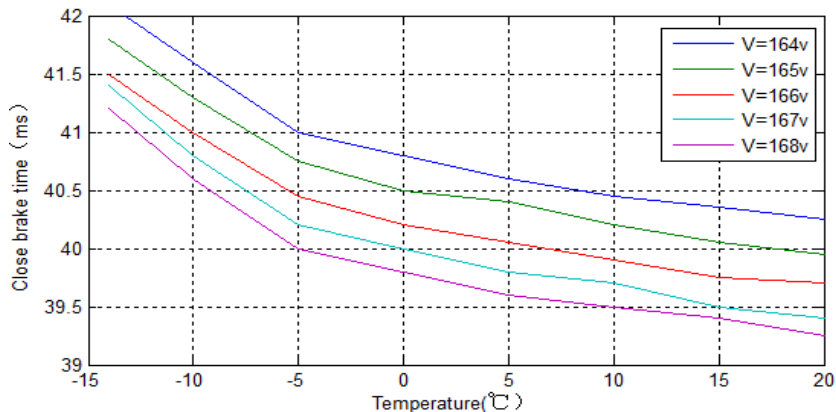


Fig. 2. The relationship among the switching time and ambient temperature and operation voltage

Circuit breaker operation structure is a typical dual steady-state operation structure, namely operation structure has the function that can make contactor from the close brake position to break-brake position, or from break-brake position to close brake position. By power electronics device control, a kind of special design, the electromagnetic system for permanent magnet suture of contactor movement with controllable fuck dynamic energy, permanent magnets, without any external energy, through closed magnetic circuit provides of the lock force, make arcing chamber keep

in close brake position, therefore permanent magnet operation structure can serve as medium-voltage phase-selection breaker operation structure.

Considering the environmental temperature and capacitor voltage on permanent magnetic actuator action time affects compensation, permanent magnetic actuator close brake, break-brake the action time of dispersion can be controlled within ± 1 ms time. Figure 2 is the relationship among the switching time and ambient temperature and operation voltage. In order to guarantee the subsection of real-time line, adopts the adaptive algorithm for every operation of recording and analysis, adjust the next opening-closing timing, such as the operation of equipment, synchronous opening-closing accuracy are gradually improve, truly realize phase-selection control purposes.

4 Control System Design

Switch controller by the following parts: the power supply voltage transform circuit, open-input circuit, analogue input circuit (optional), principal function circuit, the output circuit, open-output circuit, communication circuit, shown in figure 3.

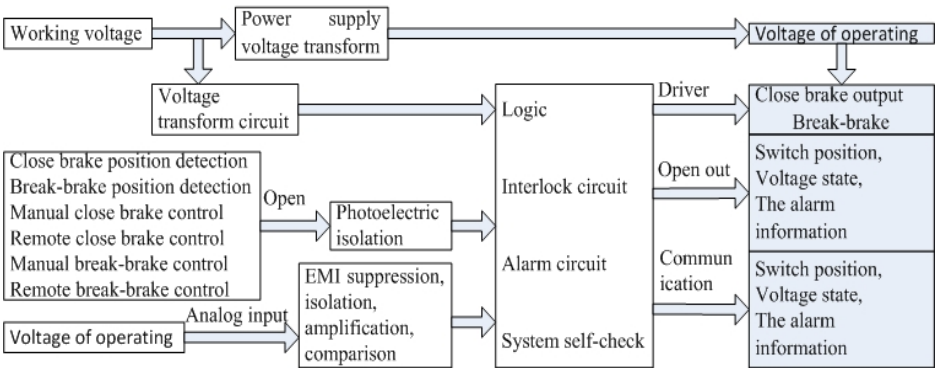


Fig. 3. The theory of the switch controller performance

Power supply voltage transform circuit using AC/DC voltage transform technique, the input voltage transform for DC operating voltage, at the same time to control circuit, power supply voltage transform circuit has the function of stabilization voltage. Open-input circuit using photoelectric isolation technology, of the input signal isolation and improves the system anti-jamming. Analogue input circuit adopt EMI suppression technology and photoelectric isolation technology, and adopts the high precision op-amp and comparator, to ensure the sampling of precision and alarm of the accuracy. Principal function circuit contains logic discriminant, interlock function, alarm circuit, system self-check etc, this part of the circuit based on the quantity and the switch quantity, simulation and system self-check case gives corresponding operation instruction. The output circuit and open-output circuit according to principal function circuit instructions for the corresponding output and open, thus completing the switch

control and relevant information of remote transmission. Communication circuit provides RS-485 communication interface, can pass method of communication to complete switch information (switch position, system related information) uploading, remote control switch close brake, break-brake and other functions.

4.1 SCM's Choice

PIC16 series microcontroller is for automobile kind of product design, thus its anti-jamming, temperature properties are more prominent. Combined with the design of its characteristics, can assure device performance meet the requirements. AD can take plugins way, to demand higher places use external AD and not high use pieces slice within AD.

Determine to PIC16 series microcontroller as development platform, plugins AD, using assembly language.

4.2 The Function of Switch Input and Output

For the position of switch and breaker, relevant state and so on, 8 channel can satisfy the requirements. Occupy MCU an input port. Using existing mature circuit.

The control of close brake and break-brake, alarm, switch state and so on. Using existing mature circuit.

4.3 Communication Function

It provides RS-485 communication port. Using existing mature circuits .(figure 4)

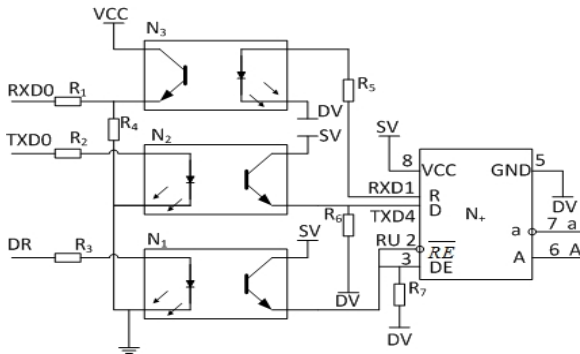


Fig. 4. Communication interface circuit

4.4 Zeropoint Examination Function

Because of the phase-selection requirements, that need to have zeropoint examination function adopts comparative way, got the signal when at zeropoint, to control switch of close brake or break-brake.

5 Conclusion

Electric railway intelligent phase-selection control permanent magnet mechanism vacuum circuit breaker successfully developed, especially after the commissioning and through the technical appraisalment of beijingrailwayboard, Get some conclusions:

Because of the intelligent phase-selection control of circuit breaker according to system current or voltage phase at predetermined phase of break-brake or close brake, so it can reduce overvoltage and inrush current. This can improve circuit breaker ability of break or close, also reduce the impact for power equipment and improve the reliability of operation.

References

1. Ding, F., Duan, X.: An Intelligent Reactive Power Compensator Based On Synchronous Vacuum Circuit Breaker. Proceedings of the CSEE (June 2005)
2. Gao, X.: The Application and Detect of 10 kV Vacuum Circuit Breaker in the Railway Automatic Blocking Power. Shanghai Railway Science & Technology (February 2003)
3. Duan, X., Zou, J., Fang, C., Cong, J.: Synchronized control tactics of phasing vacuum switches when closing capacitor banks. Journal of Dalian University of Technology 43(4), 457–460 (2003)
4. Bonjean, M., Marchal, D., Nicolaye, R., Thiry, P., Legros, W., Falzone, S.: An asymmetrical magnetic actuator for MV circuit-breaker. In: 15th International Conference & Exhibition on Electricity Distribution Cired, NICE 1999, Belgium, June 1-4, vol. 1/13, pp. 301–304 (1999)
5. Mckean, B.A.R., Reuber, C.: Magnets & vacuum – the perfect match. Trends in Distribution Switchgear, 73–79 (1998)
6. Dullni, E.: A vacuum circuit breaker with permanent magnetic actuator for frequent operations. In: Proceedings Discharges and Electrical Insulation in Vacuum, vol. (2), pp. 688–691 (1998)
7. Lammers, A.J.K., Leufkens, P.P., Schoorenberg, G.C.: MV vacuum switchgear based on magnetic actuator. Trends in Distribution Switchgear, 86–90 (1998)
8. Xin, L., Huijun, G., Zhiyuan, C.: Magnetic field calculation and dynamic behavior analysis of the permanent magnetic actuator. In: 19th Int. Symp. On Discharges and Electrical Insulation in Vacuum, Xi'an, pp. 532–535 (2000)
9. Delfino, B., Fornari, F., Gemme, C., Moratto, A.: Power quality improvement in transmission and distribution networks via synchronous switching. In: Transmission and Distribution Conference and Exposition, vol. (1), pp. 367–372 (2001)

General Information Diffusion Method in Risk Analysis

Qiong Li^{1,2}, Jianzhong Zhou^{1,*}, and Donghan Liu³

¹ College of Hydropower and Information Engineering,
Huazhong University of Science and Technology, Wuhan 430074, China

jz.zhou@mail.hust.edu.cn

² School of Mathematics and Physics, Huangshi Institute of Technology,
Huangshi 435003, China

³ Huangshi Institute of Technology, Huangshi 435003, China

Abstract. Flood is a most serious hazard to life and property. The traditional probability statistical method is acceptable in analyzing the flood risk but requires a large sample size of hydrological data. This paper puts forward a improved method based on information diffusion method for flood analysis. Information diffusion theory helps to extract as much useful information as possible from the sample and thus improves the accuracy of system recognition. Furthermore, information diffusion theory helps to extract as much useful information as possible from the sample and thus improves the accuracy of system recognition. This technique contributes to a reasonable prediction of natural disasters risk. As an example, its application is verified in the flood risk analysis in China, and the risks of different flood grades are obtained. Our model yield very good results and suggests that the methodology is effective and practical so that it has the potentiality to be used to forecast the flood risk in flood risk management.

Keywords: information diffusion, flood, risk assessment.

1 Introduction

Flood disasters are more and more frequent in our country, in ordinary flood risk assessment, probability statistics method is the main tools which is used to estimate hydrological variables' exceedance probability. This method has the advantage that its theory is mature and its application is easy. But when it comes to solving practical problems, problems exist in the feasibility and reliability without considering fuzzy uncertainty. Once encountering small sample problem, results based on the classical statistical methods are very unreliable sometimes. In fact it is rather difficult to collect long sequence of extremum data and the sample is often small. So we can use fuzzy mathematical method for comprehensively disaster risk evaluation. This paper uses information diffusion-a fuzzy mathematics method to establish flood risk assessment model with small sample and then applies it to the flood risk analysis in henan province successfully.

* Corresponding author.

2 Information Diffusion

2.1 Definition of Information Diffusion

Information diffusion is a fuzzy mathematic set-value method for samples, considering optimizing the use of fuzzy information of samples in order to offset the information deficiency. The method can turn an observed sample into a fuzzy set, that is, turn a single point sample into a set-value sample. The simplest model of information diffusion is normal diffusion model.

Information diffusion: Let V be the small sample and W be the basic universal field. The information diffusion is defined as a map $\mu : V \times W \rightarrow [0,1]$ satisfying ([1])

- (1) $\forall v_j \in V$, if w_j is the observed value of v_j , then $\mu(v_j, w_j) = \sup_{w \in W} \mu(v, w_j)$
- (2) $\forall v_j \in V$, $\mu(v_j, w_j)$ decreases with $\|w_j - w\|$ increasing
- (3) $\forall v \in V$, $\int_w \mu(v, w)dw = 1$, and \int is replaced by \sum_w in a discrete case.

2.2 General Information Diffusion Method (GIDM)

In Shang et al. ([2]), the uniform information diffusion method (UIDM) is presented and described by

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} (K \frac{\partial u}{\partial x}) \tag{1}$$

where K is a constant. Then the above equation can be written as

$$\frac{\partial u}{\partial t} = K \frac{\partial^2 u}{\partial x^2} \tag{2}$$

However, as most materials have uneven distributions and different consistency themselves, and materials move from higher to the lower level of consistency ([3]). Thus in this paper, a variable K is defined as a function of u i.e. $K(u)$, rather than a constant. Here $K(u)$ can be selected freely if it satisfies certain conditions. Here the simplest form of the quadratic function $K(u) = (u + 1)^2$ is used. It satisfies the following conditions:

- (1) Non-negative increasing function ;
- (2) Irregularity of diffusion velocity, i.e., faster diffusion as consistency increases ;
- (3) There will be a little diffusion even without any consistency.

Then general information diffusion method (GIDM) problem can be described as

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} (K(u) \frac{\partial u}{\partial x}), \\ u|_{t=0} = \delta(x) \end{cases}, \tag{3}$$

For this quasi-linear diffusion equation

$$\frac{\partial u}{\partial t} = K(u) \frac{\partial^2 u}{\partial x^2} + \frac{\partial K(u)}{\partial u} (\frac{\partial u}{\partial x})^2, \tag{4}$$

we can use the MacCormack difference scheme to obtain solutions.

The MacCormack difference scheme is a special form of the Lax-Wendroff difference scheme and it follows a two-step predictor-corrector mechanism. As this technique is used with forward differences on the predictor and with rearward differences on the corrector, it is a second-order-accurate method.

Let the information diffusion function be $v_t(w_n, x) = u(w_n - x, t)$ and $W = \{w_1, w_2, \dots, w_n\}$ be the variable of degree of the sample. Here, $u(x, t)$ is the solution of Eq. (4). Then the information carried by the sample diffuses to the whole

field and the function for the disaster degree can be written as $f_i(x) = \frac{\sum_{i=1}^n v_t(w_n, x)}{n}$

3 Application Example

3.1 Flood Disaster Index

According to the 41 years' practical series material from 1950 to 1990 in henna province, we take disaster area and direct economic loss as the disaster degree's index and by frequency analysis the floods are classified into four grades as seen in table 1, and the four flood grades are small, medium, large and extreme flood.

Table 1. Henan flood disaster rating standard

Disaster level	Inundated area (hm ²)	Direct economic losses (Billion yuan)	Grade number
small flood	0~46.7	0~9.5	1
medium flood	46.7~136.7	9.5~31.0	2
large flood	136.7~283.3	31.0~85.0	3
extreme flood	283.3~	85.0~	4

To raise the grade resolution of flood disaster loss, a new model-projection pursuit (PP) model ([4]) is used for evaluating the grade of flood disaster and the flood degree values are calculated as in table 2.

Table 2. Disaster index based on the projection pursuit model

number	Inundated area (hm ²)	Direct economic losses (Billion yuan)	Degree value	number	Inundated area (hm ²)	Direct economic losses(Billion yuan)	Degree value
i	X(1,i)	X(2,i)		i	X(1,i)	X(2,i)	
1	38.70	7.900	1.369	17	157.30	38.600	2.486
2	38.50	7.800	1.366	18	283.30	85.000	3.498
3	32.10	6.500	1.315	19	556.90	67.100	3.967
4	24.20	4.900	1.256	20	649.50	194.900	3.987
5	36.40	7.400	1.350	21	602.30	180.700	3.979
6	46.70	9.500	1.432	22	446.50	134.000	3.897
7	97.60	21.700	1.895	23	694.90	208.500	3.992
8	60.40	12.800	1.552	24	72.92	9.900	1.574
9	112.60	25.200	2.033	25	148.13	20.656	2.156
10	56.20	11.800	1.515	26	203.92	27.521	2.559
11	80.60	17.600	1.736	27	179.10	24.858	2.389
12	136.70	31.000	2.258	28	375.46	94.927	3.726
13	259.10	76.100	3.363	29	301.24	47.836	3.233
14	200.10	54.400	2.915	30	141.97	116.439	3.368
15	280.10	83.800	3.481	31	279.84	121.127	3.699
16	236.10	67.600	3.209	32	172.06	51.619	2.750

3.2 Flood Risk Evaluation Based on General Information Diffusion

Based on the disaster degree values of the 32 samples (see Table 2), that is the sample

points set $W = \{w_1, w_2, \dots, w_n\}$. We might as well define $u|_{t=0} = \sqrt{\frac{3}{\pi}} \exp(-3x^2)$

and apply the MacCormack technique, where the interval of x is set as $[-4,4]$. Then the information carried by the sample diffuses to the whole field and the function for

disaster degree can be written as $f_t(x) = \frac{\sum_{i=1}^n v_i(w_n, x)}{n}$, then disaster risk estimate

namely probability risk value in Henan province is calculated out. The relationship between the recurrence interval N and probability P can be expressed as $N = 1/P$, then the exceedance probability curve of flood to disaster degree value are shown as Figure 1.

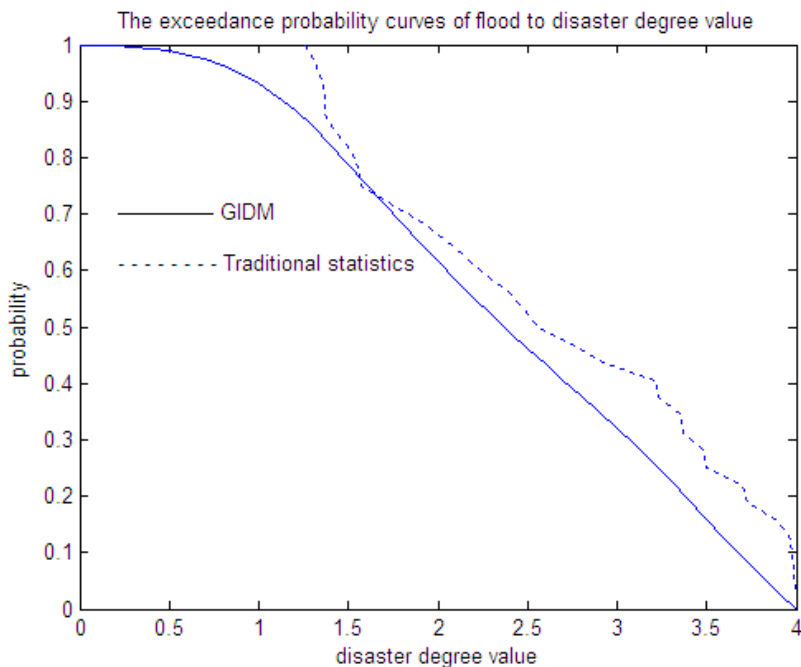


Fig. 1. The exceedance probability curves of flood to disaster degree value based on General Information Diffusion and frequency analysis

Due to the standard of four grades, so we have Chen (2009) ([5]):

- (a) If $1.0 \leq H \leq 1.5$, then desertification degree belongs to small (1 grade).
- (b) If $1.5 < H \leq 2.5$, then it belongs to medium (2 grade).
- (c) If $2.5 < H \leq 3.5$, then it belongs to large (3 grade).
- (d) If $3.5 < H \leq 4$, then it belongs to extreme (4 grade)

The result in Figure 1 illustrates the risk estimation i.e. the probability of exceeding the disaster degree value. From Figure 1 we know the risk estimation is

0.1592 when the disaster index is 3.5, in other words, in Henan Province, floods exceeding 3.5 degree value (extreme floods) occur every 6.2814 years. Similarly, the probability of floods exceeding 2.5 degree (large floods) is 0.4619, namely Henan Province suffers the floods exceeding that intensity every 2.1650 years. The frequency and the recurrence interval of the floods of the four grades are shown in Table 3. This indicates the serious situation of floods in Henan Province whether on the aspect of frequency or intensity. In Figure 1 the curve so estimated is compared to the frequency analysis based on the results of Jin et al. ([4]). Figure 1 shows that our results are consistent with those of frequency analysis. It also means that normal information diffusion is useful to analyze probability risk of flood disaster. Because the flood disaster belongs to the fuzzy events with incomplete data, therefore, the method proposed is better than frequency method to analyze the risk of the flood disaster.

Table 3. Flood disaster risk evaluation values

Disasters level	Small flood	Medium flood	Large flood	Extreme flood
Exceedance probability risk	0.9310	0.7885	0.4619	0.1592
Recurrence interval(year)	1.0741	1.2682	2.1650	6.2814

4 Conclusion

Floods occur frequently in China and cause great property losses and casualties. In order to implement a compensation and disaster reduction plan, the losses caused by flood disasters are among critically important information to flood disaster managers. This study develops a improved method of flood risk assessment disasters based on information diffusion method, and it can be easily extended to other natural disasters. It has been tested that the method is more reasonable and stable.

Acknowledgments. This work is supported by a grant from the National Basic Research Program of China (Project No.2007CB714107), a grant from the Key Projects in the National Science and Technology Pillar Program (Project No. 2008BAB29B08), and a grant from the Special Research Foundation for the Public Welfare Industry of the Ministry of Science and Technology and the Ministry of Water Resources (Project No. 201001080).

References

1. Huang, C.F., Shi, Y.: Towards Efficient Fuzzy Information Processing-Using the Principle of Information Diffusion. Physica-Verlag (Springer), Heidelberg, Germany (2002)
2. Shang, H.J., Lu, Y.C., Jin, P., Zhang, L.: Information diffusion method in risk analysis. In: Proceedings of the 5th International FLINS Conference, Computational Intelligent Systems for Applied Research, pp. 189–197. World Scientific Publishing Co. Pre. Ltd., Singapore (2002)
3. Gu, C.H., Li, D. Q.: Mathematical and Physical Equations. Higher Education Press, Beijing (2002) (in Chinese)
4. Jin, J.L., Zhang, X.L., Ding, J.: Projection Pursuit Model for Evaluating Grade of Flood Disaster Loss. Systems Engineering-theory & Practice 22(2), 140–144 (2002)
5. Chen, S.Y.: Theory and model of variable fuzzy sets and its application. Dalian University of Technology Press, China (2009)

Multi-agent Based Bank Queuing Model and Optimization

Qicong Zhang

School of Information Engineering
Shandong Institute of Trade Unions' Administration Cadres
Jinan, 250001, China
zqc198002@163.com

Abstract. In order to solve the current problem that banking customers wait too long, the agent-based simulation model of queuing system is proposed. The customer, queue, server and ATM are abstracted different agents. The operation of bank queuing system was simulated through the interaction among agents. In our model, an algorithm for adjusting the server queue dynamically was presented to reduce the average waiting time of customers. The experiment and analysis show that our model can simulate the operation of bank queuing system actually, which utilizes the existing resources and promotes the customer satisfaction.

Keywords: intelligent agent, bank queuing model, algorithm optimization.

1 Introduction

Bank queuing problem is more serious with each passing day, and the long waiting time for business one of the factors that discount the customer satisfaction [1]. At present, scholars have done some actual research on this situation and provided some advice for improvement[2]. However, it is difficult to carry out more comprehensive and further development, due to the limitation of survey samples, time, energy and scale in the practical survey. On the other hand, there is also some research to produce deterministic models and give the improving measures through mathematical modeling methods [3]. But bank queuing problem is more complex, variable and random. The emerging research of computing experimental method can be used to describe the problem better. Literature [4] uses this method to build a bank queuing model. The model constitutes different queues according to different types of business. In order to use the bank resources more reasonable, the model gives an algorithm for adjusting the business types of server dynamically. But now most of banks in our country adopt bank queuing machine, so the model can't describe the present bank queuing phenomenon effectively. In addition, the model in the literature adjusts the server queue according to the numbers of people in different business queues. And the hypothesis is that the coming rates of people of various business types. However, it is greatly different to the actual situation. Therefore, the algorithm is likely to result in the problem that the customers of a smaller type of business have no or tardy service.

In our study, we build a multi-agent based model of bank queuing system, which simulates the operation of bank queuing system through constructing different agents in the system and implementing the interaction of different agents. Although the current bank queuing machine embodies equity, it couldn't effectively improve service efficiency. In addition, literature [5] points that the allocation of bank business windows is unreasonable and the utilization of self-help service is not high. We reset the service window types, and propose an algorithm that adjusting the service types of server queues according to estimate the service time of the specific business type. Moreover, the algorithm enhanced the lobby managers' capability of providing services and shunting business. In this situation, the algorithm makes customers complete the business in ATM instead of the already busy service window. Our model can not only simulate the phenomenon of bank queuing system actually, but also control the cost of bank operation effectively and reduce the average customer waiting time.

2 Agent-Based Model of Bank Queuing System

Bank queuing system model consists of seven types of agents, which are customer agent, server agent, ATM agent, queuing machine agent, lobby manager agent, queue management agent and multi-queue agent, respectively. The system structure is shown in figure 1.

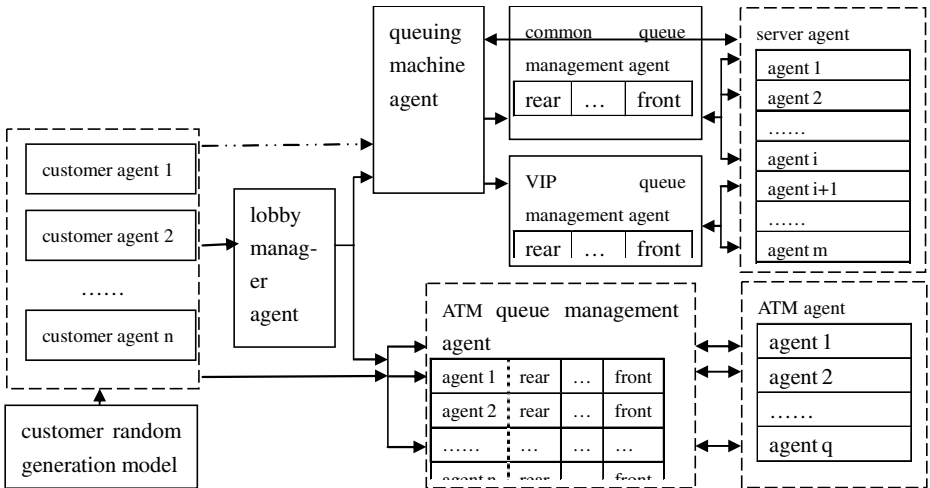


Fig. 1. The structure of bank queuing system based on agent

Random generation module generates a customer agent and assigns it initial values of relevant attributes, such as the business type to handle, the familiarity with the business and whether for advanced users(very important person, VIP), etc. If the customer agent is familiar with the business, it will directly decide which service way

to accept according to the type of business. It can choose the service of ATM agent by the interaction with ATM management agent or the service of server agent by the interaction with queuing machine agent. Otherwise, it will decide the service type through the interaction with lobby management or queuing machine agent.

The lobby manager agent decides to ask the customer agent forwardly or accept the customer agent passively, according to the initiative value that preliminarily designed. When the lobby manager agent finishes the interaction with the customer agent, it will send the shunting message based on business type to the customer agent.

Queuing machine agent divides business types into VIP business and common business. And the algorithm determines whether common business is divided into short, medium and long business based on the service time estimation. It is assumed that short business can be completed by both ATM agent and server agent while medium and long business can be completed only by server agent. When queuing machine agent interacts with customer agent, it will provide the relevant information to customer agent, such as the number of people waiting the specific business and the total number of waiting people. At the same time, the queuing machine agent will send server agent the message including the arrival of a customer agent and the business type, accept the message of business-type that the server agent has complete and update its information base.

Queue management agent consists of three different agents, which are common queue management agent, VIP queue management and ATM queue management agent, respectively. The common queue agent and VIP queue agent are in charge of the interaction between the customer agent and server agent. When the customer agent finishes the interaction with queuing machine agent, it will be inserted into the tail of common queue or VIP queue according to the type of business. When server agent sends the message of service start to common queue or VIP queue, the queue management agent provides specific information related to agent. ATM queue management agent is in charge of the interaction between the customer agent and ATM agent. The customer agent attends the tail of a queue according to its own will. After ATM agent sends the service information, ATM queue management will provide the information of the customer agent in the head of queue to ATM agent. When a customer agent begins to accept service, the queue management agent will cancel it from the queue.

Server agent is mainly responsible for the interaction among common queue management, VIP queue management agent, queue machine agent and other server agent. Because banks generally provide VIP counters, the initial model sets up several specific server agents to provide service for VIP customer agent. In general, server agents in charge of different type of business send the service message to the corresponding queue management agent. Based on first-come first-serve algorithm, the customer agent in the head of queue will be selected to obtain a service. If the implementation is optimized to cancel VIP counters, all server agents can provide services to for all types of business. A server agent uses the server queue dynamic adjustment algorithm to determine the service type of the next customer agent according to the business type which other server agent provide. When a server agent finishes a service of a customer agent, the message will be sent to queue machine agent. The server agent receives the information of a customer agent which is provided by queue machine agent.

After the interaction with the corresponding ATM queue management, ATM agent will select the customer agent in the head of queue to provide service according to the first-come first-service algorithm.

3 Server Queue Dynamic Adjustment Algorithm

At present, most of banks set up specific VIP counters. But the number of VIP customers in reality is very small, it often appears that VIP counters are free while integrated service counters are overcrowding. In addition, due to the different business at different times, some customers for the business of a short time need to wait for a very long time. Such phenomenon will greatly affect customer satisfaction. In the increasingly fierce competition of banking industry, these problems should be solved urgently. Otherwise, the loss of customers will bring significant risk to banks. Therefore, we propose the server queue dynamic adjustment algorithm based on shortest job first algorithm in the computer operating system. We make the following assumptions:

(1) Each server agent can handle all types of business, and VIP counters are canceled.

(2) Each customer agent has only one business, and the business can be completed in a counter.

(3) Based on the time estimation, common businesses are divided into short, medium and long business. Short business can be completed in ATM and server, while medium and long businesses are completed only in server.

Abstract data type of server agent is:

```
DSP{
  Agent object:  $D = \{SA_i \in \text{ElemSet}, i = 1, 2, \dots, m, m \geq 0\}$ .
               ElemSet:  $= \langle \text{Agent\_id}, \text{type} \rangle$ ;
  Basic operation P: Sget( $a$ ) is the number of server agents which provide
                    service for  $a$ -type business. And  $a = V, L$  and  $S$ 
                    denotes VIP, long business and medium and short
                    business, respectively.
                    Set( $a$ ) denotes that the service object is set to the first
                    customer agent of the queue for  $a$ -type business.
}
```

Abstract data type of common queue management agent is:

```
DSP{
  Agent object:  $D = \{CQA \in \text{ElemSet}\}$ ;
               ElemSet:  $= \langle \text{Agent\_id}, \text{type} \rangle$ ;
  Basic operation P: Qget( $t_c$ ) denotes getting the time when the head agent of
                    common queue is inserted into the queue and assigning it
                    to  $t_c$ .
                    Qget( $a, t$ ) denotes getting the time when the first  $a$ -type
                    agent is inserted into the queue and assigning it to  $t$ .
```

Long(a) denotes the number of customer agents for a -type business. If there is no parameter, it indicates the total number of customer agents in the common queue.

```

}
Abstract data type of VIP queue management is:
DSP{
  Agent object: D={VQA ∈ ElemSet};
                ElemSet:=< Agent_id,type>;
  Basic operation P: Qget( $t_v$ ) denotes getting the time when the head agent of
                    VIP queue is inserted into the queue and assigning it to  $t_v$ :
                    Long() denotes the total number of customer agents in
                    VIP queue;
}

```

The core of algorithm is described as follows:

```

Service_Adjust()
{ if VQA.Long()>0
  then { VQA.Qget( $t_v$ );
        CQA.Qget( $t_c$ );
        if ( $t_v < t_c$ ) or ( $SA_i.Sget(v) < [m/2]$ )
        then  $SA_i.Set(v)$ ; }
else if CQA.Long(s)>0
  then { CQA.Qget( $t_c$ );
        CQA.Qget( $a, t$ );
        if  $t > t_c$ 
        then if  $SA_i.Sget(l) \geq m-1$ 
        then  $SA_i.Set(s)$ ; } }

```

4 Model Implementation and Experiment Analysis

4.1 Model Implementation

This study constructs the model of bank queue system using netlogo application platform[6]. In the initialization phase of the model, it creates a bank queue to simulate the scene of business operation, including server agent, ATM agent, queuing machine agent, and lobby manager agent. The agent attributes and conduction rules are also initialized in this phase. In addition, common queue management agent, VIP queue management agent and ATM queue management agent are created to be charge of the management and interaction of different types of customer agents. The drawing and monitor tools are initialized to track records and analysis relevant data, such as

the average waiting time of customers for different types of business. The operation time of bank queue system is preset by human. Relevant records and data are exported by drawing and monitor tools.

4.2 Experiment Simulation and Analysis

The parameters of our simulation system are set as follows:

- (1) The bank business hall has 4 business server counters, 2 ATMs. That is $m=4$, $q=2$.
- (2) The number of VIP counters is equal to the number of VIP threshold in the dynamic adjustment algorithm. That is $M_2=[m/2]=2$.
- (3) A long business threshold, that is $M_1=m-1=3$.
- (4) The time of short business, medium business and long business in common business are estimated as 0-5 minutes, 5-10 minutes and more than 15 minutes, respectively. The time of VIP business is estimated as 0-15 minutes.

Due to different banks are in different locations, different business types of customers are different arrival rates. Therefore, the simulation system is implemented by setting different arrival rates of customers for the two experiments. Assume that the total of customers is 150 people. In the first experiment, the number of customers for short business, medium business, long business and VIP business are 70, 55, 18 and 7, respectively. In the second experiment, the number of customers for short business, medium business, long business and VIP business are 70, 55, 18 and 7, respectively. The system runs for 2 hours, and the experiment data are recorded as follows:

Table 1. The experiment result of simulation of bank queuing system (unit: minute)

	the first experiment		the second experiment	
	the dynamic adjustment algorithm	the original algorithm	the dynamic adjustment algorithm	the original algorithm
the average waiting time for short and medium business	11.2	15.0	14.8	23.1
the average waiting time for long business	18.7	20.2	23.2	24.0
the average waiting time for VIP business	3.3	3.4	6.0	5.7
the average waiting time for all business	13.5	17.7	17.8	21.1

From the above experiment data, we find that using the dynamic adjustment algorithm, the average waiting time of customers is generally reduced, and the average waiting time of customers for short or medium business is greatly reduced. In the second experiment result, the average waiting time of customers for VIP and long business remains unchanged. Although the number of customers for VIP and long business has increased in the second set, a portion of customers for short business are diverted to ATM. The dynamic adjustment algorithm is helpful to the customers for

short and medium customers, which is still the acceptable range of the customers for VIP and long business.

5 Conclusion and Future Work

To solve the problems such as the long average waiting time of customers, unreasonable allocation of bank counters and low utilization of self-help service, we construct and implement the simulation model of bank queuing system based on agent technology. The model takes to reset the types of service counters, enhance the lobby manager's capability of providing services and shunting business. And an algorithm for adjusting the server queue dynamically is proposed. The simulation experiment shows that our model can not only simulate the phenomenon of bank queuing system actually, but also control the cost of bank operation effectively and reduce the average customer waiting time. It should be noted that the processing time for each type of business is assumed by the model, which is different from reality. In addition, some details such as different money amount for the same type of business may result in different processing time. Thus, the future work focuses on how to subdivide business types and improve the intelligence of queuing machines.

References

1. Song, G.: Research on the Affect on of Customers' Satisfaction in Retail Banking. *Journal of Central University of Finance & Economics* (3), 33–38 (2010)
2. Jin, Y.S., Ming, X., Li, X., Wen, J.Y., Jin, D.: Customer-centric optimal resource reconfiguration for service outlet. In: *IEEE/INFORMS International Conference on Service Operations, Logistics and Informatics*, pp. 754–759. IEEE, Chicago (2009)
3. Chao, H.J.: A novel architecture for queue management in the ATM network. *IEEE Journal on Selected Areas in Communications* 9(7), 1110–1118 (1991)
4. Li, C., Yang, G., Yin, Y.: Agent Based Simulation of Bank Queuing System. *Computer Simulation* 25(12), 277–280 (2008)
5. Zheng, Q.: Commercial bank queuing problems and countermeasures in our country. *Business Culture(Academic)* (12), 128 (2008)
6. Wilensky, U.: NetLogo: Center for Connected Learning and Computer-Based Modeling[EB/OL]. Northwestern University, Evanston, IL, Center for Connected Learning and Computer-Based Modeling (1999), <http://ccl.northwestern.edu/netlogo>

Simulation of Agent-Based of Intelligent Public Transit System

Qicong Zhang

School of Information Engineering
Shandong Institute of Trade Unions' Administration Cadres
Jinan, 250001, China
zqc198002@163.com

Abstract. With agent theory and technology, a simulation model of bus transit system based on Intelligent Transport System (ITS) is proposed. The buses, passengers, stations and signal lights are abstracted as different agents, and vehicle scheduling agent is also introduced. The operation of bus transit system is simulated through the interaction among passenger agent, vehicle scheduling agent, station agent, bus agent and signal light agent. In order to solve the soaring passengers' phenomenon in some stations, an improved algorithm that adjusts vehicle scheduling algorithm dynamically is proposed. The experimental result and analysis show that our model can realistically simulate the phenomenon of transit system operation, utilizing bus resources reasonably and reducing passengers' waiting time, energy consumption and operating costs of the bus company.

Keywords: agent-based simulate, intelligent public transit system, vehicle scheduling algorithm.

1 Introduction

Priority to the development of city public transport is an important way to reduce energy consumption and alleviate city congestion. How to plan and schedule public transport system scientifically and reasonably, is a quite valuable research subject. The public transport system is a complex system, involving a wide range of issues and fields. Agent-Based Modeling and Simulation(ABMS) is very suitable for the research of this kind of problem. At present, some research achievements have been made on the bus decision support system [1], travel information and navigation [2], bus running service evaluation [3] and public transit network planning [4]. In this paper, our work focuses on the scheduler of common bus operating system.

Literature [5] constructed a bus rapid transit operation model. However, as a new public transport form, the bus rapid transit system is just at the starting stage. Because the current number of the cities that introduce the rapid transit system is not high, common bus is still the main transport form of public transport. In this paper, a more general sense of the public transport system model is built. In our method, the transport system is abstracted as a multi-agent model, consisting of bus agent, intelligent station agent, passenger agent, vehicle scheduling agent, signal light agent, pedestrian agent

and vehicle agent. Moreover, in order to solve the soaring passengers' phenomenon in some stations in a short period of time, we improved the dynamic adjustment of vehicle scheduling algorithm that proposed in literature [5]. The improved algorithm can not only increase the frequency of buses more scientifically and reasonably according to real-time information provided by intelligent platform, but also emit public transport vehicles of different capacity on the basis of the number of passengers waiting in specific station. Therefore, our algorithm can guarantee to reduce customers' average waiting time and energy consumption, making it more suitable for common bus with the advantages of low-carbon and green. Experimental result and analysis reveal that our model can better simulate the real public transport system. Meanwhile, it actually reduces bus operation costs, protects urban environment, and increases customers' satisfaction.

2 Agent-Based Model of Bus Intelligent Transit System

Agent-based model of bus intelligent transit system contains six types of agents, which are bus agent, passenger agent, station agent, signal light agent, pedestrian agent and vehicle agent, respectively. In order to reduce the burden of interactivity between passengers and buses and to meet actual needs, the vehicle scheduling agent is introduced to our system. Figure 1 shows the system structure.

In our system, the station agent is the connection of the passenger agent and the bus agent. Due to ITS, each station agent can be responsible for managing a passenger agent queue, and sending the number of waiting passengers to the vehicle scheduling agent. When a passenger agent generated and assigned randomly to a station, the station agent inserts this passenger agent to the tail of the passenger agent queue to register. After the interactive of the passenger agent and the station agent, the station agent will also send the information of the passenger agent queue to the vehicle scheduling agent. At the same time, the station agent will receive the message that whether a bus agent arrive at this station. If a bus agent arrives, for some passengers to get off the bus, the specific number of passengers getting on the bus is decided according to the current capacity surplus sent by the bus agent and the number of the passenger agent queue in this station. When a passenger agent finishes off getting on the bus, the station agent will remove it from the queue.

Passenger agent is generated by random generation module in accordance with different traffic peak period. After a passenger agent arrives at each station randomly, it will create a number of getting off station and finish the interaction with the station agent. When a bus agent arrives at this station, the station agent will decide whether this passenger agent get on the bus. Then, when this passenger agent reaches the corresponding getting off station, the passenger agent will send the message to the bus agent and get off the bus to die out.

Vehicle scheduling agent is responsible for the production and scheduling of the entire transit system. In addition, it will accept the message sent by bus agent. If a bus agent is generated toward a certain direction, the number of vehicles scheduled toward this direction will reduce 1. Otherwise, if a bus agent toward a certain reaches the terminal station, the number of vehicles scheduled toward reverse direction will add 1. Depending on the conditions such as the traffic peak, the message from station

agent and bus agent, and the number of vehicles which can be scheduled currently, the vehicle scheduling optimization algorithm determines departure time interval, how to start and authorize bus agent.

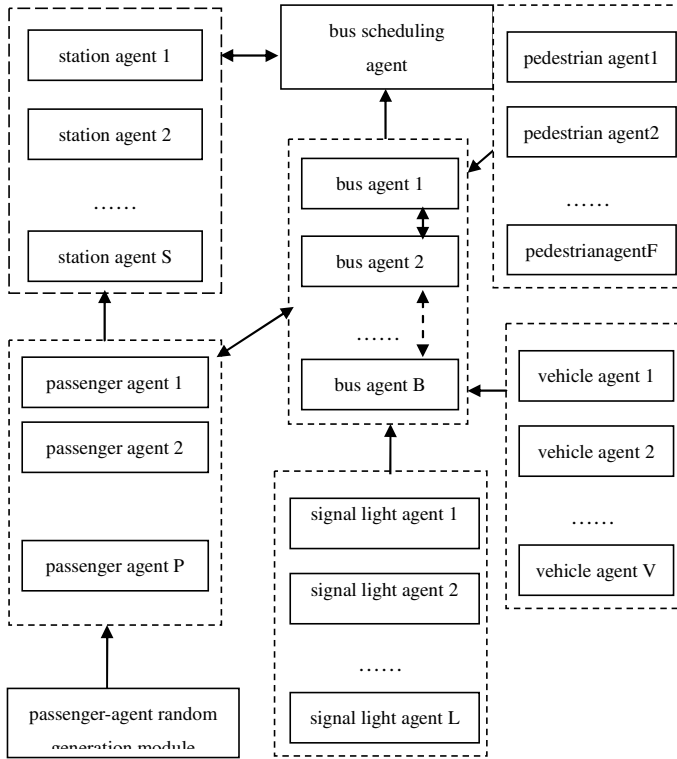


Fig. 1. The structure of agent-based bus intelligent transit system

The generation of bus agent is depended on the vehicle scheduling agent. When signal light agent appears in front during the running process, the bus agent will decide to run or stop according to the state of signal light and whether there is a pedestrian or vehicle. When the bus arrives on a station, the stop time of the bus agent is decided by the number of passengers getting off the bus and the number of passengers getting on the bus that provided by the station agent. If there is other bus agent stopping ahead in the running period, the vehicle scheduling agent will authorize the bus agent determine the overtaking or stop. In addition, the bus agent sends the message of current operation or station to the vehicle scheduling agent. When arriving at the terminal station, the bus agent will die out. The work process of the bus agent is shown in figure 2.

The signal light agent changes its state according to the initial set of traffic lights alternate cycle. The pedestrian agent and the vehicle agent decide to move or rest based on the signal light agent.

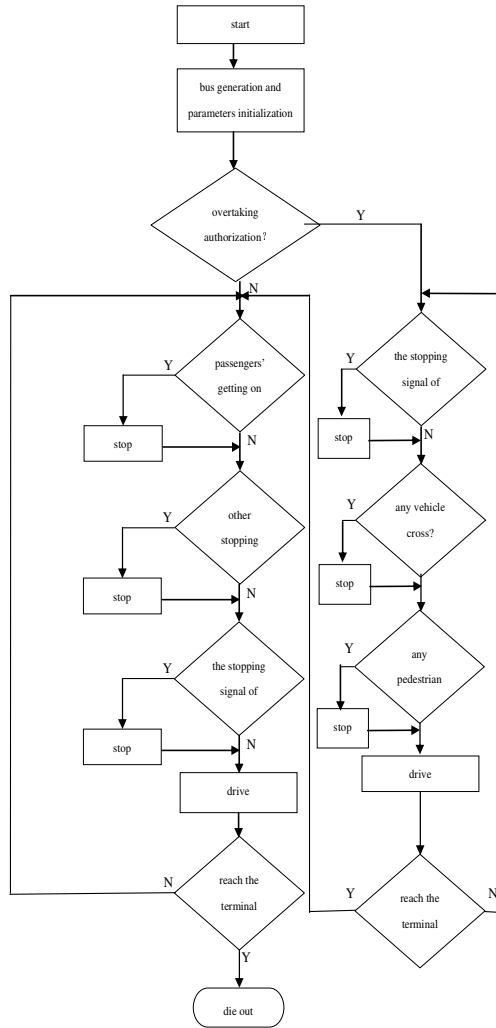


Fig. 2. The work process of the bus agent

3 Improved Dynamic Vehicle Scheduling Algorithm

For fixed-line, the bus company usually uses the same type and the same maximum capacity of vehicles, and develops a different time of departure schedules based on the common and peak period of the passenger flow. Departing time interval is generally a fixed value, which is longer in common period and shorter in peak period. However, this scheme is too simple to solve problems such as the soaring passengers' phenomenon in some stations in a short period of time. In order to reduce passengers' waiting time, bus company operating costs and energy consumption, this paper

proposes an improved dynamic vehicle scheduling algorithm. The algorithm makes the following assumptions:

- 1) Departure time intervals are divided into common period and peak period, which are T_1 and T_2 , respectively.
- 2) According to different capacity, buses are divided into two types of medium and large, whose maximal capacities are C_1 and C_2 , respectively.
- 3) The conventional vehicle speed is V , and the authorized vehicle speed is \bar{V} , where $V < \bar{V}$;
- 4) The station appears to the soaring phenomenon when the number of waiting passengers is not less than \bar{C} , and the distance between neighborhood station is equal;
- 5) No traffic jams when the vehicles normally moving.
- 6) Based on ITS, the station agent can send real-time statistical information of the number of waiting passengers.

The concepts, functions and procedures used in the algorithm description are defined as follows:

Definition 1: The station agent is abbreviated as $A_{i(s)}$ ($i = 1, 2, \dots, s, s \geq 0$), where i denotes station number.

Definition 2: The function $g(A_{i(s)})$, denotes obtaining the total number of waiting passengers in the i^{th} station.

Definition 3: The process $n(T)$ denotes adjusting departure interval. If the previous $T = T_1$, then adjusted to T_2 , and vice versa.

Definition 4: The vehicle scheduling agent is abbreviated as A_i ($i = 1, 2, \dots, w, w \geq 0$), where i denotes the of bus line number.

Definition 5: The process $s(A_{i,j}, n)$ denotes that vehicle scheduling agent sends the i^{th} bus line, j type bus with \bar{V} speed directly to the station where appears the soaring passengers' phenomenon. The value of j is assigned to K for mid-size bus and Q for large bus. n is the number of departure vehicle. If omitted, it means all inventory vehicles.

Definition 6: The function $n(A_{i,j})$ denotes the inventory of j -type vehicles in i^{th} bus line.

Suppose that the distance from the start station to the end station is 1. And it involves M station, whose station number is 1, 2, 3, M . M is an integer constant.

T is initially assigned to T_1 .

The core process of the algorithm is described as follows:

```

Bus_Adjust()
{if  $\sum_{i=1}^S g(A_{i(s)}) > N * \bar{C}$ 
then  $a(T)$ 

```

```

if  $g(A_{i(s)}) > \bar{C}$ 
then { if  $n(A_i, Q) > [g(A_{i(s)}) / C_2]$ 
      then if  $n(A_i, K) > [g(A_{i(s)}) / C_1]$ 
            then  $s(A_i, K, [g(A_{i(s)}) / C_1])$ 
            else  $s(A_i, Q, [g(A_{i(s)}) / C_2])$ 
      else if  $n(A_i, Q) > 0$ 
            then {  $s(A_i, Q)$ 
                  if  $(g(A_{i(s)}) - n(A_i, Q) * C_2) > C_2$ 
                  then if  $n(A_i, K) > (g(A_{i(s)}) - n(A_i, Q) * C_2)$ 
                        then  $s(A_i, K, [g(A_{i(s)}) / C_1])$ 
                        else if  $n(A_i, K) > 0$ 
                              then  $s(A_i, K)$ 
                  }
      }
}

```

4 Model Implementation and Simulation Experiments

NetLogo application platform is used in our experiment. NetLogo is a programmable modeling environment used for simulating the natural and social phenomena[6]. The model building process is divided into initialization phase and operation phase.

For this simulation system, we take a city bus line 3 as a simulation object. We assume that the interaction between the passenger agent and the intelligent station agent is correct operation. That is to say, the message that the vehicle scheduling agent has received from the intelligent station agent, is real and reliable. The parameters in the experiment are set as follows:

- 1) $l = 10\text{km}$, the distance from the start station to the terminal station;
- 2) $S = 15$, the total number of stations in our simulation system;
- 3) $L = 20$, the number of the signal light of no bus priority;
- 4) $\bar{C} = 50$, the number of passengers in a soaring station;
- 5) $V = 15\text{km/h}$, $\bar{V} = 20\text{km/h}$;
- 6) $T_1 = 10\text{min}$, $T_2 = 3\text{min}$;
- 7) $B_Q = 8$, the number of large buses whose capacity is 120 people; $B_K = 12$, the number of medium-sized buses whose capacity is 60 people.
- 8) $N = 3$;

The system runs 2 hours in the traffic peak period and some stations would appear the soaring phenomenon. The simulation results are recorded in Table 1.

Table 1. The experiment result of simulation of bus intelligent transit system

	the improved dynamic vehicle scheduling algorithm	the original algorithm
the capacity rate of large buses	97%	90%
the capacity rate of medium buses	98%	93%
the longest waiting time	12min	35min
the average waiting time	7min	9min

From the above experiment data, we find that using the improved dynamic vehicle scheduling algorithm, the average waiting time of passengers is slightly reduced and the longest waiting time of passengers is greatly reduced. At the same time, the capacities of both large and medium buses are improved.

5 Conclusion and Future Work

In this paper, agent-based simulation model of bus intelligent transit system is designed and implemented. In order to solve the soaring passengers' phenomenon in some stations, an improved algorithm that adjusts dynamically vehicle scheduling algorithm is proposed. Through the analysis of the experiment data, the algorithm is proved to be feasible and effective. Simulative experiment shows that, the model can not only truly simulate bus operation, but also reduce the waiting time of passengers, the operating costs of the bus company and the energy consumption. Due to our model under the assumption that passengers' arrival rate and the getting -off station number is random, it has a certain difference with the reality. In addition, with the rapid development of technology, the intelligent station can obtain not only the information of the passengers' getting-on station, but also the information of the passengers' getting-off station accurately. Therefore, how to simulate and optimized the bus scheduling more precisely, accurately and efficiently, is the further research.

References

1. Belmonte, M.-V., Pérez-de-la-Cruz, J.L., Francisco, T., Alberto, F.: Agent coordination for bus fleet management. In: The ACM Symposium on Applied Computing, vol. 1, pp. 462–466 (2005)
2. Wang, L., Chen, H., Li, Y., Shao, H.: A Survey:Application of Multi-Agent Technology to Urban Traffic System. *Computer Systems & Applications* 19(1), 198–203 (2010)
3. An, J., Yang, X., Teng, J., Liu, H.: Multi agent based Simulation of Transit Service Process. *Journal of Tongji University (Natural Science)* 38(6), 832–838 (2010)
4. Sienkiewicz, J., Hoyst, J.A.: Statistical analysis of 22 public transport networks in Poland. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 72(4), 1–11 (2005)
5. Zhang, Q., Yang, G.: Study on Agent Based Simulation of Bus Rapid Transit System. *Journal of Frontiers of Computer Science and Technology* 5(4), 363–367 (2011)
6. Wilensky, U.: NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1999), <http://ccl.northwestern.edu/netlogo>

Design and Simulation of Fuzzy Decoupling for Combustion Process of Circulating Fluidized Bed Boiler

Hao Tong, Xiaopeng Zhao, and Guozhuang Liang

College of Electrical Engineering and Information Science,
Hebei University of Science and Technology, 050018 Shijiazhuang, Hebei, China
tong.hao_001@163.com, xp.666666@163.com, liang_gz@sohu.com

Abstract. Decoupling is an important link in the control of combustion process of circulating fluidized bed boiler (CFBB). In this paper, the method of fuzzy is used in decoupling process, and the fuzzy decoupling device is designed based on the model of object with certain load. Then the control effect is compared with the simulation result which without decoupling device, and the contrast results showed that the method has a good decoupling capability. The method is easy to realize and simple to design, besides it has a good characteristics in adjusting the object model, it is valuable in the actual industry.

Keywords: Decoupling method, Fuzzy, Circulating Fluidized Bed Boiler.

1 Introduction

CFBB is an important power supply device in modern industries, and it is widely used in coal-fired power, heating and metallurgy industries. The control system of CFBB including drum water level control, main steam control and combustion control, and the combustion control is the most important and complicated control link [1]. With the characters of complex variables, long time lag and intense nonlinear, especially there exists a strong coupling relationship among the variables, and the coupling is a big trouble in the control of combustion process. In this paper the coupling characteristics of combustion process of CFBB has been analyzed, and then the fuzzy decoupling device is designed, it makes the controlled variables can be controlled alone and provides a convenience to the control of the whole system.

Lots of efforts have been put into the research of the decoupling. Jiao jian has made a brief analysis about the coupling of CFBB and some fuzzy principle involved in the reference [2]. Zhang GouZhong etc. use the forecast and adaptive control method in the model of CFBB, but the calculation of dynamic matrix is very complicated [3]. Fuzzy control method is used by Zhang ling in the temperature control system of distillation tower, and has got a good control effect [4]. At the same time, there exists a common problem. Most of the model involved in the references is simple and classical, besides the input is one step road, while the control model and signal is more complicated in actual practice, and the CFBB is a good example. The

fuzzy controller is designed in this paper for the complicated model and more practical input, and the control effect is satisfied.

2 Fuzzy Decoupling Principle

Decoupling control is one of the multivariable control methods. The so-called decoupling system is fix a certain structure, then looking for the right compensation devices to eliminate the coupling relationship between each variables, make sure the output is affected by the certain input and controlled by a certain controller.

2.1 The Basic Principle of Decoupling

The key work for decoupling system is to design decoupling devices. Figure 1 is the basic decoupling principle of double inputs and double outputs system.

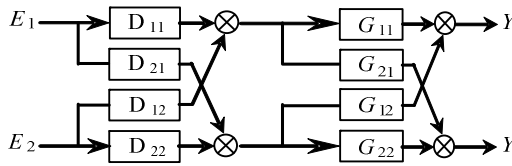


Fig. 1. The basic decoupling principle of double input and double output system

Where, G_{11}, G_{12}, G_{21} and G_{22} represent the model functions of controlled object, D_{11}, D_{12}, D_{21} and D_{22} are the matrixes of decoupling devices, the mathematical model can be got from the transmission relationships[4].

$$D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}^{-1} \begin{bmatrix} G_{11} & 0 \\ 0 & G_{22} \end{bmatrix} \tag{1}$$

Here we can get the decoupling matrix D, namely the mathematical models of the decoupling devices, the purpose of decoupling is achieved, and this is the general principle of decoupling. While in practice, the calculation of D will become more difficult and it has a deep dependence on the object model. As this, the methods with more simple and good control effect of decoupling are needed, and the static decoupling is one of them.

Based on the principle involved above, let $D_{11} = D_{22} = 1$, then we get the structure of static decoupling. The form (2) can be got by the compensation principle.

$$\begin{cases} D_{21} = -G_{21} / G_{22} \\ D_{12} = -G_{12} / G_{11} \end{cases} \tag{2}$$

It is the mathematical expression of static decoupling devices. But in this way, the adjusting range and adaptability is so low, that it can't meet the dynamic changes of the parameters in practice, so many experts introduce the fuzzy method to the decoupling.

2.2 The Principle of Fuzzy Decoupling

Based on the static decoupling, fuzzy decoupling bring the fuzzy decoupling devices to replace D_{21} and D_{12} of the static decoupling devices, to get the purpose of decoupling. The design of fuzzy decoupling devices contains three parts. the setting of quantification factors, the establishment of the fuzzy rules and the fixture of the scaling factor. The establishment of the fuzzy rules is very important, not only it needs to learn from the actual operation parameters and expert experiences, but also has to consider the nature of the devices themselves. Figure 2 is the principle of fuzzy decoupling.

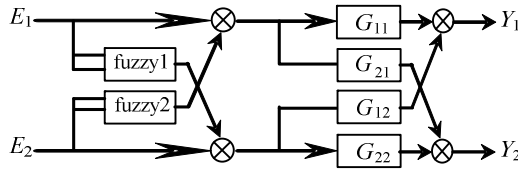


Fig. 2. The principle of fuzzy decoupling

3 The Design of Fuzzy Decoupling Device

According to the research of different references [1],[5], we choose the mathematical model as follow:

$$\begin{bmatrix} T \\ P \end{bmatrix} = \begin{bmatrix} \frac{1-b s}{1+T_o s} K_{\theta} e^{-\tau_{\theta} s} & -\frac{k_v}{1+T_v s} e^{-t_v s} \\ \frac{1-\alpha s}{(1+T_p s)^2} K_p e^{-\tau_p s} & \frac{k_u}{1+T_u s} e^{-t_u s} \end{bmatrix} \begin{bmatrix} m \\ f \end{bmatrix}. \quad (3)$$

Where m is the amount of fuel, f is the amount of air, P is the press of main air, T is the temperature of bed. As (3) express, there exists a different greed of coupling between the manipulation variables m , f and the controlled variables P , T .

Though the analyze of above, design fuzzy decoupling devices, the error between the setting value and measuring value get into the PID controller, take the output of PID controller and its differential as the input of fuzzy decoupling device, marked as E and EC . The control method of them is closed loop control, the principle of the fuzzy decoupling as figure 3.

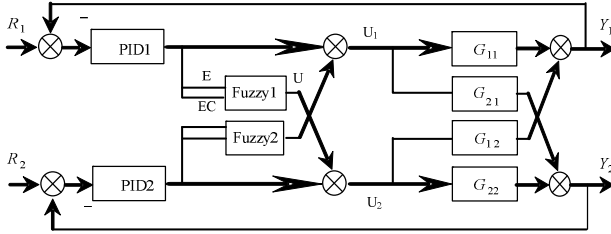


Fig. 3. The principle of the fuzzy decoupling system

Take the fuzzy controller 1 as example, the inputs are E (proportion link) and EC (differential link), and the output is marked as U. The fuzzy set is {NB,NS,ZO,PS,PB}. The membership function is trimf. The quantification factors and scaling factors can be calculated roughly according to the way that mentioned by Wang Chuanchuan[6], then turning in experiments, the quantification factors $K1=0.9, K2=0.5$, the scaling factors $K3=0.1$. The general form of the fuzzy decoupling rule is: IF E =A and EC=B, Then U=C, the basic rules of fuzzy controller 1 is shown in table 1.

Table 1. Rules of fuzzy controller 1

$\begin{matrix} U \\ E \\ EC \end{matrix}$	NB	NS	ZO	PS	PB
NB	PB	PB	PB	NB	NS
NS	PS	PB	PS	NS	ZO
ZO	PS	PB	ZO	ZO	ZO
PS	NS	ZO	NS	PS	PS
PB	NS	NB	NB	PB	PB

Table 2. Rules of fuzzy controller 2

$\begin{matrix} U \\ E \\ EC \end{matrix}$	NB	NS	ZO	PS	PB
NB	PS	PB	ZO	NB	NS
NS	PS	PB	ZO	NS	ZO
ZO	ZO	ZO	ZO	ZO	ZO
PS	ZO	ZO	ZO	PS	PS
PB	NS	NB	ZO	PB	PB

Through the analysis of above, the control parameters of fuzzy controller 2 can be got at the same time. Fuzzy set is {NB, NS, ZO, PS, PB}; the membership function is trimf; the quantification factors $K4=0.5, K5=0.5$, the scaling factors $K6=0.1$, the basic rules of fuzzy controller 2 is shown in table 2.

4 Simulation Results and Analyze

4.1 The System Simulation without Decoupling

Take the model of the CFBB in 70% lode as example, and simulate the system in MATLAB 7.0, the simulation diagram of system structure without decoupling is as figure 4.

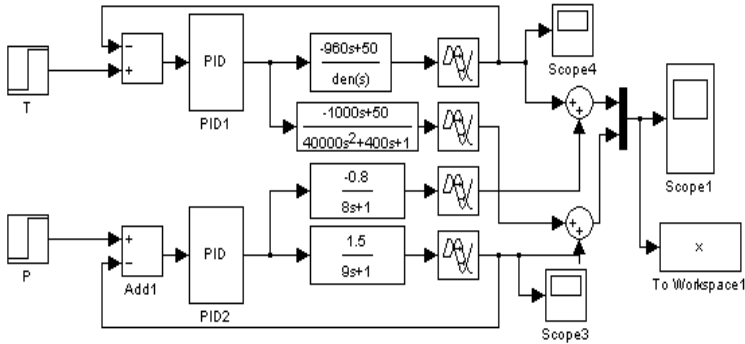


Fig. 4. The simulation diagram of system structure without decoupling

The parameter values of PID1 are $K_p=0.0568, K_i=0.000115, K_d=7$; The parameter values of PID2 are $K_p=0.000837, K_i=0.006, K_d=7$. The rendering of the stable output is as figure 5.

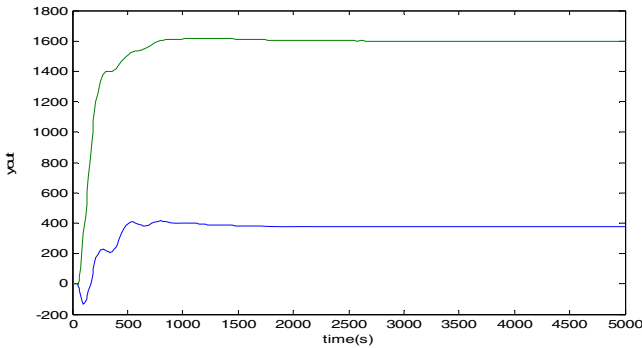


Fig. 5. The rendering without decoupling

Where $T=P=800$, if there no coupling between two input variables, then the stable simulation curve of the final value in the images should stable in 800, but the picture says the final value of bed temperature is about 400, and the final value of steam is 1600, the outcome of the experiment proved that there exists a strong coupling, if the system run as this can bring a consequence can't be imaged.

4.2 The Simulation Result of Fuzzy Decoupling

Design fuzzy decoupling device according to the design principle mentioned in part 3, the simulation rendering is as figure 6.

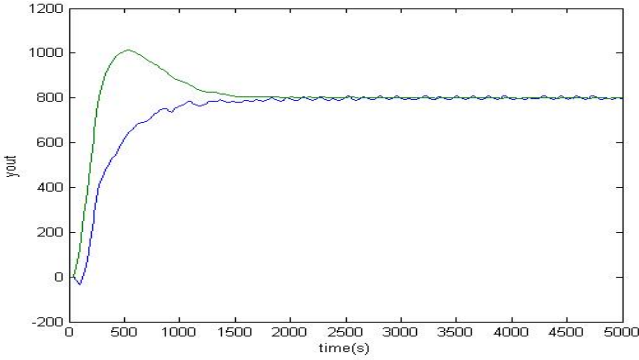


Fig. 6. The rendering of fuzzy decoupling

Obviously, the fuzzy controller has got the decoupling propose, the end value of control variables has been maintained in 800, and the peak value of steam was controlled at 1000, it provide a good safeguard to the industrial control. But there is still a problem, a wave is always appears in the final value of bed temperature. Proved by the experiment, the problem comes from the effect of the air which has an interference with the bed temperature, and the rendering curve of the interference is as figure 7.

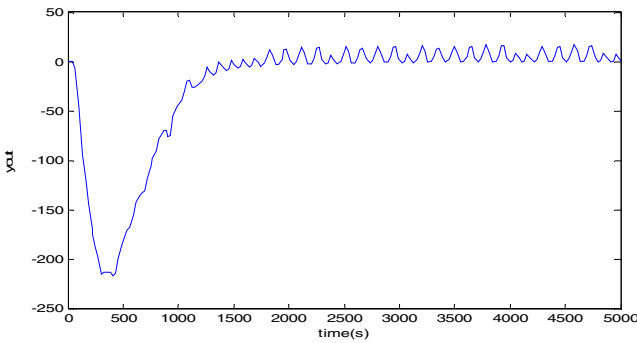


Fig. 7. The interference curve of the air

Although the final temperature of bed is controlled at the promised range of change, but the method still need to be optimized. The direction of the research will be the modification of the fuzzy rules and optimization of quantification factors and scaling factors.

5 Conclusion

In this paper, the fuzzy method is used in decoupling process, and the fuzzy decoupling device is designed for the combust process of CFBB, then the control effect is compared with the simulation result which without decoupling device. The contrast results showed that the method has a good decoupling capability, and can keep the control variables at the range of requirements. The method is easy to realize and simple to design, besides it has a good characteristics in adjusting the object model, it is valuable in the actual industry.

References

1. Zhao, X.-P., Liu, C.-Y., Song, X.-L., Yao, H.: Application and Simulation of Neural Nets in Combustion System of Circulating Fluidized Bed Boiler. In: Seventh International Conference on Natural Computation, vol. 11, pp. 79–82 (2011)
2. Jiao, J.: Decoupling Design of Combustion Control System for Circulating Fluidized Bed. *Northeast Power Technology* 5, 1–3 (2004)
3. Zhang, G.-Z., Wu, J., Liu, Y., Yang, Q.: Adaptive Control Research on Decoupled Model of CFBB. *Zhejiang Electric Power* 2, 1–4 (2008)
4. Zhang, L., Zhang, W.-Y., Zheng, E.-R.: Design and Simulation Research of a Sort Of Fuzzy Decoupling Control System. *Computer Simulation* 8, 118–121 (2010)
5. Niu, P.-F.: An Intelligence Combustion Process Control System Application and Research For Circulating Fluidized Bed Boilers of Giant Domestic Manufacture. *Proceedings of the CSEE* 12, 62–71 (2000)
6. Wang, C., Zhao, J., Qi, X.: The choice of the quantification factors and scaling factors in the design of fuzzy controller. *Journal of Sichuan Ordnance* 1, 61–63 (2009)

The Application of Ant-Colony Clustering Algorithm to Earthquake Prediction

Xiaoyan Shao, Xiang Li, Lingling Li, and Xinru Hu

Zhengzhou Institute of Aeronautical Industry Management
Zhengzhou 450015, P.R. China
xiaoyanshao@126.com

Abstract. Ant-colony clustering algorithm can well analyze the nonlinear relation between future earthquake magnitude, occurrence time and various anomalies before the earthquake, because it has stronger clustering function. This algorithm was applied to the Intelligence Decision Support System for earthquake prediction. The construction of ant-colony clustering algorithm in the system and its application in earthquake prediction was presented in this paper. Finally, the experiment showed that ant-colony clustering algorithm could achieve better results than the traditional k-means algorithm.

Keywords: ant-colony algorithm, clustering, earthquake prediction.

1 Introduction

Ant-colony clustering algorithm originated in the ant's instinct to construct its cemetery. Chretien made a lot of tests to study the ant's cemetery, who found out that worker ants could cluster dead bodies in few hours, no matter how they scattered randomly and how varied their sizes were. The basic principle was that the clustered dead bodies sent out pheromone to attract worker ants to put more similar objects, thus the smaller clusters became bigger. In that case, the distribution of clusters in the environment functioned as an indirect communication. Deneubourg and others proposed the ant-colony clustering algorithm (BM: Basic Model) to illustrate this phenomenon, in which it pointed out that single object was easier to be picked up and moved to other places where there were many similar ones. Based on the spatial distance between objects, Lumer and Faieta proposed relative algorithms, who popularized the basic model and applied it in data analysis[1].

2 Problem Description

In the ant-colony clustering algorithm, each data object was represented with an agent, which was corresponding to each ant in the colony; and the activity space for these agents was a two-dimensional grid, just as that of the ants in the colony.

The idea of the algorithm was that data object to be clustered was regarded as a classifying object in the ant nest, and each object was scattered randomly in the plane. Firstly, the ant calculated the similarity between the node and those within its viewing

radius, then it estimated the similarity between the node and those within the neighborhood, so as to make the decision whether it should pick up or drop the node. The ant was simulated by single agent and the self-organizing clustering of data objects was achieved by the interaction between each agent or each agent and the environment. [2]

The ant agent changed the similarity of the nodes within the neighborhood in the plane by moving data objects; and the results of different ant agents would exert influence on the environment around the neighborhood. An ant agent could improve the similarity of the nodes within the neighborhood by dropping a similar one and improve the similarity of dissimilar nodes within the neighborhood by moving a dissimilar one. Therefore, the ant agents communicated indirectly and exchanged information by means of influencing the similarity within the neighborhood [3, 4].

The function $f(i)$ referred to the average similarity between the object i that was picked up or dropped by the ant and the objects within the neighborhood of its viewing radius. The formula was shown as (1).

$$f(o_i) = \begin{cases} \frac{1}{\sigma^2} \sum_j \left(1 - \frac{\delta(i, j)}{\alpha}\right) & \text{if } f(o_i) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

In which, $\delta(i, j) \in [0, 1]$ represented the distance between Object i and Object j , which was usually calculated with Euclidean distance or the cosine of the included angle; $\alpha \in [0, 1]$ represented scale factor; σ^2 represented the area of Object i 's viewing field or the quantity of objects; the ant lied in the center of the neighborhood field, so $(\sigma - 1) / 2$ represented the viewing radius of the neighborhood field.

In the process of clustering analysis, the ant was always picking up the node that was the most dissimilar to those of the neighborhood, then dropped it at the place where were the nodes that were the most similar to those of the neighborhood. The probabilities of picking up and dropping referred to the probabilities that the ant made a decision to pick up or drop an object in accordance with the similarity of the neighborhood, which were expressed by P_{pick} and P_{drop} respectively.

$$P_{drop(i)} = \left(\frac{f(i)}{k^- + f(i)} \right)^2 \quad (2)$$

$$P_{pick(i)} = \left(\frac{k^+}{k^+ + f(i)} \right)^2 \quad (3)$$

The lower the average similarity between the data object and its neighborhood was, the more impossible the data object belonged to the neighborhood, so it was more possible to be picked up, and vice versa. In the formula, the influence of the similarity function $f(i)$ to P_{pick} and P_{drop} depended on the parameters k^+ and k^- [5, 6, 7].

3 Algorithm Description and Analysis

The basic idea of ant-colony clustering algorithm was elaborated as above, In this algorithm, $O(N)$ stood for the complexity of time and $O(N^2)$ stood for the complexity of space.

In addition, major measuring parameters during the ant-colony clustering included: spatial entropy, mean-fit and un-similar, which were used as the evaluation functions for clustering analysis. Particularly the first two parameters reflected the changing process of data clustering when the analysis was going on from different angles. Details as follows:

(1) Spatial Entropy

It reflected the clustering size in the sample space, whose formula was as(4).

$$E(\Theta) = - \sum_{K \in R} \frac{N_k}{N} * \log\left(\frac{N_k}{N}\right) \quad (4)$$

In which, R stood for data sample space, N_k referred to the number of nodes in the k th subclass and N referred to the total nodes in data sample space R . With the analysis of clustering, spatial entropy became smaller gradually, till it reached the minimum value.

(2) Mean-fit

It reflected the similarity between nodes in the whole sample space, whose formula was as (2.7).

$$F(\Theta) = - \frac{1}{N} \sum_{i=1}^N f(i) \quad (5)$$

In which, $f(i)$ was the similarity function for Node i , N referred to the number of nodes in the sample space. $F(\Theta)$ ranged from 0 to 1. With the analysis of clustering, mean-fit became greater gradually, till it reached the maximum value.

(3) Un-similar

It reflected the dissimilarity between nodes in the whole sample space, whose formula was as (6).

$$Unsim(\Theta) = \frac{1}{N} \sum_{i=1}^N \frac{1}{N_{nei}} \sum_j d(i, j) \quad (6)$$

In which, $d(i, j)$ referred to the distance between Node i and Node j in the neighborhood, whose value ranged from 0 to 1. With the analysis of clustering, the un-similar became smaller gradually, till it reached the minimum value.

(4) Outlier

Outliers referred to the data objects that were not in conformity with data's general behaviors or the data model, which were description for difference and particular cases, such as special cases beyond the standard, outliers other than data clusters, etc. Much potential knowledge was included in these special cases, so it was of particular

significance to explore and analyze them; for example, finding out fraud in regular commercial behaviors, intrusion detection in network security management and so on.

According to Knorr's definition: given a data set D and threshold values ξ and σ , if there were many sample points within q 's distance σ , then the sample point $q \in D$ was an outlier.

4 Application of the Algorithm in Earthquake Prediction

4.1 Data Background

In practical earthquake prediction, there was some relationship between precursors, time and types of seismological anomalies and the magnitude of future earthquake; as well as nonlinear relation between future earthquake magnitude, occurrence time and various anomalies before the earthquake. In this paper, the duration T (month) of 14 anomalies, which occurred relatively more in each earthquake, was selected as the input attributes[8]. The 14 anomalies were classified into two categories, i.e.

(1) Seismological anomalies: belt, gap, strain release, frequency, b-value, seismic window and wave velocity.

(2) Precursory anomalies: short-range leveling, ground tilt, ground resistivity, water radon, water level, strain and macroscopic precursor.

Since several-station anomalies might occur before one earthquake (i.e. the anomalies were observed at several stations), and an anomaly might occur at different stages during earthquake development. Therefore, when the precursory anomaly was observed at several stations, its weighted sum might be input:

$$T_i = \sum_j \omega_j t_{ij}, j = 1, 2, \dots, n \quad (7)$$

In which, t_{ij} was the duration of some precursory anomaly at some station. If some anomaly was observed at some station during the 3 stages of medium-term, short-term and immediate earthquake (it was regarded as 3 times in *EARTHQUAKE CASES IN CHINA*), then t_{ij} referred to the duration from the beginning of the anomaly to the occurrence of the earthquake; if the anomaly was observed during 2 stages, then t_{ij} referred to the sum of the 2 periods. ω_j was weight value, which was obtained by the following formula.

$$\omega_j = t_{ij} / \sum_j t_{ij} \quad (8)$$

In addition, the quantity of anomalies N should also be regarded as an input attribute in accordance with knowledge and experience of earthquake experts. N referred to the sum of all seismological and precursory anomalies prior to some earthquake, and their types included other anomalies besides the input attributes mentioned above. The same anomaly occurring at different stages shall be regarded as one item.

In the experiment, magnitudes were classified into three groups to verify the accuracy of clusters: the first group represented 5 earthquakes of M 7 or above; the second group represented 15 earthquakes of M6 to M 6.9; and the third group represented 25 earthquakes of M 5 to M 5.9.

4.2 Experimental Result

Table 1 showed all indicators for processing earthquake data in means of ant-colony clustering algorithm, comparing to k-means algorithm. In the experiment, parameters were as follows: $\alpha = 0.8$, $k+ = 0.8$, then $k- = 0.2$. The experiment was carried out 5 times and the data obtained were as follows.

Table 1. Comparison of Ant-colony Algorithm and K-means Algorithm for Processing Earthquake Data

Experiment	Average Distance		Time (s)	
	k-means	Ant-colony Clustering	k-means	Ant-colony Clustering
1	43.112	41.333	134	157
2	45.114	41.333	133	157
3	44.084	41.527	129	157
4	44.511	41.912	133	158
5	44.512	41.333	125	156
Average	44.267	41.488	131	157

It could be seen from the table that ant-colony clustering algorithm, on the prediction of earthquake data, moved shorter distance averagely than k-means algorithm, though it cost a little more time. It meant the former was more advantageous than the classical k-means algorithm and was very competitive in future application, so it was worthy of further study.

5 Conclusion

Clustering was actually a problem of integer linear programming. It was efficient to solve the problem of many clusters by ant-colony algorithm, particularly when the model sample and the quantity of clusters were large. Based on the theoretical analysis and experimental results in this paper, it had been found out that ant-colony clustering algorithm was a little inferior to the classical k-means clustering algorithm on calculating speed, but it moved shorter distance averagely. It was independent on the description of specific mathematic problems, and it could find the optimal solution more efficiently with its global optimization ability. Finally, it also proved the accuracy of the above conclusion when the ant-colony clustering algorithm was adopted for earthquake prediction in this paper. Therefore, the ant-colony clustering algorithm was of great value to research.

Acknowledgment. The research work is supported by Program for New Century Excellent Talents in University(2009), by Key Scientific and Technological Project of He'nan Province (No. 112102310082); by Key Scientific and Technological Project of He'nan Province (No. 112102210024), by the Foundation of He'nan Educational Committee(No.2011B520038).

References

1. Yan, Y., Fan, J., Kamel, M.: Clustering Combination Based on Ant Colony Algorithm. *Journal of the China Railway Society* 26(4) (August 2004)
2. Dorigo, M., Maniezzo, V., Colomi, A.: The Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics* 26(1), 28–41 (1996)
3. Parpinelli, R.S., Lopes, H.S., Freitas, A.A.: Data Mining With an Ant Colony Optimization Algorithm. *IEEE Transactions on Evolutionary Computing* 6(4) (2004)
4. Fan, M., Meng, X.: *Data Mining: Concepts and Techniques*. China Machine Press (2001)
5. Wu, Q., Zhang, J., Xu, X.: An Ant Colony Algorithm with Mutation Features. *Journal of Computer Research and Development* 36(10), 1240–1245 (1999)
6. Wu, B., Shi, Z.: Partition Algorithm for TSP Based on An Ant Colony Algorithm. *Chinese Journal of Computer* 24(12), 1328–1333 (2001)
7. Zhu, Q., Yang, Z.: An Ant Colony Optimization Algorithm Based on Mutation and Dynamic Pheromone Updating. *Journal of Software* 15(2) (2004)
8. Wang, W., Jiang, C., Zhang, J., Zhou, S., Wang, C.: The Application of BP Neural Network to Comprehensive Earthquake Prediction BP. *Earthquake* 19(2) (April 1999)

Rapid LED Optical Model Design by Midfield Simulation Verified by Normalized Cross Correlation Algorithm

YongJiang Di^{1,2}

¹ School of Metallurgical and Materials Engineering,
Chongqing University of Science and Technology, Chongqing 401331, China

² Department of Optics and Photonics, National Central University, Tao-Yuan 320, Taiwan
yjdee@163.com

Abstract. For the precise light behavior controlling in precise LED product design, the rapid and accurate LED model is need to build. The optical model of low and high power LEDs were built up and simulated by using Monte Carlo ray tracing, and were verified by the light intensity patterns in different midfield distances with the normalized cross correlation (NCC) algorithm. The one-dimensional intensity patterns were tested with a simple experimental set-up. The accuracy is higher than 99.55% in midfield and far-field region for NCC between the simulated light pattern and experimental measurement. The rapid optical model design assistant with NCC can be applied in varied precise LED illumination product design in midfield and far-field.

Keywords: Optical modeling, white LEDs, midfield region, cross correlation algorithm, light patterns.

1 Introduction

The LEDs (Light Emitting Diodes) are used everywhere, in many shapes, forms and powers, and with a wide range of application from indicator lights to solid-state lighting. The LEDs have more advantage than a traditional light source and have been regarded as the best potential light source for next-generation lighting [1, 2]. The accurate and rapid light patterns are needed for the forward design for indicator lights and illuminating lights [3, 4].

The traditional light patterns are obtained in the far-field region. But most case where the LEDs are not direct used as illuminating lights – they are used as component of light products and the light are manipulated with other components such as reflect cup and lenses in a distance short than the far-field distance [5, 6]. So the far-field light patterns may not agreement with the practical light patterns, and the final light patterns are not always accurate with that one want [7].

In this paper, the low and high power LEDs was simulated by Monte Carlo ray tracing in various midfield and far-field distances. The one-dimensional intensity patterns were measured with a simple set-up in midfield and far-field distances. The NCC between the simulation patterns and the experimental measurements are used to verify the validity of the second order light design.

2 Optical Modeling Process and the Experimental Measurement

The low power white LED is the most sold in electronic store with a diameter of 5 mm and power of 0.1W, called $\Phi 5$ LED. The prototype of high power white LED is PROLIGHT PR2N-3LVS-SD with power of 3W and approximately Lambertian light patterns. The establishing of LED geometric model was based on the above-mentioned LED and the configuration used in this simulation is shown in Fig. 1.

The simulation was carried out with commercial software. The original ray number is 10 million. The dependence of one dimensional light patterns of simulated on the distance is shown in Fig. 2. It can be found that the change of light patterns with the distance is very small for the high power white LED for the dimension of dome lens of high power LED is very close to that of the chip and having a Lambertian light patterns. As the distance is larger than 10 times of the LED dimension, the light pattern is on the quasi far-field for both the two LED.

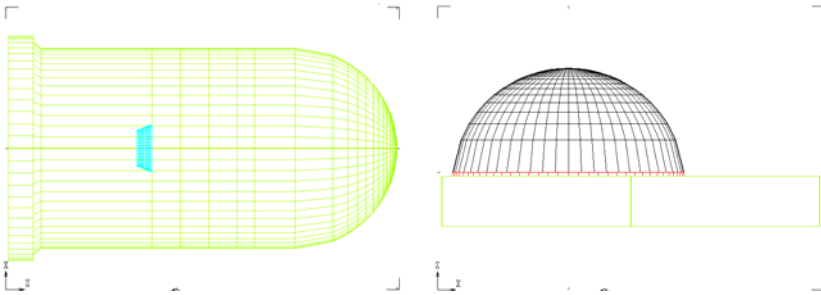


Fig. 1. Geometric model of LED: (a) low power LED, (b) high power LED.

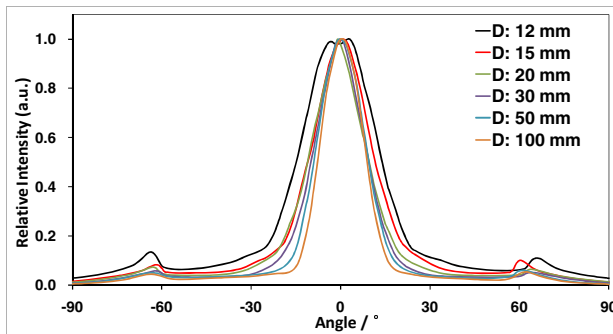


Fig. 2. Simulated angle patterns of different distance of the low power LED (left) and high power LED (right).

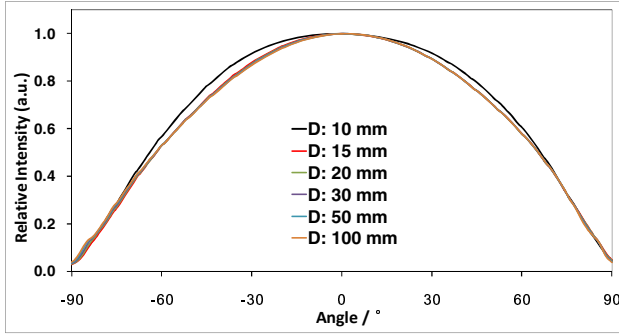


Fig. 2. (continued)

The experimental confirmation of the light patterns of the two type of LED is measured by one set up shown in Fig. 3 (a). The origin point of the distance starting is the dome vertex of the two LED. The end point of the distance is that of the sink point of the power meter. 3 LEDs of each type was used for the measurement of the light patterns for the examination of the LED light uniformity. The distance of both the two types of LED light patterns is 1cm to 100 cm. 181 mini intensity absorbing surfaces equivalent to that of the power meter of the experiment with 1° angular difference for each two neighboring detected surfaces are set in the simulation shown in Fig.3 (b).

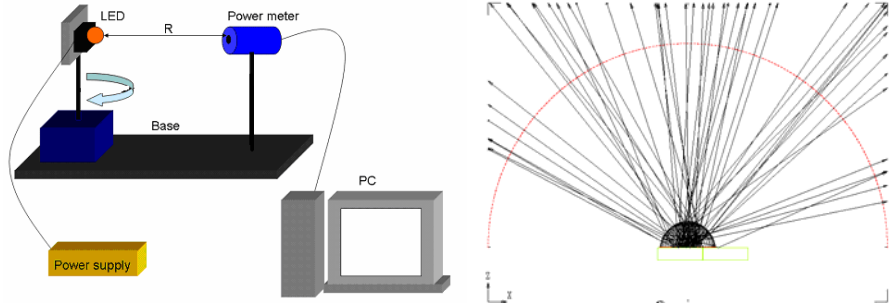


Fig. 3. Schematic of experimental set-up of the LEDs’ one-dimensional intensity patterns measurement (left), Diagram used in the simulation of the angular intensity distribution in simulation (right).

The similarity between the radiation patterns optical of the simulated and the measured was evaluated by the normalized cross correlation (NCC) algorithm [8]. NCC is frequently used to merit the similarity between two images. The NCC of the radiation pattern I_e at distance D with respect to the radiation pattern of reference I_s is

$$NCC = \frac{\sum_n [I(\theta_n)_e - \bar{I}_e][I(\theta_n)_s - \bar{I}_s]}{\sqrt{\sum_n [I(\theta_n)_e - \bar{I}_e]^2 \sum_n [I(\theta_n)_s - \bar{I}_s]^2}} \tag{1}$$

Where θ_n is the n -th view angle, e.g. the angle displacement of the detector shown in Fig. 2. \bar{I}_s and \bar{I} are the mean value of I_s and I across the angular range. Fig. 6 and Fig. 7 shows NCC vs. measurement distance for the radiation patterns of Fig. 4 and Fig. 5 for the low and power LED respectively.

3 Comparison of Simulation with Experiment Result

The NCC and the normalized radiation patterns between the simulation and the experiment were used to promote the modeling veracity. With the help of the NCC, the light model was improved to more close to that of the practicality. The light model is modified to cater to the normalized measured value at each view angles.

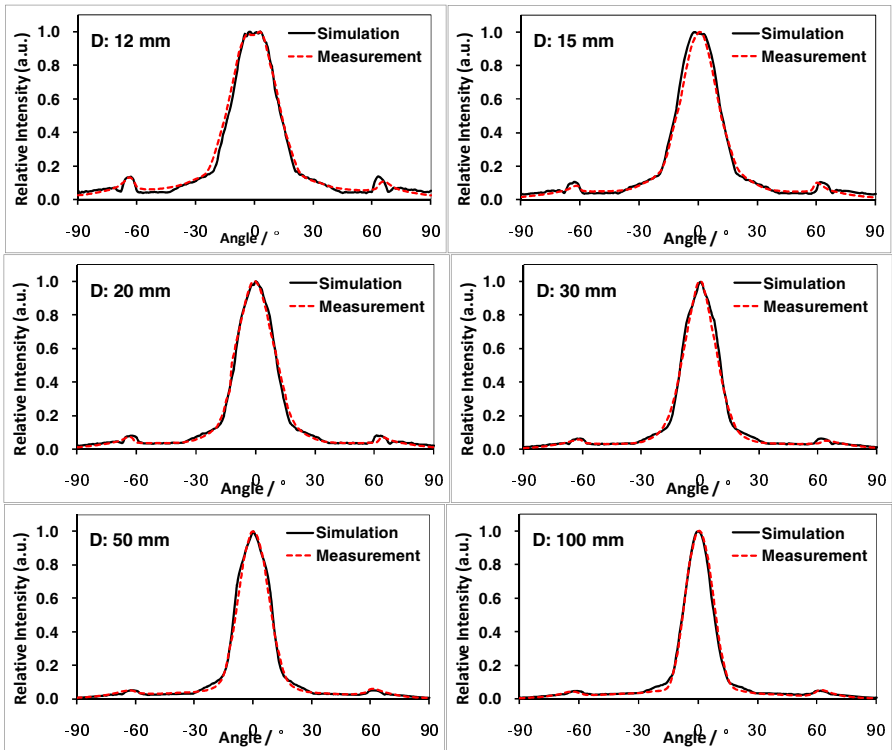


Fig. 4. Simulated light pattern versus experimental measurement for different distance of the low power LED.

One can see the simulated and measured radiation patterns of the two types LED in Fig.4 and Fig.5. The discrepancy between the simulation and experiment decreases firstly and then increase with the increasing of the distance. In the shortest distance in Fig.4 and 5, the light pattern significantly varies from other distance due to the finite

size of the LED structure for the LED can't be treated as a point light source. As the distance increases to remote midfield region, the light pattern has little change, as its quasi far-field is close to the midfield. For far distance, its behavior in the far-field direction region, the light pattern almost not change with the distance at varies angular direction [7].

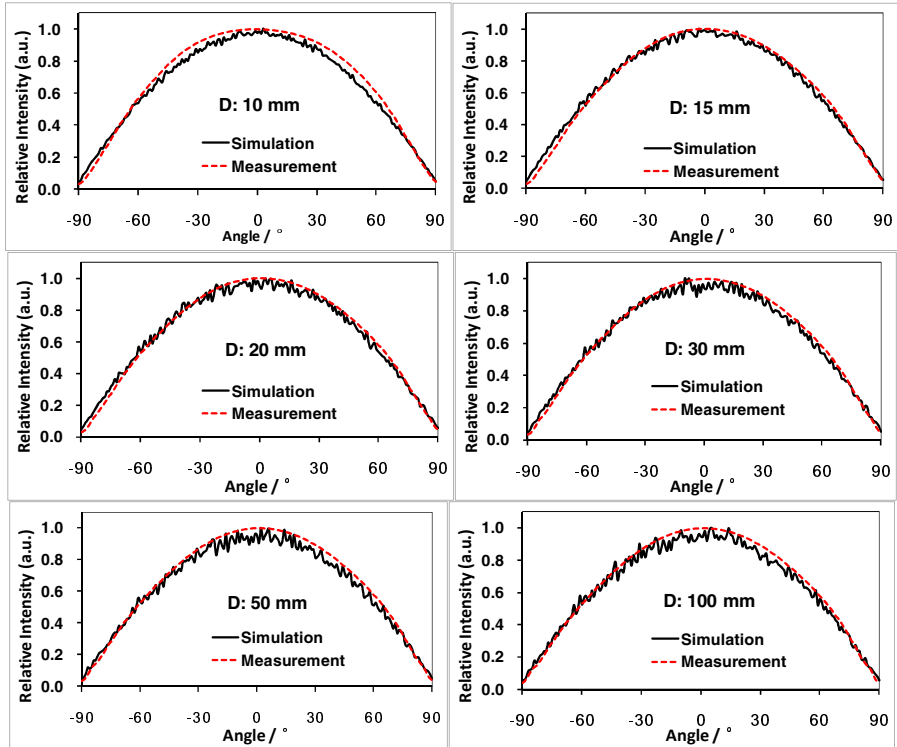


Fig. 5. Simulated light pattern versus experimental measurement for different distance of the high power LED

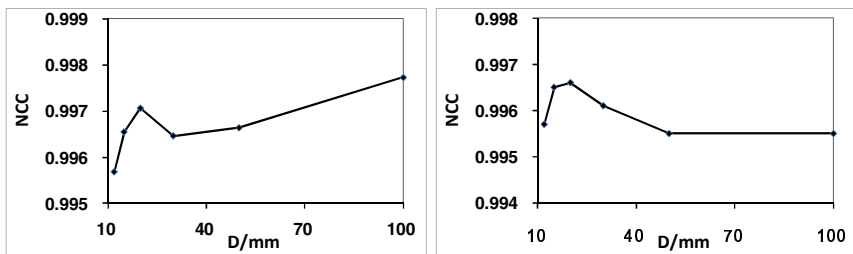


Fig. 6. NCC of the measurement and simulation of low and high power LED versus different distances.

The NCC of the two types LED is shown in Fig. 6. All the NCC data is larger than 0.9955 for the low and high power LED both. The NCC value is increase firstly as the distance is increasing, for the light pattern is more approaching to the far-field. The decrease in the mid distance is caused by the limit maximum ray number. Less ray reach to the detector for the larger distance.

The NCC of the two types LED is shown in Fig. 6. All the NCC data is larger than 0.9955 for the low and high power LED both. The NCC value is increase firstly as the distance is increasing, for the light pattern is more approaching to the far-field. The decrease in the mid distance is caused by the limit maximum ray number. Less ray reach to the detector for the larger distance.

4 Summary

One rapidly establishing of low and high power LED optical model is proposed by midfield simulation verified with normalized cross correlation algorithm. The light model can be improved with the NCC between the simulation and the experiment. The accuracy is higher than 99.55% in midfield and far-field region for NCC between the simulated light pattern and experimental measurement. The NCC among the measured LED measured in the midfield region can be used to check the lighting uniformity of the LEDs.

Acknowledgments. This research is funded by Research Foundation of Chongqing University of Science & Technology with project No. of CK2010B28, and the Cooperative Education Project of National Central University. The author acknowledgments the advising on the design and experiment given by Prof. Ching-Cherng SUN of Department of Optics and Photonics of National Central University.

References

1. Chen, Z., Zhang, Q., Wang, K., Luo, X., Liu, S.: Reliability test and failure analysis of high power LED packages. *Journal of Semiconductors* 32, 14001–14007 (2011)
2. Ramane, D., Shaligram, A.: Optimization of multi-element LED source for uniform illumination of plane surface. *Optics Express*. 848, 639–848 (2011)
3. Pan, J.W., Tu, S.H., Sun, W.S., Wang, C.M., Chang, J.Y.: Integration of Non-Lambertian LED and Reflective Optical Element as Efficient Street Lamp. *Optics Express*. 230, 22–230 (2010)
4. Teng, T.C., Kuo, M.F.: Highly precise optical model for simulating light guide plate using LED light source. *Optics Express*. 18, 22208–22214 (2010)
5. Chien, W.T., Sun, C.C., Moreno, I.: Precise optical model of multi-chip white LEDs. *Optics Express*. 15, 7572–7577 (2007)
6. Sun, C.C., Chen, C.Y., He, H.Y., Chen, C.C., Chien, W.T., Lee, T.X., Yang, T.H.: Precise optical modeling for silicate-based white LEDs. *Optics Express*. 16, 20060–20066 (2008)
7. Moreno, I., Sun, C.C., Ivanov, R.: A far-field condition for LED arrays. *Applied Optics* 48, 1190–1197 (2009)
8. Sun, C.C., Lee, T.X., Ma, S.H., Lee, Y.L., Huang, S.M.: Precise optical modeling for LED lighting verified by cross correlation in the midfield region. *Optics Letters* 31, 2193–2195 (2006)

Wireless Data Acquisition and Control System Based on DCS

Genghuang Yang, Abreham Fekibelu, Feifei Wang, Shigang Cui, and Li Zhao

Tianjin Key Laboratory of Information Sensing and Intelligent Control,
Tianjin University of Technology and Education, Tianjin, China, 300222
ygenghuang@126.com

Abstract. This paper presents a peripheral interface control (PIC)-based data acquisition and control tool that uses a Peripheral Interface Controller (PIC) microcontroller, data acquisition software's i.e. LABVIEW. Specifically, a library of PIC microcontroller functions for GUI is created. Moreover, the PIC microcontroller and LABVIEW are merged, by exploiting their serial communication capability, to produce an inexpensive data acquisition and control platform based on Distributed Control Systems, so that the user can observe, monitor and record the necessary datum. A custom serial communication protocol was designed and implemented. The hardware design of the system was optimized to reduce the amount of noise on power supply pins and analog inputs.

Keywords: data acquisition, measurement, wireless, distributed control system.

1 Introduction

Physical distribution of the sensors, controllers and actuators of fundamental control systems, and closing the loops through communication networks, form fundamental Distributed Control Systems (DCS)[1-2]. DCS has a wider scope and greater potential in terms of efficiency, flexibility, reliability and scalability for industrial applications. Combining smart sensor with wireless technology is a compelling addition to structure a large-scale smart sensor networks for real-time data acquisition, distributed detection and estimations. This requires a distributed data acquisition and control system containing edge-sensing devices to interact with different types of sensors, which poses significant challenges to the system architect.

Data acquisition is the processing by which physical phenomena from the real world are transformed into electrical signals that are measured and converted into a digital format for processing, analysis, and storage by a computer.

In a large majority of applications, the data acquisition (DAQ) system is designed not only to acquire data, but to act on it as well [3-4]. In defining DAQ systems, it is therefore useful to extend this definition to include the control aspects of the total system. Control is the processing by which digital control signals from the system hardware are converted to a signal format for use by control devices such as actuators and relays. These devices then control a system or process. Where a system is referred to as a data acquisition system or DAQ system, it is possible that it includes control functions as well.

Data acquisition hardware does not work without software, because the software running on the computer that transforms the system into a complete data acquisition, analysis, display, and control system.

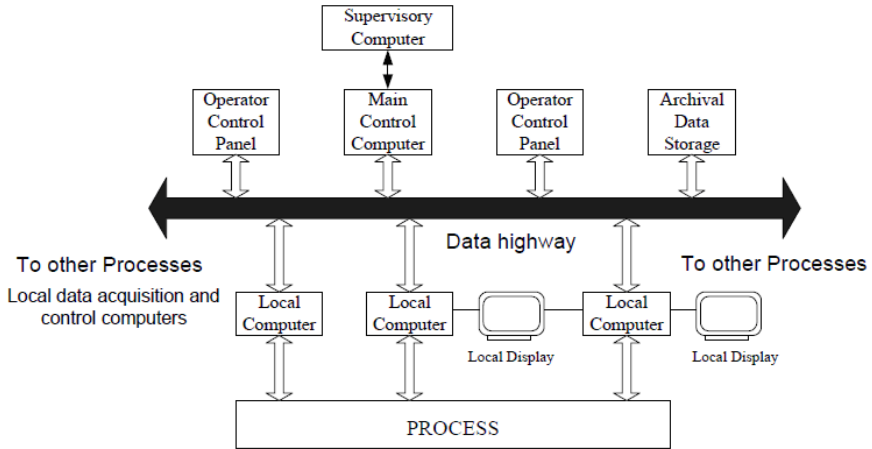


Fig. 1. Distributed control system architecture

The Operation Management subsystem provides a common interface for plant operators, engineers and maintenance personnel. The same HMI station presents each user class with critical information when and where it is needed. This includes real-time and historical process data, intuitive alarm indicators, sequence of events (SOE), and reports.

2 DCS Structure

A data acquisition and control system, built around the power and flexibility of the PC may consist of a wide variety of diverse hardware building blocks from different equipment manufacturers. It is the task of the system integrator to bring together these individual components into a complete working system. The DCS system consists of one or more of the following elements.

- (1) Sensors and transducers;
- (2) Field wiring;
- (3) Signal conditioning;
- (4) Data acquisition hardware;
- (5) PC (operating system);
- (6) Data acquisition software.

2.1 Transducers and Sensors

Transducers and sensors provide the actual interface between the real world and the data acquisition system by converting physical phenomena into electrical signals that the signal conditioning and/or data acquisition hardware can accept.

Transducers available can perform almost any physical measurement and provide a corresponding electrical output. For example, thermocouples, resistive temperature detectors (RTDs), thermistors, and IC sensors convert temperature into an analog signal, while flow meters produce digital pulse trains whose frequency depends on the speed of flow.

2.2 Field Wiring and Communicating Cabling

Field wiring represents the physical connection from the transducers and sensors to the signal conditioning hardware and/or data acquisition hardware. When the signal conditioning and/or data acquisition hardware is remotely located from the PC, then the field wiring provides the physical link between these hardware elements and the host computer. If this physical link is an RS-232 or RS-485 communications interface, then this component of the field wiring is often referred to as communications cabling.

Since field wiring and communications cabling often physically represents the largest component of the total system, it is most susceptible to the effects of external noise, especially in harsh industrial environments. The correct earthing and shielding of field wires and communications cabling is of paramount importance in reducing the effects of noise. This passive component of the data acquisition and control system is often overlooked as an important integral component, resulting in an otherwise reliable system becoming inaccurate or unreliable due to incorrect wiring techniques.

2.3 Signal Conditioning

Electrical signals generated by transducers often need to be converted to a form acceptable to the data acquisition hardware, particularly the A/D converter which converts the signal data to the required digital format. In addition, many transducers require some form of excitation or bridge completion for proper and accurate operation.

The principal tasks performed by signal conditioning are filtering, amplification, linearization, isolation and excitation.

2.4 Data Acquisition Hardware

Data acquisition and control (DAQ) hardware can be defined as that component of a complete data acquisition and control system, which performs any of the following functions:

- (1) The input, processing and conversion to digital format, using ADCs, of analog signal data measured from a system or process – the data is then transferred to a computer for display, storage and analysis

- (2) The input of digital signals, which contain information from a system or process
- (3) The processing, conversion to analog format, using DACs, of digital signals from the computer – the analog control signals are used for controlling a system or process
- (4) The output of digital control signals

2.5 Data Acquisition Software

Data acquisition hardware does not work without software, because it is the software running on the computer that transforms the system into a complete data acquisition, analysis, display, and control system. Application software runs on the computer under an operating system that may be single-tasking (like DOS) or multitasking (like Windows, UNIX, and OS2), allowing more than one application to run simultaneously. The application software can be a full screen interactive panel, a dedicated input/output control program, a data logger, a communications handler, or a combination of all of these.

3 DAQ and Control System Configuration

In many applications, and especially for data acquisition and process control, the power and flexibility of the PC, allows DAQ systems to be configured in a number of ways, each with its own distinct advantages. The key to the effective use of the PC is the careful matching of the specific requirements of a particular data acquisition application to the appropriate hardware and software available.

The choice of hardware, and the system configuration, is largely dictated by the environment in which the system will operate (e.g. an R&D laboratory, a manufacturing plant floor or a remote field location). The number of sensors and actuators required and their physical location in relation to the host computer, the type of signal conditioning required, and the harshness of the environment, are key factors. The most common system configurations are computer plug-in I/O, distributed I/O, stand-alone or distributed loggers and controllers and IEEE-488 instruments.

4 Modes of Computer Control

4.1 Computer Control Networks

The computer control network performs a wide variety of tasks: data acquisition, servicing of video display units in various laboratories and control rooms, data logging from analytical laboratories, control of plant processes or pilot plant, etc. The computer network can be as simple as an array of inexpensive PC's or it could be a large commercial distributed control system (DCS).

4.2 Small Computer Network

In small processes such as laboratory prototype or pilot plants, the number of control loops is relatively small. An inexpensive and straightforward way to deal with the

systems is to configure a network of personal computers for data acquisition and control. The network consists of a main computer linked directly to the process in two-way channels. Other local computers are linked to the main computer and are also connected to the process through one-way or two-way links. Some of these local computers can be interconnected. Each of the local computers has a video display and a specific function. For example, some local computers are dedicated for data acquisition only, some for local control only and some other for both data acquisition and local control. The main computer could have a multiple displays. All computers operate with a multitasking operating system. They would be normally configured with local memory, local disk storage, and often have shared disk storage with a server.

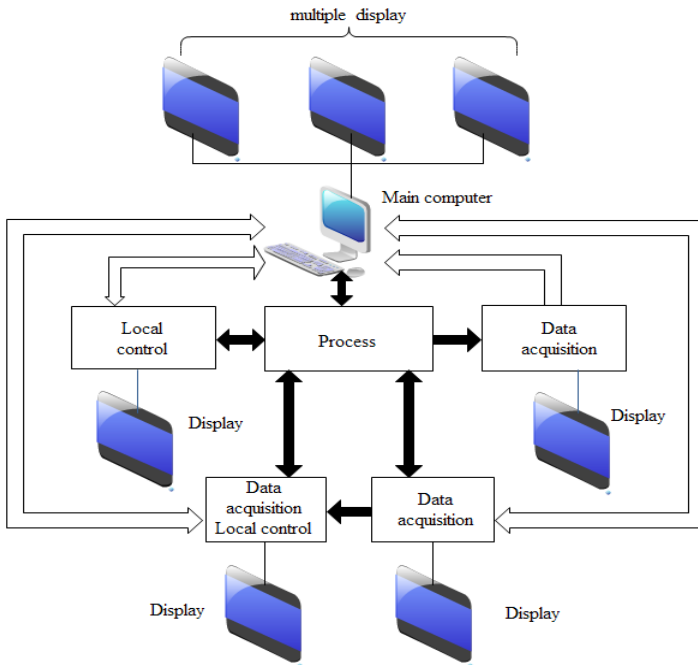


Fig. 2. Computer network

5 Application and Conclusion

A DCS is then a powerful tool for any large commercial plant. The engineer or operator can immediately utilize such a system to:

- (1) Access a large amount of current information from the data highway.
- (2) See trends of past process conditions by calling archival data storage.

(3) Readily install new on-line measurements together with local computers for data acquisition and then use the new data immediately for controlling all loops of the process.

(4) Alternate quickly among standard control strategies and readjust controller parameters in software.

Acknowledgement. This work is supported by China Postdoctoral Science Fund (20090460501), Tianjin Higher School Science & Technology Development Fund (20090704) and Fund of Tianjin University of Technology and Education (YJS10-03, YJS10-05, KJY11-08).

References

1. Korsaha, K., Wetheringtona, R., Wooda, R.: Emerging Technologies in Instrumentation and Controls: An Update Manuscript Completed (November 2005, January 2006)
2. Paraskevopoulos, P.N.: Modern Control Engineering. Marcel Dekker, Inc., New York (2002)
3. Korsaha, K., Wetheringtona, R., Wooda, R.: Emerging Technologies in Instrumentation and Controls: An Update; Manuscript Completed (November 2005, January 2006)
4. Gu, R.: Upgrading Process of Distributed Control System Based on TCP/IP. Industrial Control Computer (June 2007)

ARM and DSP Based Device for Power Quality Monitoring

Genghuang Yang, Feifei Wang, Shigang Cui, and Li Zhao

Tianjin Key Laboratory of Information Sensing and Intelligent Control,
Tianjin University of Technology and Education, Tianjin, China, 300222
ygenghuang@126.com

Abstract. This paper brings forward ARM and DSP structure for power quality monitoring in power station or factory. Different automatic production in industry is sensitive to different case of power quality. The new device is designed to be spot programmable to resolve the problem caused by the variation. DSP does data sampling and processing. ARM deals with the communication with high-level server by Internet. The device can work as a network terminal. Networking for power quality monitoring provides basic data for site-level and net-level assessment. FPGA simplifies the peripheral electro-circuit. The hardware schematic and software design are introduced especially the module related with voltage and current sampling, data exchange, interface for communication.

Keywords: Power quality, ARM, DSP, voltage variation.

1 Introduction

Nowadays the embedded system is more and more popular in information technology and industrial control. ARM and DSP are two main microprocessors in the application of embedded system. ARM runs with much higher speed than the primary microprocessor. ARM has been used in many fields in real-time application such as image manipulation, speech processing [1]. DSP has the advantage of RISC core in data processing especial numerous multiplies and additions.

DSP is the coprocessor in this device. DSP does data processing such as measurement of basic parameters and analysis of power quality [2]. The measurement includes power, power factor, amplitude value, phase of voltage and current. As for the power quality monitoring, there are several types. Different type decides different method to analyze. The type for analysis is activated by ARM. ARM7 works as the main core to manage the data transmission in the device especially the result from DSP after data processing. FPGA integrates the peripheral electro-circuit and establishes the credible control signals.

uClinux with TCP/IP is embedded in ARM7. The device can work as a terminal of networking. The device can transmit the data to high-level device or server for site-level power quality assessment [3]. Also the integration of data from each terminal in the network supplies information for net-level assessment.

2 Frame of the Device

SAMSUNG S3C4510B [4] incorporates the ARM7TDMI thumb processor with high-performance 32-bit RISC architecture. The chip integrates 10/100M Ethernet controller, 2 HDLCs, 1 I2C interface, 8kbyte CACHE/SRAM, 2 UART interfaces, 2 DMA controllers, 6 ROM/SRAM/FLASH interfaces, 4 DRAM/SDRAM interfaces and 4 I/O banks. S3C4510B is popular in the application of networking.

TI TMS320F2812 [5] is a 32-bit CPU with 150MIPS. The chip integrates 128kword FLASH, 128kword ROM, 16kword SARAM, 1 SPI, 2 UARTs, 16 12-bit ADCs. TMS320F2812 is popular in the field of industrial control. Fig.1 shows the structure of hardware.

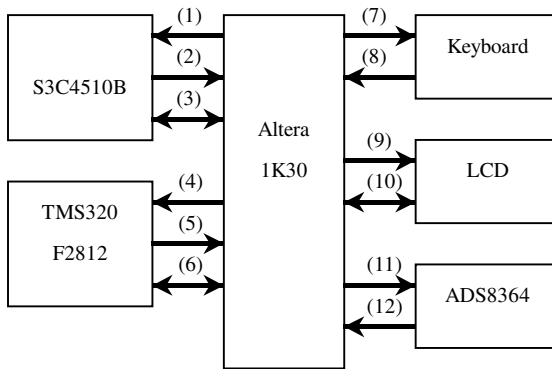


Fig. 1. Structure of hardware

(1)(4): interrupt signal 0 and nWait; (2)(5) (9) (11): controlling signals;(7): row bus for keyboard, R0-R3; (8): line bus for keyboard, L0-L3;(3)(6) (10) (12): data bus, D0-D7

S3C4510B controls the processing of TMS320F2812, LCD displayer and keyboard by 2 I/O banks. The main control signals include ECS2, ECS3, R/W signals nOE, nWBE0 and low 8-bit address bus for chip selection. Interrupt signal 0 is used for the handling of keyboard and nWait is the required waiting time for the low-speed modules. D0-D7 is the low 8-bit data bus.

ADS8364 [6] with 6 channel 16-bit AD converters is used for AC sampling. ALTERA ACEX1K30 [7] replaces the peripheral electro-circuit. As Fig.1 shows, ACEX1K30 simplifies the peripheral electro-circuit of S3C4510B to establish the control signals for other modules, line and row scan signals for the keyboard. Also the electro-circuit between TMS320F2812 and ADS8364 is replaced by FPGA to make the AC sampling more credible and precise. FPGA acts as the bridge for the data transmission between TMS320F2812 and S3C5410B.

There are 6 AC sampling interfaces, 32 DC sampling channel. Internet interface, USB1.1, RS-232 and RS-485 interface are also available.

2.1 Data Transmission between Two Cores

ARM-DSP module in FPGA accomplishes the data transmission between ARM and DSP. The signal /ARM-INT from ARM requests an interrupt handling of DSP. In the ARM write cycle, the 8-bit counter from 0 to 255 addresses the RAM and the data from ARM write to the RAM. Then in the DSP read cycle, DSP read the data from the addressed RAM. Thus the data transmits from ARM to DSP. Data transmission from DSP to ARM is the same as the above. ARM gets the data from DSP when necessary. If DSP checks the case of power quality, DSP will request ARM to get the sampling data or result of data processing in DSP.

2.2 Internet Interface

As for hardware of Internet application, the electro-circuit for Internet interface is made up of MAC controller and physical layer (HPY). The 10/100M Ethernet controller in ARM has media independent interface (MII) and buffered DMA interface (BDI). Internet MAC control is available in ARM and only a simple PHY interface chip is needed in the application. RTL8201 is adopted to be the MII interface. The isolated transformer separates the receiver and transmitter of signal from RJ45.

2.3 DSP- ADC Interface

AC sampling is the most important for the measurement in power quality monitoring. For the same ADC, the sampling frequency decides the precision. More samples in a cycle means more precise. As the frequency of AC waveform varies from time to time, the interval time to sample is the main factor for FFT algorithm. Phase-locked logic control electric-circuit gets the interval time for sampling based on the real-time frequency of AC waveform.

The waveform of AC voltage or current may be exceptional especially when frequency deviation occurs. In this case, the pulses from PLL module are not fit for sampling. Threshold of interval time to sample sets to handle the exception. In this application, when the frequency offset exceeds 5Hz (referenced frequency is 50Hz), the sampling frequency becomes fixed. ADS8364 has 6 channel 16-bit AD converters inside. To protect the ADC from over-voltage damage, dual amplitude limiters are used before the analog signals input to ADC.

3 Software Design of Hardware

3.1 Code Reload for DSP

Investigation from IEC shows that the probability when each type of power quality occurs varies with electromagnetic condition. Also different automatic production in industry is sensitive to different case of power quality. The power quality monitoring may be instant or continuous. It is unnecessary to add all type of power quality

monitoring in a device. The spot programmable function makes sense. The basic type of power quality monitoring such as voltage variation analysis, harmonic analysis and imbalance of three phases are included in all application. If needed for specific time and place, some types such as noise estimation, flicker can be added to the application.

In the application, there is a segment of code named code loader to load a new application. The segment of code for application and code loader are oriented in different page of FLASH EEPROM. The code loader includes data receiving, FLASH erasure and FLASH write processing. Once code reload order is received, the old application will interrupt and copy the segment of code loader to RAM. The code loader receives the segment of the new code for application from PC by communication such as RS-232 or USB. The code loader solidifies the segment to the address for the old application. After the code reload is finished, the new application starts up again.

3.2 uClinux in ARM

uClinux [8-9] is the mini operation system based on Linux. The uClinux for microprocessor project is a part of Linux to systems without a Memory Management Unit (MMU). uClinux is popular in application of ARM. In this application, uClinux is embedded in ARM and makes ARM run more credibly. The embedded OS manages the connected modules such as the keyboard, the display, the Internet interface or the other communication interfaces and the communication with DSP. TCP/IP is embedded in the network file system of uClinux.

4 Application of Power Quality Analysis

Voltage sag, swell and voltage interrupt are the common types in voltage variation analysis. As for the voltage sag, two parameters are included: one is magnitude of the voltage sag and the other is duration. Voltage swell and interrupt are the same as the sag. RMS computing gains the magnitude. The device detects and captures the voltage variation by comparing the RMS value to the threshold setting for voltage sag or swell. IEEE Standard 1159-95 [10] recommends the threshold value is 0.9 and 1.1 (of the nominal/rated voltage) for the sag and swell, respectively. If the voltage is lower than 0.1 (of the nominal/rated voltage), it means voltage interrupt. The distribution of voltage-duration in voltage variation cases is shown as Fig.2.

As the voltage variation is more popular than the other types, the device classifies the cases of voltage variation to gain the site-level power quality assessment based on SARFI-ITIC, SARFI-CBEMA and SARFI-SEMI. The display shows the result of assessment. When necessary for voltage variation analysis in net-level power quality assessment, the device transmits the data to high-level sever.

The data comes from a branch of Tianjin power supply Co.,Ltd. There are 124 cases of voltage variation from 2002 to 2004. 4 cases 3-phase power interrupt and 52 cases of power variation within 0.9 to 1.1 of rating voltage value. There are 26 cases of voltage variation in SARFI-ITIC.

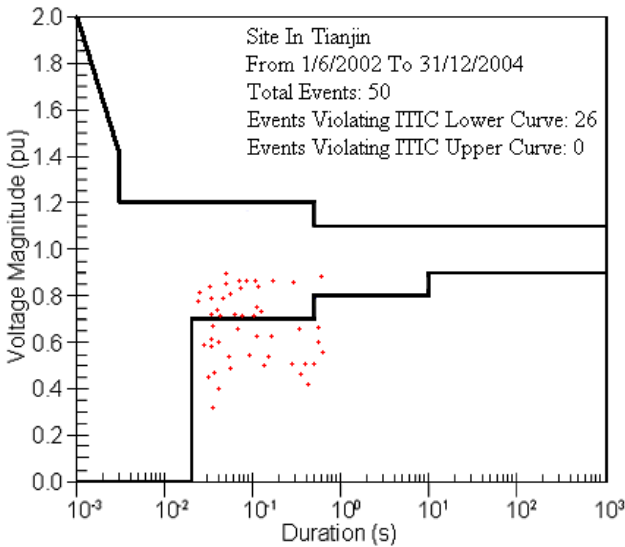


Fig. 2. Distribution of voltage-duration in voltage variation cases

5 Conclusions

(1) The task of ARM and DSP is clear, ARM acts as the core for communication and human-computer interaction and DSP runs for data processing especially realization of the algorithms for power quality analysis.

(2) File system of uClinux embedded in ARM includes the TCP/IP. The Internet interface makes the device works as a terminal of network. Networking of all the devices not only establishes the platform for power quality assessment, but also supplies the basic data for the future analysis.

(3) Spot programmable function makes the device work in a more flexible way. The initialization of device can be based on the special place or special purpose. Also the software in DSP can be upgraded. Besides, the application of power quality monitoring is made up of modules in software design, when needed, module for instant analysis can be added to the device.

(4) In hardware design, FPGA replaces much electric-circuit to establish the control signals. The design makes the device work more credible. The PLL module gets the synchronous frequency for sampling and reduces the influence of frequency variation on FFT.

Acknowledgement. This work is supported by China Postdoctoral Science Fund (20090460501), Tianjin Higher School Science & Technology Development Fund (20090704) and Fund of Tianjin University of Technology and Education (YJS10-03, YJS10-05, KJY11-08).

References

1. Cao, X.L., Wu, P., Ding, T.F.: A Design of a Communication Terminal Based on ARM and DSP. *Chinese Journal of Electron Devices* 28(2), 428–431 (2005)
2. Fan, J.T., Chen, J.Y., Wei, B.Q.: Microcomputer-based Feeder Protection on ARM & DSP Architecture. *Automation of Electric Power systems* 29(2), 77–80 (2005)
3. Baran, M.E., Kim, J., Hart, D.G., et al.: Voltage Variation Analysis for SiteLevel PQ Assessment. *IEEE Trans. on Power Delivery Application* 19(4), 1956–1961 (2004)
4. SAMSUNG Group, S3C4510B High Performance Network Controller,
<http://www.samsung.com/Products/Semiconductor/SystemLSI/Networks>
5. Texas Instruments Incorporated, TMS320F2812 - 32-Bit Digital Signal Controller with Flash,
<http://focus.ti.com/docs/prod/folders/print/tms320f2812.html>
6. Texas Instruments Incorporated, ADS8364-16-Bit 250 kSPS 6 ADCs, Parallel Out, W/6 x FIFO W/6 Ch.,
<http://focus.ti.com/docs/prod/folders/print/ads8364.html>
7. Altera International Limited, ACEX 1K Programmable Logic Device Family Data Sheet,
<http://focus.ti.com/docs/prod/folders/print/tms320f2812.html>
8. Arcturus Networks, CyberGuard and Analog Devices Inc sponosed, Embedded Linux/Microcontroller Project, <http://www.uclinux.org/index.html>
9. Ma, Z.M., Li, S.P., Kang, K., Ye, N.: Tutorial of Embedded System Based on ARM & Linux. ch. 7, pp. 236–252. Beijing University of aviation & avigation Press, Beijing (2005)
10. IEEE Recommended Practice for Monitoring Electric Power Quality. IEEE Standard 1159 (1995)

Chinese Ecological Risk Early-Warning Indicator System and Empirical Research

Zeng Yongquan

College of Sociology, Huazhong Normal University, Wuhan City,
Hubei Province P.R. China 430079

Abstract. In order to monitor the ecological risk, it is very necessary to establish the Chinese ecological risk early-warning indicator system (CEREWIS). Firstly, the article proposes the principle that should be followed in establishing CEREWIS and designs the basic framework of CEREWIS on the basis of expert consultation method. Secondly, the article determines the weights of each indicator through the further expert consultation and the analytic hierarchy process(AHP), makes dimensionless transformation through the maximum and minimum law, and constructs the comprehensive evaluation models. Finally, the article comprehensively evaluates the risk situations of the Chinese ecology in recent 10 years.

Keywords: Ecological risk, Early-warning, Index system, Comprehensive evaluation.

CLC number: C91 **Document code:** A

1 Introduction

Coordination development of human beings and their living environment is essential for the society's healthy operation and sustainable development. Ecological balance development results in the natural environment adaptive for human existence, and the natural environment is often the necessary conditions for the development of human society. Protecting the natural environment just likes controlling the population, is a China's basic national policy. China is confronted with the hard realities that the economy has not yet fully developed, but the natural and ecological environment has been substantially damaged in transition period, which means ecological risk has been serious risks the public faces. Strengthening early-warning of ecological risk, establishing ecological risk early-warning indicator system, as well as timely monitoring and disseminating emergency alarms, can provide scientific basis for government departments' pre-control, and correct decision making.

2 Principles for the Construction of Ecological Risk Early-Warning Indicator System

Scientificity. The construction of ecological risk early-warning indicator system must have a scientific basis in theory, in practice must be feasible and effective, so as to gather information and get the performance of the amounts, then make right inductive analysis and applications accordingly.

Purposiveness. When determining each individual indicator of ecological risks, should consider this indicator's status and role in the whole index system, according to its nature and characteristics which the particular object of study reflects, to determine the name, meaning and diameter range of the indicator,.

Affinity. Requires that various indicators of ecological risk early-warning indicator system must be interrelated and linked with each other in the aspects of meaning, diameter range, calculation method, calculation time and space range, etc., in order to understand the magnitude relations, internal relations and its regularity between ecological phenomena in an integrated and comprehensive manner.

Comprehensiveness. This calls for a high degree of generality of early-warning indicators, which can accurately reflect the production and development status of ecological risk. Thus requires combining static and dynamic indicators, objectively and comprehensively reflecting the conditions and laws of ecological risk's on-going development.

Maneuverability. Requires each indicator has precise value performance, with a full consideration of statistical work and the investigation of social status, to ensure that indicators' raw data can be collected.

Sensitivity. High levels of sensitivity are required for indicators, and an indicator value's subtle change can directly map out the development change status of a risk or a certain kind of risks. The choice of indicators focuses on simplicity, prevents to be well-rounded or too complicated, and highlights the functions of risk early-warning forecasting.

Comparability. The designation of indicators must refer to and be with the convergence of the statistical standards and norms domestically and abroad, to carry out international and domestic comparative analysis. Comparable with the historical data is also required, early-warning indicators should be relatively stable, which not only fully reflects the characteristics, conditions and needs of local social development at that time, but also possesses relative stability to maintain certain continuity for predicting future development recently.

3 The Structure of Ecological Risk Early-Warning Indicator System

This thesis adopts Delphi method to pre-select the indicators. A total of 12 consultants all have associate professor's professional title and above, who are engaged in sociology, social risks, and social problems for 10 years of research age and above. On this basis, this thesis establishes a set of ecological risk early-warning indicators system as follows:

Table 1. Ecological risk early-warning indicator system

Primary indicators	Secondary indicators	Unit
Natural disasters	Natural disasters mortality	1 / 10,000 people
	Economic losses rate of natural disasters	%
Resources	Unit GDP energy consumption	ton standard coal (TSC) /10, 000 yuan
	Cultivated area index	%
Environment	Forest coverage	%
	Urban air quality compliance rate	%
	Compliance rate of surface water	%
	Urban sewage treatment rate	%

In this indicator system, ecological risks include natural disasters, resources and environment these three primary indicators and eight secondary indicators.

4 Methods of Ecological Risk Early-Warning

In order to evaluate and monitor ecological risks currently in China., we must adopt scientific evaluation methods. Comprehensive evaluation method is a method being widely used currently. There is a need to determine the weights of each indicator and conduct non-dimensional transformation of indicator values and construct comprehensive evaluation models.

4.1 The Determination of Weights of Ecological Risk Early-Warning Indicators

This thesis adopted the analytic hierarchy process (AHP) and clustering methodology to empower every indicator, based on expert consultation. 11 experts participated in this expert consultation, respectively from universities and research institutions of Wuhan, Beijing, Shanghai and Chongqing, where 9 of them have positive senior professional titles.

Ecological risk primary indicators include natural disasters, resources and environment index. Through calculation, acquire the similarity coefficient of 11 experts; every similarity coefficient is very close to each other, so there is no need to remove the individual expert opinions, CR are all less than 0.1, indicating that the judgment matrixes of experts all pass the consistency. This thesis will use the method of weighted similarity coefficient to calculate the weights of each indicator, and normalize to get the final weights. The results are as follows:

Table 2. The determination of weights of ecological risk indicators

Experts	Similarity coefficients	Corresponding weights of experts judgment matrix			The largest eigen-value	CR
Expert 1	10.1795	0.3333	0.3333	0.3333	3.0000	0.0000
Expert 2	9.4825	0.3913	0.3043	0.3043	3.0000	0.0000
Expert 3	9.7984	0.2800	0.3600	0.3600	3.0000	0.0000
Expert 4	10.1795	0.3333	0.3333	0.3333	3.0000	0.0000
Expert 5	9.7975	0.3043	0.3043	0.3913	3.0000	0.0000
Expert 6	10.1795	0.3333	0.3333	0.3333	3.0000	0.0000
Expert 7	9.3715	0.3684	0.2632	0.3684	3.0000	0.0000
Expert 8	10.1795	0.3333	0.3333	0.3333	3.0000	0.0000
Expert 9	10.1795	0.3333	0.3333	0.3333	3.0000	0.0000
Expert 10	9.0069	0.2381	0.3333	0.4286	3.0000	0.0000
Expert 11	8.8637	0.2381	0.4286	0.3333	3.0000	0.0000
Final weights		0.3181	0.3323	0.3496		

Through the above study, we can draw the weights of ecological risk early-warning indicator system at all levels, as follows:

Table 3. The determination of weights of ecological risk early-warning indicator system

Indicators	Primary indicators	Secondary indicators
Ecological risk 1	Natural disasters 0.3181	Natural disasters mortality 0.5360
		Economic losses rate of natural disasters 0.4640
	Resources 0.3323	Unit GDP energy consumption 0.5323
		Cultivated area index 0.4677
	Environment 0.3496	Forest coverage 0.2335
		Urban air quality compliance rate 0.2515
		Compliance rate of surface water 0.2645
	Urban sewage treatment rate 0.2505	

4.2 The Non-dimensional Approach

After constructing the ecological risk early-warning indicator system, in order to achieve comprehensive assessments of ecological risk, the dimensions of each indicator must be unified. According to the meanings and data collection’s different difficulty degree of various indicators, this thesis will use three dimensionless approaches to treat every indicator data. (1) Fuzzy processing method, such as natural disasters mortality and economic losses rate of natural disasters; (2) Minimax method, such as unit GDP energy consumption and forest coverage; (3) Interval treatment method, such as cultivated area index, urban air quality compliance rate, compliance rate of surface water and urban sewage treatment rate.

4.3 Ecological Risk Assessment Model

The paper use multi-level linear weighted comprehensive assessment method to calculate its evaluation value required for ecological risk evaluation. The formula is

$$v_t = \sum_{i=1}^3 w_i \sum_{j=1}^m w_{ij}^{(1)} v_{jt}^{(2)}$$

Where, v_t denotes the comprehensive assessment value of ecological risk system in time period t , $v_{jt}^{(2)}$ represents the j secondary sub-index risk value in time period t , w_i represents the weight of the i primary sub-index corresponding to ecological risk system, $w_{ij}^{(1)}$ represents the weight of the j secondary sub-index corresponding to the i primary sub-system.

5 Empirical Study of Ecological Risk Comprehensive Assessment in China

The data collection is a key element of comprehensive evaluation. The data collected according to ecological risk early-warning indicator system is as follows:

Table 4. Data sheet of ecological risk early-warning indicators

Specific indicators	Units	ln	ln	ln	ln	ln	ln	ln	ln	ln	ln
		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Natural disasters mortality	1/100000	6.6	6.81	6.3	4.54	6.63	6.09	7.33	5.84	18.61	1.93
Economic losses rate of natural disasters	%	2.06	1.78	1.43	1.39	1	1.1	1.17	0.89	3.74	0.74
Unit GDP energy consumption	ton standard coal (TSC) /10 , 000 yuan	1.40	1.33	1.30	1.36	1.43	1.43	1.41	1.34	1.32	
Cultivated area index	%	100	99.51	98.20	96.22	95.48	95.20	94.96	94.93	94.91	
Forest coverage	%	16.55	16.55	16.55	18.21	18.21	18.21	18.21	18.21	18.21	20.36
Urban air quality compliance rate	%		34.4	34.5	41.5	41.4	65.8	62.4	69.8	76.9	82.4
Compliance rate of surface water	%		51.7	52.9	52.5	41.6	41	46	50.0	55	57.1
Urban sewage treatment rate	%				42.1	43.6	48.4	56	59.0	65.3	72.3

The data is mainly derived from “China Statistical Yearbook”, “Statistical Yearbook of Chinese Society”, “National Economic and Social Development Statistics Bulletin”, Chinese statistical yearbook database, the relevant years of national statistical database. Unit GDP energy consumptions are computed based on prices in 2000.

By calculating, the risk values from 2000 to 2009 are as follows:

Table 5. Natural ecological risk values from 2000 to 2009

Time	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Risk value	4.6685	4.5808	4.6952	5.0627	5.4809	5.2964	5.2555	4.9340	6.3232	4.2124

The results can also be depicted by the chart as follows:

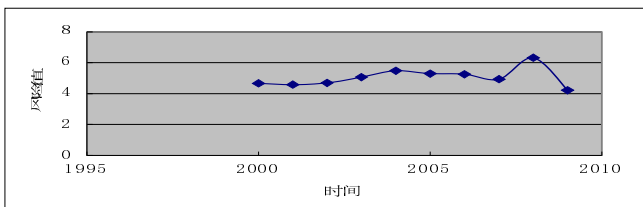


Fig. 1. Ecological risk values from 2000 to 2009

From the above chart we can see that the ecological risk of the past decade in China has been hovering in the 4-6 point intervals, reaches a peak in 2008 and declines in 2009. According to the risk warning level evaluation sheet, the warning level of ecological risk of the past decade has been the medium warning level, which means that social risks in ecological areas may occur at any time, in need of public policy-making sector’s analysis and review of relevant policy decisions and to take measures to reduce the risk losses to avoid other social risks caused by ecological risk.

References

1. Van der Voet, E., van Oers, L., Moll, S., et al.: Policy review on decoupling: Development of indicators to assess decoupling of economic development and environmental pressure in the EU-25 and AC-3 countries. CML report 166, Institute of Environmental Sciences, Leiden University, Department Industrial Ecology, Leiden (2005)
2. von Weizsacker, E., Lovins, A.B., Lovins, L.H.: Factor Four. Doubling Wealth-Halving Resource Use. Earthscan, London (1997)
3. OECD, Indicators to measure decoupling of environmental pressure from economic growth. OECD, Paris (2002)
4. World Bank. World Development Indicators, [EB/OL] (2010), <http://data.worldbank.org/data-catalog> (August 31, 2010)
5. Bertollo, P.: Assessing Landscape Health: A Case Study from Northeastern Italy. *Environmental Management* 27(3) (2001)
6. Lu, X.: The quantitative relationships between the economic growth and environmental burden. *Environmental Protection* 7, 13–18 (2007)
7. Wang, Y.: Tests and Analysis of Dynamic Relationships between China's Resource Consumption and Economic Growth. *Resources Science* 33(1) (2011)
8. Min, Q., Jiao, W., Cheng, S.: Pollution Footprint: A Type of Ecological Footprint Based on Ecosystem Services. *Resources Science* (2) (2011)
9. Zuo, W., et al.: Study on Regional Ecological Security Assessment Index and Standard. *Geography and Territorial Research* 18(1) (2002)
10. Guo, Z.: To Build the Early Warning and Maintaining System of National Ecological Security. *Science & Technology Review* 1 (2001)
11. Jiang, Y., Fan, Z.: The decision theory and method based on the judgment matrix. Science Publishing Company (2008)
12. Yue, C.: The decision theory and method. Science Publishing Company (2003)
13. Nu, X., et al.: The Chinese Social position analyzing and forecasting in 2010. Social Science Literature Publishing Company (2009)

The Optimization on the Multiperiod Mean Average Absolute Deviation Fuzzy Portfolio Selection

Zhang Peng and Pan Jin

School of Management Wuhan University of Science and Technology
Wuhan P.R. China, 430081

Abstract. Considering the transaction costs and the constraints on trade volumes, the paper proposes the multiperiod mean average absolute deviation fuzzy portfolio selection model which the fuzzy return rates are denoted as trapezoidal fuzzy numbers and the new algorithm—the discrete approximate iteration method is proposed to solve the model. The algorithm is as follows: According to the network method, discredits the state variables and transforms the model into multiperiod weighted digraph; uses max-plus algebra to solve the maximal path that is the admissible solution; continues iterating until the two admissible solutions based on the admissible solution are very close. At last, the paper proves the linear convergence of the algorithm.

Keywords: Multiperiod fuzzy portfolio selection, Mean Average absolute deviation, Discrete approximate iteration, Fuzzy number.

CLC number:F224.9 **Document code:**A

1 Introduction

Portfolio selection is an effective way to decentralize the investment risk. In the 1950s, Markowitz measured the investment risk by the variance and proposed the mean-variance formulation in a single period. This approach proposed by Markowitz paved the foundation of the modern portfolio theory [1]. However, portfolio selection inevitably exists fuzzy randomness. Because of the complexity of the stock market and many subjective factors, it is very reluctant to predict the future trend by the former return rates.

With the development of fuzzy mathematics, people began to study the portfolio selection under fuzzy uncertainty situation utilizing fuzzy set theory. In 1997, Watada [2] studied the portfolio selection on fuzzy return rates. From then on, numerous studies have focused on promoting the mean-variance approach proposed by Markowitz under fuzzy investment environment. For instance, Tanaka and Guo [3], Wang and Zhu [4], Zhangs [5,6], Bilbao-Teros [7], Lacagnina, Pecorella [8] and Huang [9] respectively proposed mean-variance approach under differently fuzzy investment environment; Fangs [10], Verchers [11], Guptas [12] proposed Mean Average absolute deviation approach under differently fuzzy investment environment; Huang [13-15] proposed fuzzy Mean -semivariance, Mean -risk approach;

Zhang[16] proposed mean - standard deviation fuzzy model;Lis[17-18] proposed Mean-variance - partial degrees, mean - entropy ,mean - cross entropy fuzzy model.Meanwhile, Some scholars study portfolio selection under random fuzzy mixed situation. For instance, Ammar[19], Hao[20], Hasuikes[21]. But the study above considered portfolio selection as single phase,namely,assuming investment returns and risks are changeless and investors haven't any investment strategy adjustments in the whole investment period. Obviously,this assumption is not conform to reality.

Considering the transaction costs and the constraints on trade volumes, the paper proposes the multiperiod mean average absolute deviation fuzzy portfolio selection model which the fuzzy return rates are denoted as trapezoidal fuzzy numbers .Then this paper use the new algorithm—the discrete approximate iteration method to solve the model.

2 Fuzzy Mean Average Absolute Deviation

Definition 1. Let A_i be the rate of return on the i th security, x_i be the investment proportion on i th security, $i = 1, \dots, n$, then $\omega(x)$ satisfies: x_i hypothesis fuzzy variables A_i for investment yields on the first I kind of assets, $I = 1, \dots, n$, x_i for investment first I kind of assets, the proportion of the investment portfolio of yields is

$$\omega(X) = \bar{M}(|\min\{0, \sum_{i=1}^n (A_i x_i - \bar{M}(A_i) x_i)\}|) = \bar{M}(\max\{0, \sum_{i=1}^n (\bar{M}(A_i) x_i - A_i x_i)\}) \quad (1)$$

where A_i is fuzzy variable, $\omega(x)$ is the mean absolute deviation of the fuzzy returns.

Definition 2. Assume A_i be trapezoidal fuzzy number, $A_i = (a_i, b_i, c_i, d_i)$,

For a given level γ , $[A_i]_\gamma = [a_i - (1 - \gamma)c_i, b_i + (1 - \gamma)d_i], \forall \gamma \in [0, 1]$. then the possibility mean of A_i satisfies:

$$\bar{M}(A_i) = \int_0^1 \gamma (a_i - (1 - \gamma)c_i + b_i + (1 - \gamma)d_i) d\gamma = \frac{a_i + b_i}{2} + \frac{d_i - c_i}{6} \quad (2)$$

According to (1), (2), we have the mean absolute deviation on investment portfolio as follows $\omega(x) = \frac{\sum_{i=1}^n b_i x_i - \sum_{i=1}^n a_i x_i}{4} + \frac{\sum_{i=1}^n c_i x_i + \sum_{i=1}^n d_i x_i}{12}$.

3 Fuzzy Portfolio on Multi-stage Mean Average Absolute Deviation

As is known to all, financial market are influenced by many improbability factors. in fuzzy uncertainly environment, risk assets' returns couldn't predict accurately. Based on the above reasons, assuming that there are various risk assets. Let r_{it} be the rate of return about the stage numbered t and risk asset numbered i , where r_{it} is denoted as trapezoidal fuzzy number, $r_{it} = [a_{it}, b_{it}, c_{it}, d_{it}]$, left width $c_{it} \geq 0$, right width $d_{it} \geq 0, t = 1, \dots, T, i = 1, \dots, n$; Let x_{it} be the investment proportion about the stage numbered t and risk asset numbered i, y_{it} and z_{it} respectively represent buy and sell quantity about the stage numbered t and risk asset numbered $i, x_{it} + y_{it} - z_{it}$ represent investment proportion about the stage numbered t and risk asset numbered

$i.C_i(y_{it})$ and $C_i(z_{it})$ respectively represent unity transaction costs of buy and sell about the stage numbered t and risk asset numbered i , S_0 and S_t respectively represent assets about the stage numbered 0 and numbered t . Let $M(r_{pt})$ represents fuzzy return rates, $f(r_{pt})$ represents mean average absolute deviation, then

$$M(r_{pt}) = \sum_{i=1}^n \left[\frac{(a_{it}+b_{it})}{2} + \frac{(d_{it}-c_{it})}{6} \right] (x_{it} + y_{it} - z_{it}) \tag{3}$$

$$f(r_{pt}) = \sum_{i=1}^n \left(\frac{b_{it}-a_{it}}{4} + \frac{c_{it}+d_{it}}{12} \right) (x_{it} + y_{it} - z_{it}) \tag{4}$$

If I_{pt} represents net fuzzy return rates, then

$$I_{pt} = M(r_{pt}) - \sum_{i=1}^n [c_i(y_{it}) + c_i(z_{it})] \tag{5}$$

Let $U(I_{pt}, f(r_{pt}))$ represents utility function about the stage numbered t , then

$$U(I_{pt}, f(r_{pt})) = \beta_t((1 - \omega)I_{pt} - \omega f(r_{pt})) \tag{6}$$

where $\beta_t(0 < \beta_t \leq 1)$ means discount factors, $\omega(0 < \omega \leq 1)$ means risk preference coefficient. $\beta_t = 1$ means a utility units is the same in each time, $\omega = 0$ means investors prefer the largest return rates regardless of risk, $\omega = 1$ means investors extreme dislike risk.

Financial market exists transaction costs and volume limit etc. Different investors in different circumstances has various limit. Its volume and transaction cost function is multiform. This paper proposes the multiperiod mean average absolute deviation fuzzy portfolio selection model as follows:

$$\begin{aligned} &max \sum_{t=1}^T \beta_t((1 + \omega)I_{pt} - \omega f(r_{pt})) \\ &s.t. \begin{cases} S_t = (1 + I_{pt})S_{t-1} \\ 0 \leq y_{it} \leq \sum_{j=1, j \neq i}^n x_{jt}, i = 1, \dots, n \\ 0 \leq z_{it} \leq x_{it} \\ y_{it}z_{it} = 0 \end{cases} \end{aligned} \tag{7}$$

In model (7), the first constraints means the state of final wealth transfer equation about the stage numbered $t - 1$ and numbered t ; the second constraints means that the buy quantity of the i th asset can't excess others' total quantity, i.e. don't allow bought, namely, don't allow borrow money to buy; the third constraints means don't allow shorting about the j th stock, i.e., don't allow borrow shares to sell; the fourth constraints means does not allow buy and sell the same assets at the same time. The economic implication of Model (7) means: To meet the four constraint conditions above, how to allocate various assets to achieve maximum utility for investors.

When there are transaction costs, the fourth constraints of Model (7) constant established [24].

4 Discrete Approximate Iteration Method

The basic steps of the discrete approximate iteration method [24] :

(1) Discretizing state variables into four equal parts by ascending order, i.e., forming five values;

(2) Using rotation algorithm can get the objective function values on different status values. Then constructing a multi-stage empowerment figure; ;

(3) Using algebraic methods to work out the shortest path of multi-stage empowerment digraph [26] ; If the difference between the $k+1$ path numbered $F^{(k+1)}$ and the k path numbered $F^{(k)}$ less or equal $\varepsilon(\varepsilon \leq 10 - 6)$, then stop the iteration ,i.e., the longest path is $F^{(k+1)}$; Otherwise continuing iteration based on the shortest path.

(4) Divided state variables the stage numbered $k+1$ and the minimum and maximum into equal two parts, then go to step 2.

Theorem. If U_i is concave function, then the discrete approximate iteration method satisfies linear convergence.

Proof. Let $U_1(0, j_1)$ be the longest edge of the first stage, $U_t(i_{t-1}, j_t)$ be the longest edge of the other stages, $t = 2, \dots, T - 1$, then $U_T(i_{T-1}, i_T)$ be the longest edge of the T stage, the upper bound of the optimal solution for model (7) is $U_1(0, j_1) + \dots + U_T(i_{T-1}, i_T)$.

By discretizing the state variable, we can construct multi-stage empowerment digraph. And through the great algebra, we can get the longest path from starting point to terminal point, which is the feasible solution of model (7). Based on the feasible solution, we can get another longest path through continuing iteration, then the longest path not less than the previous optimal path. Therefore, feasible solution of discrete approximate iteration method is monotonic increasing. Furthermore, as optimal solution of model (7) with bounded, so the discrete approximate iteration method is convergence.

Assuming the model's optimal value is U_0^* , then the optimal value of the stage numbered i is U_t^* ; the optimal value of the iteration numbered $k+1$ is U_0^{k+1} , the optimal value of the stage numbered $k+1$ is U_t^{k+1} ; the optimal value of the iteration numbered k is U_0^k , the optimal value of the stage numbered t is $U_t^{(k)}$. And $U_t^{(k)} \leq U_t^{(k+1)}$. Because U_t is concave function, so $0 \leq \frac{|U_t^{(k+1)} - U_t^*|}{|U_t^{(k)} - U_t^*|} \leq 1$, i.e., $U_t^{(k+1)} - U_t^* \leq U_t^{(k)} - U_t^*$. Then $\sum_{t=1}^T (U_t^{(k+1)} - U_t^*) \leq \sum_{t=1}^T (U_t^{(k)} - U_t^*)$, so $\frac{\sum_{t=1}^T |U_t^{(k+1)} - U_t^*|}{\sum_{t=1}^T |U_t^{(k)} - U_t^*|} \leq 1$, i.e., $\frac{|U_0^{(k)} - U_0^*|}{|U_0^{(k-1)} - U_0^*|} \leq 1$. So the discrete approximate iteration method satisfies linear convergence.

Demonstration finished.

5 Empirical Research

Example: In order to illustrate the proposed methods, let us consider the following thirty securities which chosen from Shanghai 30 stocks:

- $S_1(600000), S_2(600005), S_3(600015), S_4(600016), S_5(600019), S_6(600028), S_7(600030), S_8(600036), S_9(600048), S_{10}(600050), S_{11}(600104), S_{12}(600362), S_{13}(600519), S_{14}(600900), S_{15}(601088), S_{16}(601111), S_{17}(601166), S_{18}(601168), S_{19}(601318), S_{20}(601328), S_{21}(601390), S_{22}(601398), S_{23}(601600), S_{24}(601601), S_{25}(601628), S_{26}(601857), S_{27}(601919), S_{28}(601939), S_{29}(601988), S_{30}(601998)$.

We'll use the rate of return of the end of each quarter as the sample data from April 2006 to December 2010 (Data from the basic database of Guangdong development securities Limited company). Assuming $\beta = 1, \omega = 0.8$, let P_5, P_{40}, P_{60} and P_{95} to

represent 5% , 40% , 60% and 95% of rate of return , and $a_{it} = P_{40}, b_{it} = P_{60}, c_{it} = P_{40} - P_5, d_{it} = P_{95} - P_{60}$. So how to develop the optimal investment strategy about the five stages on the mean- absolute deviation portfolio selection? This paper adopts moving average method to estimate the next four stages of rate of return of thirty securities.

Solution: Using moving average method to estimate the next four stages of rate of return of thirty securities. Separately calculating five stages of each stock's a_{it}, b_{it}, c_{it} and d_{it} , results.

Using discrete approximate iterative method can work out the optimal five stages investment strategy as follows:

The first phase of the optimal investment strategy for: $a_{i3} = 1, a_{i1} = 0, i = 1, \dots, 30, \text{ and } i \neq 13$;

The second phase of the optimal investment strategy for: Without any trading;

The third phase of the optimal investment strategy for: Without any trading;

The fourth phase of the optimal investment strategy for: Without any trading;

The fifth phase of the optimal investment strategy for: Without any trading.

6 Conclusion and Expectation

Considering the transaction costs and the constraints on trade volumes, the paper proposes the multiperiod mean average absolute deviation fuzzy portfolio selection model which the fuzzy return rates are denoted as trapezoidal fuzzy numbers and the new algorithm—the discrete approximate iteration method is proposed to solve the model. At last, the paper proves the linear convergence of the algorithm and it puts forward a new idea for solving the multi-stage fuzzy portfolio. Through an empirical study shows that: the discrete approximate iteration method is a good solution to solve the "dimension disaster" problem of dynamic programming and its convergence is very good. It's also used for solving continuous dynamic programming problems, stochastic dynamic programming problems and sequential decision problems, which have great application prospect.

References

1. Markowitz, H.M.: Portfolioselection. Journal of Finance 7, 77–91 (1952)
2. Watada, J.: Fuzzy portfolio selection and its applications to decision making. Tatra Mountains Mathematical Publication 13, 219–248 (1997)
3. Tanaka, H., Guo, P.: Portfolio selection based on upper and lower exponential possibility distributions. European Journal of Operational research 114, 115–126 (1999)
4. Wang, S., Zhu, S.: On fuzzy portfolio selection problems. Fuzzy optimization and Decision Marking 1, 361–377 (2002)
5. Zhang, W.G., Wang, Y.L., Chen, Z.P., Nie, Z.K.: Possibilistic mean-variance models and efficient frontiers for portfolio selection problem. Information Science 177, 2787–2801 (2007)

6. Zhang, W.G., Xiao, W.L., Xu, W.J.: Possibilistic portfolio adjusting model with new added assets models and efficient frontiers for portfolio selection problem. *Economic Modelling* 27, 208–213 (2010)
7. Bilbao-Terol, A., Pérez-Gladish, B., Arenas-Parra, M., Rodríguez-Uriá, M.V.: Fuzzy compromise programming for portfolio selection. *Applied Mathematics and Computation* 173, 251–264 (2006)
8. Lacagnina, V., Pecorella, A.: A stochastic soft constraints fuzzy model for a portfolio selection problem. *Fuzzy Sets and Systems* 157, 1317–1327 (2006)
9. Huang, X.: Mean-variance model for fuzzy capital budgeting. *Computers & Industrial Engineering* 55, 34–47 (2008)
10. Fang, Y., Lai, K.K., Wang, S.: Portfolio rebalancing model with transaction costs based on fuzzy desicion theory. *European Journal of Oprational Research* 175, 879–893 (2006)
11. Vercher, E., Bermudez, J., Segura, J.: Fuzzy portfolio optimization under downside risk measures. *Fuzzy Sets and Systems* 158, 769–782 (2007)
12. Gupta, P., et al.: Asset portfolio optimization using fuzzy mathematical programming. *Information Science* 178, 1734–1755 (2008)
13. Huang, X.: Mean-Semivariance models for fuzzy portfolio selection. *Journal of Computational and Applied Mathematics* 217, 1–8 (2008)
14. Huang, X.: Portfolio selection with a new definition of risk. *European Journal of Operational Research* 186, 351–357 (2008)
15. Huang, X.: Risk Curve and Fuzzy Portfolio Selection. *Computers and Mathematics with Applications* 55, 1102–1112 (2008)
16. Zhang, W.G.: Possibilistic mean–standard deviation models to portfolio selection for bounded assets. *Applied Mathematics and Computation* 189, 1614–1623 (2007)
17. Li, X., et al.: Ahybrid intelligent algorithm for portfolio selection problem with fuzzy returns. *Journal of Computational and Applied Mathematics* 233, 264–278 (2009)
18. Li, X., Qin, Z., Kar, S.: Mean-variance-skewness model for portfolio selection with fuzzy returns. *European Journal of operational Research* 202, 239–247 (2010)
19. Ammar, E.E.: On solutions of fuzzy random multiobjective quadratic programming with applications in portfolio problem. *Information Science* 178, 468–484 (2008)
20. Hao, F., Liu, Y.: Mean-variance models for portfolio selection with fuzzy random returns. *Journal of Applied Mathematics Computation* 30, 9–38 (2009)
21. Hasuike, T., Katagiri, H., Ishii, H.: Portfolio selection problems with random fuzzy variable returns. *Fuzzy Sets and Systems* 160, 2579–2596 (2009)
22. Vercher, E., Bermúdez, J.D., Segura, J.V.: Fuzzy portfolio optimization under downside risk measures. *Fuzzy Sets and Systems* 158, 769–782 (2007)
23. Speranza, M.G.: Linear programming models for portfolio optimization. *Finance* 14, 107–123 (1993)
24. Zhang, P.: Discrete Approximate Iteration Method On the Mean-Absolute Deviation Multiperiod Portfolio Selection and the Empirical Research. *Journal of Systems & Management* 3, 266–271 (2010) (in Chinese)
25. Zhang, P., Zhang, Z.-Z., Yue, C.-Y.: Optimization of the Mean Semi - absolute Deviation Portfol io Selection Model with the Restricted Short Sell ing Based on the Pivoting Algorithm. *Chinese Journal of Management Science* 2, 7–11 (2006)

$[r, s, t]$ – Colouring of Join Graphs $S_n + O_m$

Mingzhong Mo

Department of Mathematics and Computer Science,
Liuzhou Teachers College, Liuzhou, Guangxi, 545004, P.R. China
momingzh@163.com

Abstract. The concept of $[r, s, t]$ -colourings was introduced by A. Kemnitz and M. Marangio in 2007 as follows: Let $G = (V(G), E(G))$ be a graph with vertex set $V(G)$ and $E(G)$. Given non-negative integers r, s and t , an $[r, s, t]$ -colouring of a graph $G = (V(G), E(G))$ is a mapping c from $V(G) \cup E(G)$ to the colour set $\{0, 1, 2, \dots, k-1\}$ such that $|c(v_i) - c(v_j)| \geq r$ for every two adjacent vertices v_i, v_j , $|c(e_i) - c(e_j)| \geq s$ for every two adjacent edges e_i, e_j , and $|c(v_i) - c(e_j)| \geq t$ for all pairs of incident vertices and edges, respectively. The $[r, s, t]$ -chromatic number $\chi_{r,s,t}(G)$ of G is defined to be the minimum k such that G admits an $[r, s, t]$ -colouring. In this paper, we determine the $[r, s, t]$ -chromatic number for join graphs $S_n + O_m$.

Keywords: Star, Empty graph, Join graphs, $[r, s, t]$ -colouring, $[r, s, t]$ -chromatic number.

1 Introduction

The $[r, s, t]$ -colouring of graph is a new coloring introduced in [1] by Kemnitz and Marangio.

Let $G = (V(G), E(G))$ be a graph with vertex set $V(G)$ and $E(G)$. Given non-negative integers r, s and t , an $[r, s, t]$ -colouring of a graph $G = (V(G), E(G))$ is a mapping c from $V(G) \cup E(G)$ to the colour set $\{0, 1, 2, \dots, k-1\}$ such that $|c(v_i) - c(v_j)| \geq r$ for every two adjacent vertices v_i, v_j , $|c(e_i) - c(e_j)| \geq s$ for every two adjacent edges e_i, e_j , and $|c(v_i) - c(e_j)| \geq t$ for all pairs of incident vertices and edges, respectively. The $[r, s, t]$ -chromatic number $\chi_{r,s,t}(G)$ of G is defined to be the minimum k such that G admits an $[r, s, t]$ -colouring.

It is a generalization of classical vertex coloring, edge coloring and total coloring of a graph. Obviously, a $[1, 0, 0]$ -coloring is an ordinary vertex coloring, a $[0, 1, 0]$ -coloring is an edge coloring, and a $[1, 1, 1]$ -coloring is a total coloring.

In [1], A. Kemnitz and M. Marangio gave the $[r, s, t]$ -chromatic number of general graphs of the boundary and some related properties, also discussed the $[r, s, t]$ -chromatic number of the complete graph of order n . Other results on $[r, s, t]$ -colouring

are presented in [2-4]. In our paper, we determine the $[r,s,t]$ -chromatic number for join graphs $S_n + O_m$.

2 Definitions and Preliminaries

The graphs we shall consider are finite, simple and undirected, unless stated otherwise, we follow the notations and terminologies in [5,6]. Let G be a graph. We denote its vertex set, edge set, minimum degree, maximum degree, order, size by $V(G), E(G), \delta(G), \Delta(G), p(G)$ and $q(G)$, respectively, and vertex chromatic number, edge chromatic number and total chromatic number by $\chi(G), \chi'(G)$ and $\chi_T(G)$ respectively.

A graph G is an empty graph if $p(G) = m (\neq 0)$ and $q(G) = 0$, that is, no two vertices are adjacent, is denoted by O_m . A star is a complete bipartite graph $K_{1,n}$, is denoted by S_n .

The join $G + H$ [7] of two disjoint graphs G and H is the graph having vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H) \cup \{xy \mid x \in V(G), y \in V(H)\}$.

We begin by stating some useful results from the literature. A. Kemnitz and M. Marangio [1] have proved the following properties for $[r, s, t]$ -coloring.

Lemma 1. If $H \subseteq G$ then

$$\chi_{r,s,t}(H) \leq \chi_{r,s,t}(G).$$

Lemma 2. If $r' \leq r, s' \leq s, t' \leq t$ then

$$\chi_{r',s',t'}(G) \leq \chi_{r,s,t}(G).$$

Lemma 3. For the $[r, s, t]$ -chromatic number of a graph G , there holds

$$\max\{r(\chi(G)-1)+1, s(\chi'(G)-1)+1, t+1\} \leq \chi_{r,s,t}(G) \leq r(\chi(G)-1)+s(\chi'(G)-1)+t+1.$$

Lemma 4[5]. Let G and H are simple graphs, then

$$\chi(G+H) = \chi(G) + \chi(H).$$

3 Main Results and Proofs

Let S_n be a star with order $n, V(S_n) = \{v_0, v_1, \dots, v_n\}$ and $E(S_n) = \{v_0v_i \mid i=1,2,\dots,n\}$, O_m be an empty graph with order $m, V(O_m) = \{v_1, v_2, \dots, v_m\}$. The join graphs $S_n + O_m$ is shown in Figure 1. We are so marked the join graphs $S_n + O_m$ throughout this paper.

Obviously, $\chi(S_n + O_m) = \chi(S_n) + \chi(O_m) = 2 + 1 = 3$ by lemma 4, $\Delta(S_n + O_m) = n + m$.

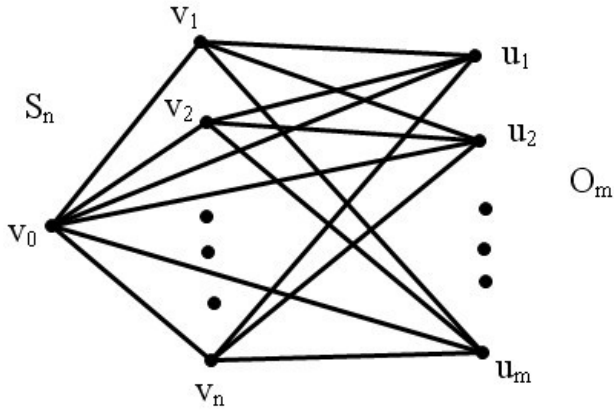


Fig. 1. Join graphs $S_n + O_m$

Next we determine the chromatic index for join graphs $S_n + O_m$.

Lemma 5. Let $S_n + O_m$ be a join graphs of star and empty, then

$$\chi'(S_n + O_m) = \Delta = n + m.$$

Proof. Let f be a coloring of join graphs $S_n + O_m$ with $n + m$ colours, According to n, m of the size, the proof is divided into two cases as follows:

Case 1. if $n \geq m$, coloring the edge of $S_n + O_m$ as follows:

$$f(v_0 v_i) = i, \quad i = 1, 2, \dots, n;$$

$$f(v_0 u_j) = n + j, \quad j = 1, 2, \dots, m;$$

$$f(v_i u_j) = \begin{cases} i + j & \text{if } i = 1, 2, \dots, n-1 \\ (n + j + 1) \bmod (n + m) & \text{if } i = n, \text{ modulus from } \{1, 2, \dots, n + m\} \end{cases},$$

$$j = 1, 2, \dots, m.$$

Case 2. if $n < m$, coloring the edge of $S_n + O_m$ as follows:

$$f(v_0 u_j) = j, \quad j = 1, 2, \dots, m;$$

$$f(v_0 v_i) = m + i, \quad i = 1, 2, \dots, n;$$

$$f(v_i u_j) = \begin{cases} i + j & \text{if } j = 1, 2, \dots, m-1, \quad i = 1, 2, \dots, n; \\ (m + i) \bmod m & \text{if } j = m \end{cases}$$

modulus from $\{1, 2, \dots, m\}$.

Since adjacent edges receive distinct colours, hence the edge coloring f is proper, then

$$\chi'(S_n + O_m) = \Delta = n + m.$$

Now we determine the $[r,s,t]$ -chromatic number for join graphs $S_n + O_m$.

Corollary 1. For join graphs $S_n + O_m$, it holds

$$\max\{2r + 1, s(\Delta - 1) + 1, t + 1\} \leq \chi_{r,s,t}(O_m + C_n) \leq 2r + s(\Delta - 1) + t + 1$$

Proof. By Lemma 3, Lemma 4 and Lemma 5, the conclusion is obvious.

Theorem 1. Let $S_n + O_m$ be a join graphs of star and empty, if $r \geq 2t + \left(\left\lceil \frac{\Delta}{2} \right\rceil - 1\right)s$ and $s \leq 2t$, then $\chi_{r,s,t}(S_n + O_m) = 2r + 1$.

Proof. Let c be a coloring of $S_n + O_m$ as follows:

Firstly, colour the vertices of $S_n + O_m$ with three colours $0, r, 2r$ as follows:

$$\begin{aligned} c(v_0) &= 0; \\ c(v_i) &= r, \quad i = 1, 2, \dots, n; \\ c(u_j) &= 2r, \quad j = 1, 2, \dots, m \end{aligned}$$

Secondly, colour the edges of $S_n + O_m$ with the $\Delta = n + m$ colours as follows:

In the interval $(0, r)$ is to select $t, t + s, \dots, t + \left(\left\lceil \frac{\Delta}{2} \right\rceil - 1\right)s$ this $\left\lceil \frac{\Delta}{2} \right\rceil$ colours, in the interval $(r, 2r)$ is to select $r + t, r + t + s, \dots, r + t + (\Delta - \left\lceil \frac{\Delta}{2} \right\rceil - 1)s$ this $\Delta - \left\lceil \frac{\Delta}{2} \right\rceil$ colours. Since $\chi'(S_n + O_m) = \Delta = n + m$, hence the edge colouring c is proper, so we obtain a proper $[r,s,t]$ -colouring of $S_n + O_m$, hence $\chi_{r,s,t}(S_n + O_m) \leq 2r + 1$. It holds $\chi_{r,s,t}(S_n + O_m) \geq 2r + 1$ by Corollary 1, therefore $\chi_{r,s,t}(S_n + O_m) = 2r + 1$.

Theorem 2. Let $S_n + O_m$ be a join graphs, if $r \geq 2t + (\Delta - 1)s$, then

$$\chi_{r,s,t}(S_n + O_m) = 2r + 1.$$

Proof. Let c be a coloring of $S_n + O_m$ as follows:

Firstly, colour the vertices of $S_n + O_m$ with three colours $0, r, 2r$ as follows:

$$\begin{aligned} c(v_0) &= 0; \\ c(v_i) &= r, \quad i = 1, 2, \dots, n; \\ c(u_j) &= 2r, \quad j = 1, 2, \dots, m \end{aligned}$$

Secondly, colour the edges of $S_n + O_m$ with the $\Delta = n + m$ colours $t, t + s, t + 2s, \dots, t + (\Delta - 1)s$.

Since $\chi'(S_n + O_m) = \Delta = n + m$, hence the edge colouring c is proper, so we obtain a proper $[r,s,t]$ -colouring of $S_n + O_m$, hence $\chi_{r,s,t}(S_n + O_m) \leq 2r + 1$. It holds $\chi_{r,s,t}(S_n + O_m) \geq 2r + 1$ by Corollary 1, therefore $\chi_{r,s,t}(S_n + O_m) = 2r + 1$.

Theorem 3. Let $S_n + O_m$ be a join graphs , if $s \geq \max\{r, 2t\}$, then

$$\chi_{r,s,t}(S_n + O_m) = (\Delta - 1)s + 1 .$$

Proof. Let c be a coloring of $S_n + O_m$ as follows:

Firstly , to colour the edges of $S_n + O_m$ with the $\chi'(S_n + O_m) = \Delta = n + m$ colours $0, s, 2s, \dots, (\chi'(O_m + C_n) - 1)s$, we can make c is proper edge colouring.

Secondly, to colour the vertices of $S_n + O_m$ with three colours $t, t + s, t + 2s$ as follows:

$$c(v_0) = t ;$$

$$c(v_i) = t + s, \quad i = 1, 2, \dots, n ;$$

$$c(u_j) = t + 2s, \quad j = 1, 2, \dots, m ,$$

Since $s \geq \max\{r, 2t\}$ we have $(t + s) - t = s \geq r$, $(t + 2s) - t = 2s \geq r$, $(t + 2s) - (t + s) = s \geq r$, $|ks - t| \geq t$, $|ks - (t + s)| \geq t$, $|ks - (t + 2s)| \geq t$, which $k = 0, 1, 2, \dots, \chi'(S_n + O_m)$, so c is a proper $[r, s, t]$ - colouring, hence $\chi_{r,s,t}(S_n + O_m) \leq (\Delta - 1)s + 1$, by Lemma 3 we have $\chi_{r,s,t}(S_n + O_m) \geq (\Delta - 1)s + 1$, then $\chi_{r,s,t}(S_n + O_m) = (\Delta - 1)s + 1$.

4 Summaries

In this paper, we first present the chromatic number and chromatic index of the join graph $S_n + O_m$, then we give the exact values of the $[r, s, t]$ - chromatic number of the join graphs $S_n + O_m$ if r, s, t meet certain conditions.

References

1. Kemnitz, A., Marangio, M.: $[r, s, t]$ - colorings of graphs. Discrete Mathematics 307, 199–207 (2007)
2. Dekar, L., Effantin, B., Kheddouci, H.: $[r, s, t]$ -colorings of trees and bipartite graphs. Discrete Mathematics 310, 260–269 (2010)
3. Kemnitz, A., Marangio, M., Mihók, P.: $[r, s, t]$ - Chromatic numbers and hereditary properties of graphs. Discrete Mathematics 307, 916–922 (2007)
4. Xu, C., Ma, X., Hua, S.: $[r, s, t]$ -coloring of $K_{n,n}$. J. Appl. Math. Comput. 31, 45–50 (2009)
5. Bondy, J.A., Murty, U.S.R.: Graph Theory. Springer, Heidelberg (2008)
6. Chartrand, G., Lesniak, L.: Graphs and digraphs. Greg Hubit Bookworks, California (1979)
7. Yap, H.P.: Total coloring of graphs. LNCS. Springer, Heidelberg (1996)

Study on Product Design for Seniors

Chunqing Yang

Eastern Liaoning University, Dandong Liaoning, China, 118000
ycqdeja@126.com

Abstract. The article analysis the problems of the product design for seniors in our country, such as the lack of elderly ergonomics research data, no concern about the changing needs and lifestyles of the elderly changes in product design, Lack of emotional and humanity design, and determine the direction of older design , toy design, assisted living products, care products and medical devices, fitness and travel goods design, and give the suggestions when design.

Keywords: product design, seniors, ergonomics.

1 Introduction

The global trend of an aging population, the market of elderly consumer market is increasingly large. According to National Bureau of Statistics predicts that the next 40 years, China will accelerate the aging of the forecast, table1 is the stream of the aging in china. Related to age and type of product consumption in the rapid increases in older industries is a growing potential market. Characteristics of research and development for the elderly, and the function of reasonable、 good quality products have become competitive in the market for development opportunities. From the ergonomic point of view, give full consideration to older people's physical characteristics, psychological characteristics, designed for the elderly to use the comfortable, convenient, safe, healthy products is the social development needs.

Table 1. The stream of the aging in china

Year	Total Population	>60years old		>65years old		>80years old	
		Population (billion)	Proportion %	Population billion	Proportion %	Population billion	Proportion %
2000	12.70	1.31	10.34	0.91	7.13	0.14	10.44
2010	13.76	1.73	12.54	1.15	8.38	0.21	12.23
2020	14.72	2.45	16.61	1.74	11.83	0.30	12.07
2030	15.24	3.55	23.30	2.44	15.96	0.43	12.07
2040	15.43	4.09	26.52	3.24	20.98	0.64	15.64
2050	15.21	4.38	28.76	3.32	21.81	1.00	22.91

1.1 The Lack of Elderly Ergonomics Research Data

The physical and psychological aspects of the elderly and young and middle-aged are very different, the product design need to get the whole social groups (including the

elderly and disabled) capacity-related aspects of the data, including: physiological aspects (such as: physical activities, strength, vision, hearing), psychological aspects (such as: perception, reaction time, memory), and needs to anthropometric data (human range of body and appearance). This research data can not blindly use foreign data, with localized features. Therefore, we should focus on the elderly ergonomic research, with the above information, the formation of a knowledge database, you can always extract the desired data, provide the basis for the design.

1.2 No Concern about the Changing Needs and Lifestyle of the Elderly Changes in Product Design

With technological development and social progress, the elderly, changes in consumer demand. If the product's features, usage behavior, usage and other aspects of the past, elderly people with different, and constantly changing. On the one hand with the middle-aged, young people are different; the other hand, the previous differences in older age groups also. At the present, the products are lack of a clear market positioning. Such as the market of cheap substitutes, some of the designers of "lazy design" is also equivalent to the old design,. Seniors product shape, color the design is too boring, monotonous, modeling language, blunt, can not give a sense of security, intimacy.

1.3 Lack of Emotional and Humanity Design

Product design should include the experience of the elderly, respect for cultural User backgrounds; friendly products and services used.

2 Design Direction of Seniors Products

2.1 Toy Design for Seniors, the Changes of Concept Is the Primary Problem

Seniors toy that is developed specifically for older toys, and some activities can help the elderly wrist, waist, and some have educational functions, for old hands-on brain, slow thinking, degradation, prevention of Alzheimer's disease.

Psychologists believe that staying at home for elderly people have a tendency to a return to innocence. The curiosity of the elderly with special emphasis on toys to satisfy their spiritual needs; some of the elderly living alone, if they can cultivate interest in favorite toys, flavoring agents will increase the number of life for people with mild dementia of the elderly, not only can improve the quality of life, but also promote health and longevity.

In Western countries developed for adults and older toys, toy market has become hot. In the United States, 40 percent of the toys are specifically designed for adults. Japan is also a lot of toys in the development of new features for the elderly, such as electric toys, playground common "combat crocodile," attach a blood pressure measurement function, as both entertainment and sports features, not only for stroke patients future nursing home can also be used. In this respect our country still in the blank.

2.2 Assisted Living Products

Physical characteristics and lifestyles of older people the correct analysis is to develop products in the market based on age. China more than 70% of the elderly are healthy, able to live independently, the demand for assisted living products is outstanding. Good living environment pleasant for the elderly can be physical and mental health, maximize their self-care period. Aging industry in more developed Western countries, the success of this design is very worth learning from. Manufactures really understands how to cast the like elderly: the French coffee with his grandmother, the United States, chewing gum dentures for the elderly, the elderly Japanese diaper production of urine bag. There are also older refrigerators use foot switches, push-button automatic latch, and so on.

2.3 Care Products and Medical Devices

Care products and medical devices for elderly and sick, the elderly living can not be completely self-designed products. Such as: wheelchairs, physiotherapy apparatus, the elderly with the diapers, blood circulation machine. Aging of the population caused by family size and family structure, family pension functions to continue to weaken, with the aging population, aging population is bound to the development of sick disability, life can not take care of the increasing proportion of elderly, with nursing and medical products can be part of the elderly to achieve self-care.

2.4 Fitness and Travel Goods Design

According to the U.S. Department of Health and Human Services issued a "guidance report on physical activity of citizens," said long-term regular physical activity participation can be effective in preventing cardiovascular disease, increase life and maintain good health. This is the first issued by the authority of the government's public information, on the positive role of sports fitness.

Scientific studies have shown that regular physical activity with cardiovascular training, strength training can reduce heart disease, myocardial infarction, hypertension, diabetes, colon cancer, depression and other risk of complications. Proven, the effect is particularly evident for the older age groups, for the prevention of Alzheimer's disease, improve quality of life of older persons has a very positive role. However, fitness center, sports rehabilitation, and even sporting goods manufacturers, few of the old market of specialized segments, nor in the products designed.

3 Suggestions

Design issues involved in the elderly can not but arouse our attention. Related to age and type of product consumption in the rapid increases in old age has become the industry's investment in the direction of the 21st century, is one of a growing potential market. Characteristics of research and development for the elderly, and the function of reasonable, good quality products have become competitive in the market for development opportunities. Give full consideration to the elderly physical

characteristics, psychological characteristics, designed for the elderly to use the comfort, convenience, safety, health product needs of social development, but also the social responsibility of the designer. Deal with the problems, given the following design recommendations.

3.1 Concern the Elderly Consumption Patterns and Behavior Research Found the Demand Rather Than creating Demand

Development of new products taking into account the age when the user is very important is to teach them how to be considered in practical design. Only when the designer has added to the older users, join a design and prototype evaluation of the experiment, they only fully realize the challenges of significance.

3.2 The Necessary Respect

The United Nations stressed the principle of independence of the elderly, participation, care, self-fulfillment and dignity. Because the elderly, the disabled, the "vulnerable" status, the community did not need to do their daily behavior research. Modern "accessible design", and some specifically for the elderly, the disabled, design appliances so that they most enjoy everyday life with the same treatment and normal power.

3.3 The Use of Modern Technology

With the development of technology, there will be more and more new technology into the lives of the elderly. While older people can not be sure how strong the demand for technology products, many products design services focus only on the mainstream consumer group of young, ignore the special needs of elderly consumers, thereby reducing the elderly desire for consumption, I am afraid that is an indisputable fact. Considerable number of enterprises in the use of high technology spare no effort in understanding and to facilitate the use of consumer products is very high-tech do not mind. Almost half of the elderly suffer from varying degrees of "technology phobia". How to use new technology to optimize product capabilities, simplified user interface in the elderly is an important topic of product design.

Acknowledgments. Liaoning Provincial Department of Education Science and Technology Research Project: Number: **2010058**.

Liaoning Education Science Eleventh Five-Year Plan 2010 research topic.

Comprehensive university where students study the training mode. **JG10DB040**.

References

1. Yang, C.: Mechanical and electrical product development and innovation, Beijing, pp. 57–58 (2007)
2. Yang, C.: Art and Design, Beijing, pp. 172–173
3. Ma, L.: On the old product design performance in the psychological study

A Time-Efficient FPGA-Based FFT Processor

Xudong Wang

College of Electronic and Information Engineer,
Nanjing University of Aeronautics & Astronautics, Nanjing 210016, China
xudong@nuaa.edu.cn

Abstract. A time-efficient 256 point fast Fourier transform (FFT) processor realized in programmable devices(FPGA) based on a novel architecture is presented. Simulation results indicate that the use of a 16x16 parallel structure for realizing 256-point FFT leads to a 16 times higher processing speed compared to its counterparts employing other series techniques.

Keywords: 256-point FFT, Time-efficient, FPGA.

1 Introduction

The fast Fourier transform (FFT) is one of the most popular algorithms in digital signal processing and it is used in communications, radar and reconnaissance applications. Field programmable gate arrays (FPGAs) have long been attractive for accelerating FFT processing speed[1]. FFT implementations on FPGA have been performed by using distributed arithmetic[2], complex multipliers[3], CORDIC algorithm[4], and global pipeline architecture[5]. However, there is no high speed FFT processor among them, which is critical for wideband system such as radar, SDR (software define radio) and reconnaissance[6].

This paper proposes a high speed 256-point FFT processor based on FPGA using a hybrid-parallel and pipeline architecture. The study has been particularised to decimation-in-frequency (DIF) FFTs of length 256 points and a 8-bit word-size has been considered. The whole processor has been implemented using two parallel 16-point FFT and 16 complex-multipliers between them. Its performance is found to be suitable with 2GHz streaming input data that is a very speediness effective option for wideband system.

2 Hybrid Parallel FFT Algorithm

To improve the system operation speed, a hybrid parallel FFT algorithm is used in this processor. Let us consider a DFT $x(n)$ of dimension N [7].

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{nk} \quad (1)$$

Where $W_N^{nk} = e^{-j2\pi kn/N}$ and $k=0\dots N-1$. If N is the product of two factors, with $N=N_1*N_2$, the indices n and k we can redefined as follows: $n=N_1*n_2+n_1$, where

$n_2=0\dots N_2-1$ and $n_1=0\dots N_1-1$, $k=N_2*k_1+k_2$, $k_2=0\dots N_2-1$ and $k_1=0\dots N_1-1$. Then we can split W_N^{nk} as follows

$$\begin{aligned} W_N^{nk} &= e^{-j\frac{2\pi kn}{N}} = e^{-j\frac{2\pi(k_1N_2+k_2)(n_2N_1+n_1)}{N_1N_2}} = e^{-j[2\pi(k_1+\frac{k_2}{N_2})(n_2+\frac{n_1}{N_1})]} \\ &= e^{-j[2\pi k_1n_2+\frac{2\pi k_2n_2}{N_2}+\frac{2\pi k_1n_1}{N_1}+\frac{2\pi k_2n_1}{N_1N_2}]} = W_{N_2}^{k_2n_2} W_{N_1}^{k_1n_1} W_N^{k_2n_1} \end{aligned} \quad (2)$$

Afterwards, we get:

$$X(k_1N_2+k_2) = \sum_{n_1=0}^{N_1-1} [(\sum_{n_2=0}^{N_2-1} x(n_2N_1+n_1)W_{N_2}^{n_2k_2}] W_N^{k_2n_1} W_{N_1}^{n_1k_1} \quad (3)$$

The N point DFT coefficients $X(k)$ can be calculated by N_1 channels of N_2 point FFT in parallel.

3 FPGA Implementation

The structure of the routines implemented into FPGA can be described as follow. Firstly, a DEMUX operation is used to de-series the input data into 16 parallel channels. Secondly, a global pipeline 16-point FFT is employed to these parallel data. Thirdly, the outputs of global pipeline FFT are multiplied with 16 complex coefficients from FPGA internal ROM. Lastly, another global pipeline 16-point FFT processor is employed to the outputs of the 16 complex multiplies. Afterwards, a MUX operation is used to make a serial output from FPGA. The 256-point FFT is calculated using 2 number of global pipeline 16-point FFTs. Synthesis and implementation of the design proved to take very high speed and consume less power when implemented on a xc4vsx55 device from the Xilinx family.

Table 1. Device utilization summary for the FFT processor

Source Type	Used	Available	Utilization
Number of Slice Flip Flops	10,717	49,152	22%
Number of 4 input LUTs	4,615	49,152	10%
Number of bonded IOBs	502	640	78%
Number of BUFG/BUFGCTRLs	4	32	12%
Number of FIFO16/RAMB16s	32	320	10%
Number of DSP48s	120	512	21%

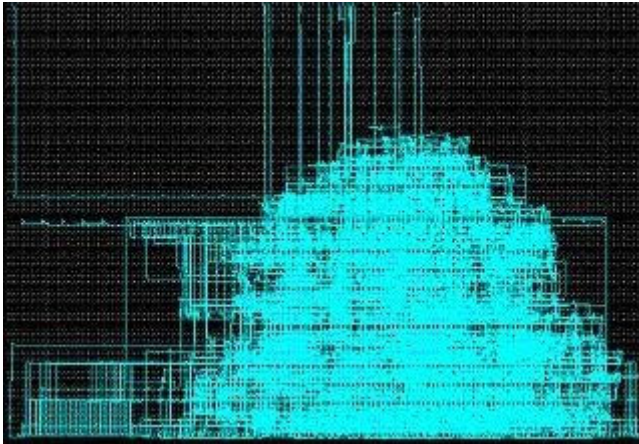


Fig. 1. Implementation results of this time-efficient FFT processor

4 Simulation Results

After the design entry of the respective blocks in the FFT processor, applying certain test patterns to verify the correctness of each block is functionally simulated. This level of simulation helps in testing the functionality of the design without including the gate level delays. Simulation of the design after implementation of the design is also done in order to obtain the delay associated with the FFT processor at the transistor level. This level of simulation takes more time compared to the functional simulation and gate level simulation. This level of simulation is known as timing simulation. Functional verification of the designed FFT processor, is done by simulating the top level RTL by applying various test patterns using a test bench. Simulation consumes 108 clock cycles to complete the 256-point FFT using the proposed hybrid parallel architecture after the data has been applied. The time taken

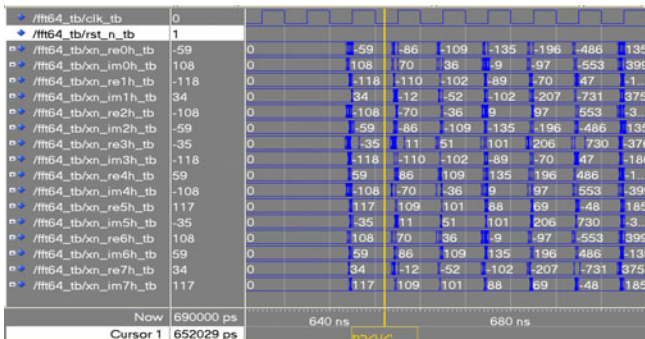


Fig. 2. Simulation results of the FFT processor

for the computation is 128ns(16 periods) at 125MHz internal clock frequency. Thus the input and output data rates can be as fast as 2GHz, which is 16 times of the calculation speed inner FPGA chip. Fig.2 shows the timing simulation result of this design with 3dB Signal-Noise-Ratio.

5 Conclusion the References Section

A 256-point high speed pipeline FFT processor for wideband system is designed, simulated, synthesized and implemented on FPGA. The whole design is done on VHDL. The architecture of this time-efficient FFT processor is given. VHDL design is on Synplify and simulation results are achieved on ModelSim and then synthesis is done on ISE to get the logic level. The overall resulted design has been implemented in a Xilinx xc4vsx55-12ff1148 FPGA. The approach presented here reduces the hardware complexity in FFT processor architectures and implements circuits more efficiently than global pipeline proposals. Main results indicate that the proposed time-efficient FFT processor performed with hybrid parallel structure gives up to 16 times operation frequency than the Xilinx cores. On the other hand, the area efficiency of the proposed 256-point FFT approach is 120% better than for the global pipeline FFT processor proposed in [5]. With our strategy, a time efficient FFT processor is implemented that can process data at a sample rate of up to 2G sample/s using patient resources.

Acknowledgments. This work is supported by China NSF Grants (60801052), Aeronautical Science Foundation of China (2009ZC52036), Ph.D. Programs foundation of China's ministry of education (200802871056) and Nanjing University of Aeronautics & Astronautics Research Funding (NS2010109, NS2010114). Thank Prof. Zhang Xiaofei for your help.

References

1. Chien, C.D., Lin, C.C., Yang, C.H., Guo, J.I.: Design and realization of a new hardware efficient IP core for the 1-D discrete Fourier transform. *IEE Proc. Circuits, Dev. Syst.* 152(3), 247–258 (2005)
2. Zhou, Y., Noras, J.M., Shepherd, S.J.: Novel design of multiplier-less FFT processors. *Signal Processing* 87(1), 1402–1407 (2007)
3. Cheng, L.S., Miri, A., Yeap, T.H.: Efficient FPGA implementation of FFT based Multipliers. In: *Proc. IEEE CCECE/CCGEI, Saskatoon*, pp. 1300–1303 (2005)
4. Banerjee, A., Dhar, A.S., Banerjee, S.: FPGA realization of a CORDIC based FFT processor for biomedical signal processing. *Microprocessor and Microsystems* 25(1), 131–142 (2001)
5. Szadkowski, Z.: 16-point Discrete Fourier transform based on the Radix-2 FFT algorithm implemented into Cyclone FPGA as the UHERC trigger for horizontal air showers. In: *Proc. of the 29th ICRC, Pune* (2005)
6. Sanchez, M., Garrido, M., Lopez-Vallejo, M., Grajal, J., Lopez-Barrio, C.: Digital channelised receivers on FPGA platforms. In: *Proc. IEEE International Radar Conference*, pp. 816–821 (2005)
7. Oppenheim, A.V., Schaffer, R.W., Buck, J.R.: *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs (1999)

The Research on the Impact of Project Communication to Project Performance

Zhongqing Wang¹ and Qiang Hu²

¹ School of Urban & Rural Construction, Chengdu University,
Chengdu, 610106, China
georgehq@163.com

² Teaching Affairs Department, Chengdu University, Chengdu, 610106, China

Abstract. The data of the relations between project communication and project performance can be collected by questionnaire. Then SPSS software was used to analysis the data. The research shows that project communication has the most significant impact on project schedule. And different project communication factors have the different impacts on project performance.

Keywords: Project communication, Project performance, Correlation analysis.

1 Introduction

Project communication is an organizational communication. Organizational communication has some features, which communication groups are not only the objects of the relationships, but also the targets of the organizational tasks. In addition to have characteristics of organizational communication above mentioned, project communication is flexible and dynamic, that make it have more difficulties and more impacts than organization communication. Just like Chester Barnard strongly asserted 60 years ago in his famous book *The Function of the Executive*, “In the exhaustive theory of organization, communication would occupy a central place, because the structure, extensiveness, and scope of organizations almost entirely determined by communication techniques.”

Project communication has been widely recognized that it had the very important impacts to the relations within the project organization [1], [2], as well as the external relations of the organization, as owners, partners etc. [3], [4]. Although a lot of research and practice improved that project communication has the important influence to the project success [5], [6], [7], there is no evidence shows how communication factors affect the project performance. This study aimed to the relation between project communication and performance. Significant correlation between project communication and project performance should be improved by the quantitative methods. Then, we further explored the connections among the communication ways, frequency, channels, feedback and project performance.

2 Methodology

Reference to the empirical methods at home and abroad for the communication, this study used the questionnaire to collect data and make the quantitative analysis.

2.1 Questionnaire

A particular questionnaire was designed to explore the relationships between project communication and performance. There were two parties in this questionnaire. One was about project performance, another was about project communication.

1. The project performance factors

We must determine what factors were used to define project performance. In engineering, three factors usually used to estimate the project performance: project schedule, project cost and project quality. Therefore, we measure the project performance according to the three factors. That is, shorter schedule (shorted as P1), lower cost (P2), and higher quality (P3) mean the better project performance.

2. The project communication factors

It must be determined what factors were chose to represent project communication. Project communication is very complex, and includes a widely range of contents. It is very difficult and critical to determine the project communication factors. In order to get the typical communication factors, we have following references:

1) The Communication Satisfaction Questionnaire (shorted as CSQ), designed by Downs and Hazen (1977). The questionnaire had been used in the more than 50 communication study of different countries, different industries and different organizations since it came out in 1977. The stability of its communication factors had been improved by a lot of research [8].

2) On the base of above questionnaire, Xiaojun Qian designed the questionnaire of employee satisfaction according with China's actual situation. He made an empirical study of employee satisfaction for Chinese employee [9].

3) The questionnaire designed by Linda S. Henderson for the exploratory study of project managers' competency in two core communication processes – encoding and decoding [6].

4) The communication advices proposed by J. Rodney Turner and Ralf Muller who study the relationship between project team and project owners [3].

5) Phillip G. Glampitt and Cal W. Downs explored the relationship between communication and productivity. The result showed that communication was perceived to have an above average impact on productivity and the communication satisfaction factors differentially impacted productivity. The personal feedback factors had a significant impact on productivity [10].

Combined with the above research results and the features of project communication, our questionnaire determined 4 factors to represent project communication: communication climate (shorted as C1), communication channels (C2), communication frequency (C3), and personal feedback (C4).

3 Communication Items

In order to collect the specific data effectively, we needed some communication items to descript the four factors. According to the related researches and the features of project communication, 18 items (showed by Table 1) were chose to descript the communication.

Table 1. Project Communication Items

Project Communication factors	Communication items
Communication Climate (C1)	the degree of getting the correct information (X1) Communication climate satisfaction (X2) Emergency communication satisfaction (X3) the degree of getting the change information (X4) the way of reporting exception (X5)
Communication Channels (C2)	the degree of joining the plan meeting (X6) the degree of preparing meeting files (X7) the degree of changing the unsuited plan (X8) the degree of repeating communication (X9) the way of getting the change information (X10)
Communication Frequency (C3)	the frequency of publishing the project performance (X11) the frequency of reporting the progress (X12) the summary of phrase experience (X13) the publication of summary report (X14)
Personal feedback (C4)	the content of publish report (X15) the degree of understanding the execution of whole plan (X16) the degree of understanding the organization' targets and policies (X17) the degree of understanding the personal work estimation (X18)

4 Questionnaire Designed

Based on the factors and items mentioned above, we developed a questionnaire with three parts to collect related information.

1) The context of the participants' basic situation.

In this part, the participants were asked to provide biographical information (e.g. gender, age, and years of project experience).

2) Project performance evaluated

The participants were asked to give the information about the schedule, cost and quality of projects which they were involved over the past 3 years.

3) Project communication evaluated

The participants were asked to estimate the project communication (e.g. the feeling about the communication climate, the degree to get the change information).

The questionnaire used Likert Scale, which measuring the degree to that people agree or disagree with a statement, usually have 5 items: 5. Strongly agree; 4. Agree; 3. Neither agree nor disagree; 2. Disagree; 1. Strongly disagree.

The participants were asked to evaluate the questions according to their experience and select the correct option. At last, we should get a score. The higher score stand for the higher degree. We could make use of these scores to do quantitative analysis.

To ensure the credibility of the questionnaire, we did a small sample survey at first, handled the collected data and modified the survey questions. Then, reliability and validity of the questionnaire were tested. The test result showed the questionnaire passed the examination. Finally, the formal survey questionnaire was formed.

2.2 Participants Selection

The purpose of this research was explored the project communication impacts on the project performance. So the sample should be these men or women who ever joined projects and have the project experience.

The study chose a data collection method previously utilized by Posner, Sotiriou, and Linda to generate a convenience sample of project management professionals. In their respective researches, these researchers surveyed a subject participant during project management seminars. Similarly, we surveyed participants of MBA or Masters of Engineering of one famous university. The participants had different background of work and project, and met the demand of the study.

180 questionnaires were sent out and 95 available questionnaires were retrieved, including 79 males, or 83.2% of the total, and 16 females, or 16.6% of the total. They are all project managers.

3 Results

The survey items were factors analyzed using Correlation Method by SPSS software. The collected data was analyzed from two aspects described by Fig. 1.

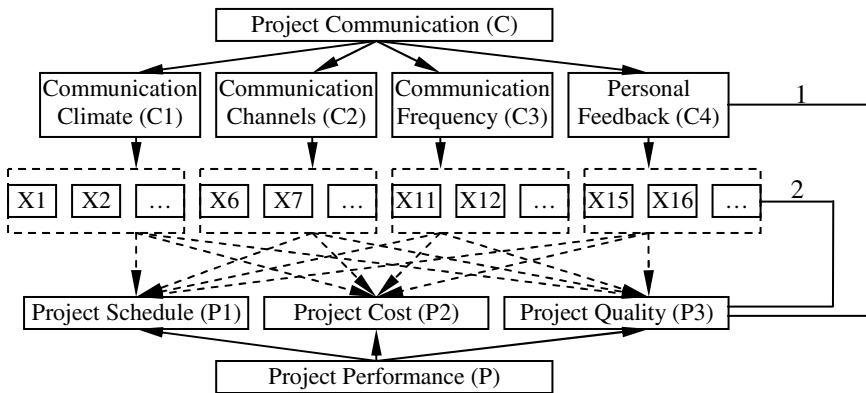


Fig. 1. The Factors Analysis System

3.1 The Analysis of the Relations between the Project Communication Factors and Project Performance Factors

We test the correlation between project communication factors and project performance factors. The Table 2 showed the results.

Table 2. Correlation Coefficients between Communication Factors and Performance Factors

The Project Communication Factors	The Project Performance Factors			
	Schedule	Cost	Quality	Total mount
Communication Climate	0.434**	0.385**	0.295**	1.114
Communication Channels	0.223*	0.282**	0.205*	0.71
Communication Frequency	0.286*	0.066	0.086	0.438
Personal Feedback	0.332**	0.308**	0.352**	0.992
Total mount	1.275	1.041	0.938	

* In confident coefficient (double measurement) of 0.05, the correlation is significant

** In confident coefficient (double measurement) of 0.01, the correlation is significant

From the table, we could get 2 suggestions:

1) Communication had the most significant impact on project schedule.

The biggest vertical total mount of correlation coefficients was 1.275, which between communication and project schedule. That mean the greatest impact of communication was on project schedule.

2) The relations between the communication factors and performance factors were differently close.

We could see the communication factor which had the most significant impact on project performance was communication climate, which was significantly positive with all three performance factors at the confidence score of 0.01, and the total correlation coefficient was 1.114, the largest figure in horizontal total numbers. Meanwhile, communication climate had the more impact on project schedule than cost and quality.

Followed by communication climate, the factor that also had the significant positive with all three performance factors at the confidence score of 0.01 was personal feedback. More over, the closest performance factor to personal feedback was quality, followed by schedule and cost.

The third factor was communication channels which was significant with cost at the confident score of 0.01, with schedule and quality on the confident score 0.05. That meant the communication channel had the greatest impact on project cost.

Finally, the factor that had the weakest impact on performance factors was communication frequency, which only was significant with schedule at the confident score 0.05.

3.2 The Analysis of the Relations between the Communication Items and Project Performance Factors

The result was showed in Table 3.

The communication items were the specific communication behaviors. The data analysis might give us some suggestions to improve performance by changing the items. We could get some results below:

1) The item *the degree of getting the correct information* had the most significant positive with all three project performance factors in all communication items.

Table 3. The Correlation Coefficients between Communication Items and performance

No.	Project Communication Items	Project Performance			
		Schedule	Cost	Quality	Total Amount
X1	the degree of getting the correct information	0.341**	0.378**	0.355**	1.074
X2	Communication climate satisfaction	0.241*	0.311**	0.297**	0.849
X3	Emergency communication satisfaction	0.249*	0.263**	0.174	0.686
X4	the degree of getting the change information	0.409**	0.264**	0.261*	0.934
X5	the way of reporting exception	0.205*	0.164	0.046	0.415
X6	the degree of joining the plan meeting	-0.077	0.162	-0.045	0.04
X7	the degree of preparing meeting files	0.102	0.13	0.198	0.43
X8	the degree of changing the unsuited plan	0.173	0.232*	0.233*	0.638
X9	the degree of repeating communication	0.236*	0.102	0.086	0.424
X10	the way of getting the change information	0.29**	0.21*	0.107	0.607
X11	the frequency of publishing the project performance	0.262*	0.048	-0.064	0.246
X12	the frequency of reporting the progress	0.138	-0.16	-0.137	-0.159
X13	the summery of phrase experience	0.293**	0.103	0.158	0.554
X14	the publication of summary report	0.154	0.084	0.133	0.371
X15	the content of publish report	0.189	0.204*	0.104	0.497
X16	the degree of understanding the execution of whole plan	0.247*	0.187	0.174	0.608
X17	the degree of understanding the organization' targets and policies	0.282**	0.229*	0.292**	0.803
X18	the degree of understanding the personal work estimation	0.149	0.145	0.294**	0.588
	Total Amount	3.883	3.056	2.3465	

The result meant that it was very important to getting the correct project information for project staff. To get correct information not only improved the productivity, but also had the positive affection to project team member, just like “the work support of project team was good, I can get the correct information as soon as quickly that make my job easier” Or “the team pay attention to my job, what I needing can get soon”.

2) The items *the degree of getting the change information*, *Communication climate satisfaction*, and *the degree of understanding the organization' targets and policies* had the most significant correlation with two performance factors in confident score of 0.01, and had significant correlation with other performance factor in confident score of 0.05.

These items could affect the project performance. The project managers might make these items better to improve performance.

3) The items had the correlation more or less with project performance factors except four communication items. E.g. we could see that the item *the degree of understanding the personal work estimation* had the significant correlation with quality factor; item *the summery of phrase experience* had the great impact on project schedule etc. The project manager might take corresponding measures to mend performance.

4) There were some negative correlation coefficients in the analytical data just like the coefficients of the items as *the degree of joining the plan meeting, the frequency of reporting the progress*. The result suggested that not more meeting, more progress reporting, the better project performance.

4 Conclusion

The analytic study made use of the project communication questionnaire designed by ourselves to collect data about project performance and communication, took the project management staff for the survey samples. The correlations between project communication and project performance were analyzed. The results suggested that the communication had the most significant impact on project schedule, and different communication factors had the different effects on performance. The study could provide the communication methods to improve project performance.

References

1. Barki, H., Rivard, S.: Toward an assessment of software development risk. *Journal of Management Information System* 10(2), 203–305 (1993)
2. Loosemore, M.: The three ironies of crisis management in construction project. *International Journal of Project management* 16(3), 139–144 (1998)
3. Rodney Turner, J., Muller, R.: Communication and co-operation on project between the project owner as principal and the project managers as agent. *European Management Journal* 22(3), 327–336 (2004)
4. Carolyn, B., Akintoye, A.: An analysis of success factors and benefits of partnering in construction. *International Journal of Project Management* 18 (2004)
5. Willoughby, K.A.: Process improvement in project expediting: there must be a better way. *International Journal of Project Management* 23, 231–236 (2005)
6. Henderson, L.S.: Encoding and decoding communication competencies in project management – an exploratory study. *International Journal of Project Management* 22, 469–476 (2004)
7. Thompson, P., Perry, J.: World Bank funded projects, 1974-1988. *Engineering construction risks: SERC Report*. British Institution of Civil Engineering (1992)
8. Downs, C.W., Hazen, M.D.: A factor analytic study of communication satisfaction. *The Journal of business communication* 14(3) (1977)
9. Qian, X., Zhan, X.: The factor analysis and Empirical study of communication satisfaction. *Management Comments* 17(6) (2005) (in Chinese)
10. Clampitt, P.G., Downs, C.W.: Employee perceptions of the relationship between communication and productivity: a field study. *The Journal of Business Communication* 30(1) (1993)

Time-Frequency Filtering for LFM Based on Time S-Transform and Time Inverse S-Transform

HuaShan Chi^{1,2}, HongXing Wang¹, Min Zhang², and PeiHong Zhao¹

¹ Department of Electronic Information Engineering,

Naval Aeronautical and Astronautical University, Yantai 264001, China

² Optical Fiber Sensor Laboratory, Department of Electronic Engineering,
Tsinghua University, Beijing 100084, China

yaohsin1979@gmail.com, 13371368601@189.cn,

minzhang@mail.tsinghua.edu.cn, zph1980@163.com

Abstract. The Linear Frequency Modulation (LFM) signal is a classical non-stationary random signal. And the processing method to LFM must expand from frequency domain to time-frequency domain. The S transform (ST) is a time-frequency analysis method which has an adaptive window, and the width changes with frequency. The classical ST algorithm is Time S-transform, and the corresponding inverse ST (IST) is Time IST. In this paper, the Time ST and Time IST algorithms are applied to the filtering of LFM. Simulation results show us that the method is effective to the LFM filtering.

Keywords: Linear Frequency Modulation, Time S transform, Time inverse S transform, time-frequency filtering.

1 Introduction

The Linear Frequency Modulation (LFM) is widely used in many fields, such as radar, sonar and communications. It is important to research the LFM filtering method. The LFM signal is a classical non-stationary signal. The traditional signal processing method is based on the Fourier transform, which can only describe the whole time characters of the signal and the local time characters are neglected. And the time-frequency analysis method must be adopted to describe the local time characters of the signal.

The S transform (ST) is a time-frequency analysis method which have many good characters, such as it has an adaptive window, and the width changes with frequency. In this paper, the Time ST algorithm, the Time IST algorithm and the time-frequency filtering method are applied to the LFM filtering.

This paper is organized as follows. In Section 2, the time ST and its inverse transform algorithms are described. In Section 3, the time-frequency filtering approach and its application to the LFM filtering are described. In Section 4, the conclusion is made.

2 Algorithms of Time S Transform and It's Inverse

The ST time-frequency analysis method was proposed by Stockwell [1], and it has been applied in various fields, such as geophysics [2], power quality analysis [3], and climate analysis [4].

The definition of ST is called Time ST, and the corresponding inverse algorithm is called Time IST.

2.1 Time S-Transform Algorithm

The ST of a continuous time signal is defined as [1]

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t)w(t - \tau, f) \exp(-i2\pi ft) dt, \tag{1}$$

where $w(t - \tau, f)$ is usually the Gaussian window

$$w(t - \tau, f) = \frac{|f|}{\sqrt{2\pi}} \exp(-(t - \tau)^2 f^2 / 2), \tag{2}$$

The center of the Gaussian window is τ , and the width of the window is controlled by f . The Gaussian window is employed in [1], but other window functions can also be used [5], [6].

Substitute (2) into (1), the Time-ST in the continuous form is derived as

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} \exp(-(t - \tau)^2 f^2 / 2) \exp(-i2\pi ft) dt. \tag{3}$$

In (3), letting $\tau \rightarrow jT$, $f \rightarrow n / NT$ and $t \rightarrow pT$, j and $n = 0, \dots, N-1$, where T denotes the sampling period, N denotes the number of samples, the Time-ST in the discrete form is derived as

$$S(jT, n / NT) = \sum_{p=0}^{N-1} h(p) \frac{|n|}{N\sqrt{2\pi}} \exp(-n^2 (p - j)^2 / 2N^2) \exp(-i2\pi pn / N), n \neq 0. \tag{4}$$

To make the following inverse operation exact, the average of the discrete time series is assigned to $S(jT, 0)$ with the zero frequency.

$$S(jT, 0) = \frac{1}{N} \sum_{p=0}^{N-1} h(p). \tag{5}$$

2.2 Time Inverse S-Transform Algorithm

To derive the Time inverse ST algorithm, a time-time function is defined as

$$x(\tau, t) = h(t) \exp[-(t - \tau)^2 f^2 / 2]. \tag{6}$$

It is obvious that (6) is transformed to $x(t, t) = u(t)$ at $\tau = t$ for any frequency f . It supplies a way to derive $h(t)$ via

$$h(t) = x(\tau, t)|_{\tau=t} . \tag{7}$$

From the window function expression (2), we can derive

$$\exp[-(t - \tau)^2 f^2 / 2] = \frac{\sqrt{2\pi}}{|f|} w(t - \tau, f) . \tag{8}$$

Substitute (8) into (6), the time-time function can be rewritten as

$$x(\tau, t) = \frac{\sqrt{2\pi}}{|f|} h(t) w(t - \tau, f) . \tag{9}$$

Applying the Fourier transform to both sides of (9), we can derive

$$X(\tau, f) = \int_{-\infty}^{\infty} \frac{\sqrt{2\pi}}{|f|} h(t) w(t - \tau, f) \exp(-i2\pi ft) dt = \frac{\sqrt{2\pi}}{|f|} S(\tau, f) . \tag{10}$$

where $X(\tau, f)$ denotes the Fourier transform of function $x(\tau, t)$. And then via inverse Fourier transform of $X(\tau, f)$, the function $x(\tau, t)$ can be derived as

$$x(\tau, t) = \int_{-\infty}^{\infty} X(\tau, f) \exp(i2\pi ft) df = \sqrt{2\pi} \int_{-\infty}^{\infty} \frac{S(\tau, f)}{|f|} \exp(i2\pi ft) df . \tag{11}$$

From (7) and (11), the Time ST is derived as

$$h(t) = x(\tau, t)|_{\tau=t} = \sqrt{2\pi} \int_{-\infty}^{\infty} \frac{S(t, f)}{|f|} \exp(i2\pi ft) df . \tag{12}$$

Letting $t \rightarrow jT$ and $f \rightarrow n / NT$ in (12), the Time inverse ST in the discrete form is derived as

$$h[jT] = \sum_{n=-N/2}^{N/2-1} \frac{\sqrt{2\pi}}{|n|} S[j, n] \exp(i2\pi nj / N) . \tag{13}$$

3 Application to the LFM Filtering

In the actual communication system, the LFM signal is interfered by noise. We suppose that all noise is additive white Gaussian noise (AWGN) in this paper unless otherwise stated. Fig.1a shows the ideal LFM signal without noise. Fig.1b illustrates the ST spectrum of the signal in Fig.1a in the form of contour. Fig.1c illustrates the ST 3D spectrum. Fig. 2a shows the LFM signal interfered by noise and the SNR is set to 5dB. Fig.2b shows the ST spectrum of the signal interfered by noise in Fig.2a. The ST spectrum in Fig.2b is interfered by a lot of cluttered curves which are caused by the noise. And the goal of the designed filter is to suppress the cluttered curves caused by the noise and consequently the noise in the LFM signal is suppressed. Fig.2c illustrates the ST 3D spectrum of the signal with noise.

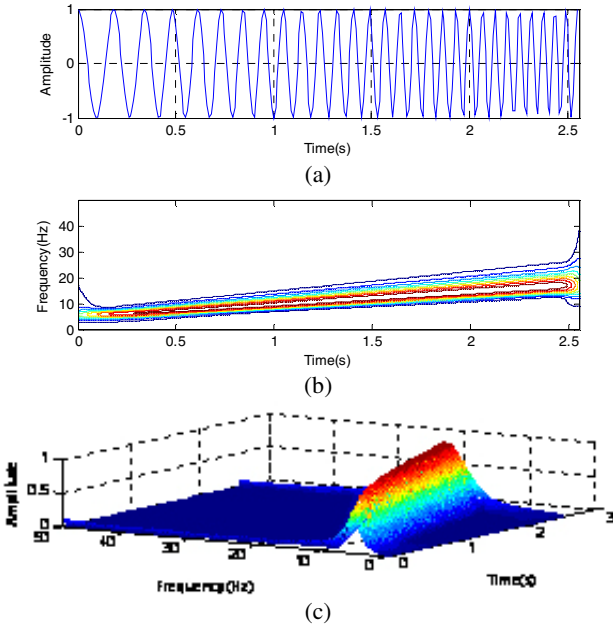


Fig. 1. The ideal LFM signal. (a) The waveform of the ideal LFM signal. (b) The ST spectrum of the signal in Fig.1a in the form of contour. (c) The ST spectrum of the signal in Fig.1a in the 3D form.

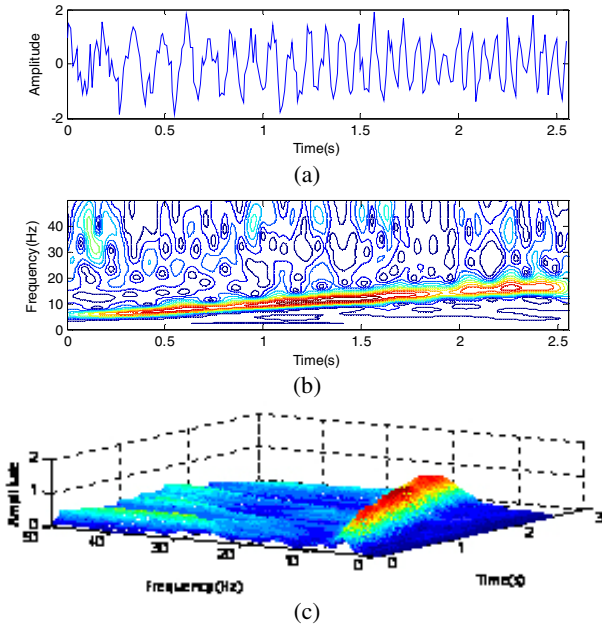


Fig. 2. The LFM signal interfered by noise. (a) The waveform of the LFM signal interfered by noise. (b) The ST spectrum of the signal in Fig.2a in the form of contour. (c) The ST spectrum of the signal in Fig.2a in the 3D form.

A filtering approach named self-filtering is proposed in [7]. The designed filter is defined by

$$TFF(\tau, f) = \frac{|S(\tau, f)|}{\max_{\tau, f}(|S(\tau, f)|)}, \tag{14}$$

where $TFF(\tau, f)$ denotes the designed filter, $|S(\tau, f)|$ denotes the amplitude of the ST representation of the signal interfered by noise, and $\max_{\tau, f}(|S(\tau, f)|)$ denotes the maximum of $|S(\tau, f)|$ over the plane (τ, f) . In (14), dividing $|S(\tau, f)|$ by its largest value over the (τ, f) plane ensures that the peak amplitude of $TFF(\tau, f)$ is 1. Those points in the (τ, f) plane with smaller amplitude are considered to be derived from the noise. After the filtering operation, the amplitude gets smaller or becomes zero for the points with smaller amplitude, so the noise attenuation filtering is achieved.

Multiplying the ST representation of the LFM signal interfered by noise by the user designed time-frequency filter, the ST representation after filtering is derived via

$$S_f(\tau, f) = S(\tau, f) \cdot TFF(\tau, f) = S(\tau, f) \frac{|S(\tau, f)|}{\max_{\tau, f}(|S(\tau, f)|)}, \tag{15}$$

where $S_f(\tau, f)$ denotes the ST representation of the signal after filtering, $S(\tau, f)$ denotes the ST representation for the signal interfered by noise, and $TFF(\tau, f)$ denotes the self-filtering filter.

Fig.3a and Fig.3b show the ST spectrum after filtering. We can see that the ST spectrum after filtering gets cleaner after filtering operation, and the cluttered curves caused by the noise are mostly restrained.

The ST has a very useful character which makes the ST time-frequency filtering possible, that is to say, the integral of ST over τ is the Fourier transform of $h(t)$. We can derive the Fourier transform function of the filtered signal

$$H_f(f) = \int_{-\infty}^{\infty} S_f(\tau, f) d\tau = \int_{-\infty}^{\infty} S(\tau, f) TFF(\tau, f) d\tau, \tag{16}$$

where $H_f(f)$ denotes the Fourier transform of the filtered signal. And the filtered signal is derived from an inverse Fourier transformation,

$$h_f(t) = \int_{-\infty}^{\infty} H_f(f) \exp(i2\pi ft) df. \tag{17}$$

Synthesizing (14), (15), (16) and (17), the filtered signal is derived as

$$h_f(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} S(\tau, f) \frac{|S(\tau, f)|}{\max_{\tau, f}(|S(\tau, f)|)} d\tau \right\} \exp(i2\pi ft) df. \tag{18}$$

Fig.3c illustrates the waveform of the filtered signal. Comparing Fig.3c with Fig.2a, we can see that the signal interfered by noise is improved after filtering.

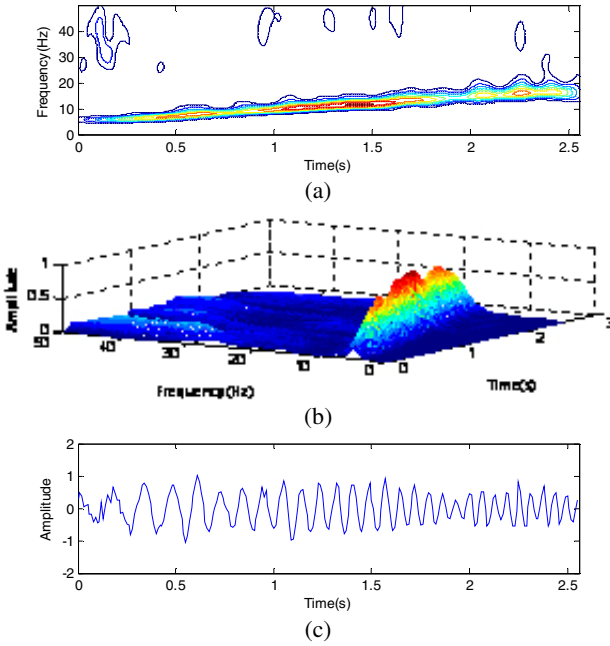


Fig. 3. The ST spectrum and the signal waveform after filtering. (a) The ST spectrum after filtering. (b) The ST spectrum after filtering in 3D form. (c) The signal waveform after filtering.

To evaluate the performance of the filtering method, the criterion of the mean square error (MSE) is introduced. The MSE of is defined as

$$MSE_h = \frac{1}{N} \sum_{k=1}^N [h(k) - h_p(k)]^2, \tag{19}$$

where $h_p(k)$ denotes the ideal noiseless LFM signal. Considering that the real parts of the complex signals, $h(k)$ and $h_p(k)$, have the practical significance, the real parts should be extracted and used in the Matlab simulation for (19).

Fig.4 illustrates the MSE comparison for the LFM signal with noise and the filtered signal. In Fig.4, the subscript ‘n’ indicates the signal with noise, the subscript ‘f’ indicates the signal after filtering. The simulation was carried out for 20 times.

From Fig.4, we can see that in the 20 times of simulation, the MSEs of the LFM signal after filtering, MSE_{hf} , are less than that of the LFM signal with noise, MSE_{hn} , which shows us that the noise is restrained after filtering.

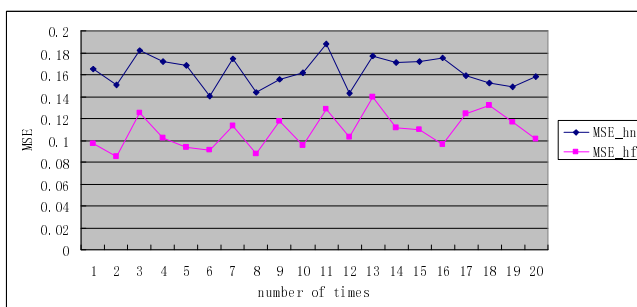


Fig. 4. MSE comparison of LFM signal with noise and the filtered signal

4 Conclusion

In this paper, we applied the time-frequency filtering based on the Time S-transform and its inverse algorithm to the filtering of the LFM signal, the result of experimental simulation shows that the noise is distinctly restrained after filtering and the Time ST time-frequency filtering approach is effective to noise attenuation and useful to the LFM filtering.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Grant No.60772056).

References

1. Stockwell, R.G., Mansinha, L., Lowe, R.P.: Localization of the complex spectrum: the S transform. *IEEE Trans. Signal Process.* 44(4), 998–1001 (1996)
2. Osler, J., Chapman, D.: Seismo-acoustic determination of the shear-wave speed of surficial clay and silt sediments on the scotian shelf. *Canad. Acoust.* 24, 11–22 (1996)
3. Dash, P.K., Panigrahi, B.K., Panda, G.: Power quality analysis using S-transform. *IEEE Transactions on Power Delivery* 18(2), 406–411 (2003)
4. Espy, P.J., Hibbins, R.E., Jones, G.O.L., Riggins, D.M., Fritts, D.C.: Rapid, large-scale temperature changes in the polar mesosphere and their relationship to meridional flows. *Geophys. Res. Lett.* 30(5) (2003)
5. McFadden, P.D., Cook, J.G., Forster, L.M.: Decomposition of gear vibration signals by the generalized S transform. *Mechan. Syst. Signal Process.* 13, 691–707 (1999)
6. Pinnegar, C.R., Mansinha, L.: The S-transform with windows of arbitrary and varying shape. *Geophysics* 68, 381–385 (2003)
7. Pinnegar, C.R., Eaton, D.W.: Application of the S transform to prestack noise attenuation filtering. *J. Geophys. Res.* 108(B9), 2422 (2003)

A BP-LZ77 Compression Algorithm Based on BP Network

HongWei Yan, HuiShan Lu, and Qiang Gao

College of Mechanical Engineering & Automatization, North University of China,
Taiyuan, Shanxi, China
Yhwei-hh@163.com, 13934597379@139.com, gaoqiang919@sohu.com

Abstract. LZ77 algorithm uses limited window to find matches in previous text. It must perform string matching with previous cache for each location of sliding window, the process time of which will increase. Based on analysis of LZ77 dictionary compression algorithm as well as advantages and disadvantages of BP neural network, a novel dictionary compression algorithm BP-LZ77 optimized by BP network was presented. The performance of BP-LZ77 algorithm is greatly improved while inherits time advantages of LZ77 compression algorithm. The combination reason of LZ77 and BP network was given. The general idea of BP-LZ77 was described and details of LZ and BP parts were also provided. Computation and analysis verify feasibility and practicality of BP-LZ77.

Keywords: Dictionary compression, BP, LZ77, data compression.

1 Introduction

Data compression technology decrease redundancy by re-encoding data. In 1977, Israeli mathematician Jacob Ziv and Abraham Lempel proposed a new data compression coding method Lempel-Ziv algorithm family, including LZ77, LZ88 and some variants. For its superior characteristics as simple and efficient, it is now the basis of primary data compression algorithms. LZ series algorithm belongs to lossless data compression algorithm. It is achieved with dictionary encoded technology, which mainly includes four major algorithms as LZ77, LZSS, LZ78 and LZW. Among them, LZ77 algorithm is notable for short compression time. The major compression tools are impacted from it. The rapid development of neural network technology becomes application of it in data compression technology reality. Based on through research on LZ77 and BP network, a novel data compression program BP-LZ77 that combines these two was presented.

The paper is organized as follows: section 2 introduces advantages and disadvantages of LZ77 and BP network; section 3 designs BP-LZ77 algorithm; section 4 analyzes on performance parameters as compression ratio and compression time and section 5 concludes our work.

2 LZ77 Algorithm and BP Network

2.1 LZ77 Algorithm

Jacob Ziv and Abraham Lempel described a technique based on sliding window cache. The cache is used to store recently processed text. The algorithm is commonly known as LZ77. LZ77 and its variants found that words and phrases in text stream are likely to repeat. In case of one repeat, the repeated sequence can be substituted by a short code. Compression program scanned such repeat and generate code to replace repetitive sequences. As time goes, code can be reused to capture new sequence. Algorithm should be designed as extraction program to derive current mapping of original data sequence.

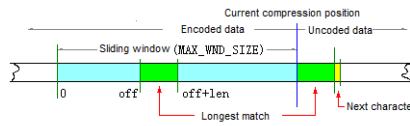


Fig. 1. LZ77 algorithm

The compression algorithm of LZ77 and its variants use two caches. Sliding history cache contains N source characters been processed previously. Prior cache contains next L characters to be processed. The algorithm tries to match two or more characters of prior cache with that in sliding history cache. If no match is found, the first character in prior cache is outputted as 9 bit character and into slide window, and the oldest character is shifted out. If matches found, algorithm continues scan to find the longest match. Then the matching string is outputted as triple, namely indicator, pointer and length. For string with K characters, the longest characters in sliding window are shifted out. K characters been encoded are shifted into window.

Although the LZ77 is valid and appropriate for current status, it has show deficiencies. The algorithm uses limited window to find match in previous text. For very long text blocks related to window size, many possible matches may be lost. The window size can be increased, which will cause two losses. One is that the algorithm processing time will increase for it should perform once character matching with prior cache for each location of sliding window. The second is the field should be longer to permit longer jump.

2.2 BP Neural Network

BP neural network is also known as back-propagation network, which is a kind of iterative gradient algorithm to solve minimum mean square difference between actual output of forward network and expected output. BP is a kind of multiple mapping network with reverse transform and can correct error, which constitutes input layer, one or more hidden layer and output layer. In case of appropriate parameters, BP

neural network can converge to smaller mean square error, which is one of the most widely used networks [2]. Theoretical studies have shown that neural network has a strong nonlinear mapping ability. Hornik et al has proved with functional analysis that three-layer BP neural network can approximate any continuous function and its infinite derivation as long as enough neurons of hidden layer [3].

2.3 LZ77 Algorithm Analysis

The control on compression time of LZ77 algorithm is stronger than most other compression algorithms. The dictionary size should be increased to improve matching probability so as to improve compression ratio. The manner result in enormous pressure for seeking maximum match of character, such as greatly increase traversal time and lost advantages of compression time. Furthermore, for some complex data object, the compression ratio can only be limited increased simply increasing dictionary size.

The principle and characters of BP neural network determine that it can be used in data compression. Many literatures have researched on application of BP neural network in data compression and designed some algorithms. The experiments results show that compression algorithm based on neural network can indeed obtain a higher compression ratio. But it has shortcomings as too long compression time and too many system resources be occupied. In order to get better compression effect, the number of neurons will be increased exponentially. As to some data file, the number of neurons may reach 1012 [4], which is clearly unacceptable. This is also the main reason that mainstream compression tool based on neural network technology has not been launched. Although it is difficult to overcome bottleneck of compression when BP neural network been used for data compression, it can be used to optimize dictionary compression algorithm. Thus, the advantages of these two can be integrated to achieve a satisfactory result.

3 BP-LZ77 Compression System Design

3.1 Overall Design

BP-LZ77 includes two parts of LZ77 and BP network. After data enter into LZ77, the maximum match founded output in the form of triple. For the data that has not found maximum match, it will be inputted into BP with a group of 8 characters. The specific flow is shown in Fig. 2.

3.2 LZ77 Design

Appropriate reduction of dictionary length will further improve compression rate. Set sliding window size as 18KB and minimum value of maximum match length as 8 characters. It means the match length greater than or equal to 8 can find the longest match. For the data has found longest match, record its match length and location information. The data that has not found longest match is input into BP part.

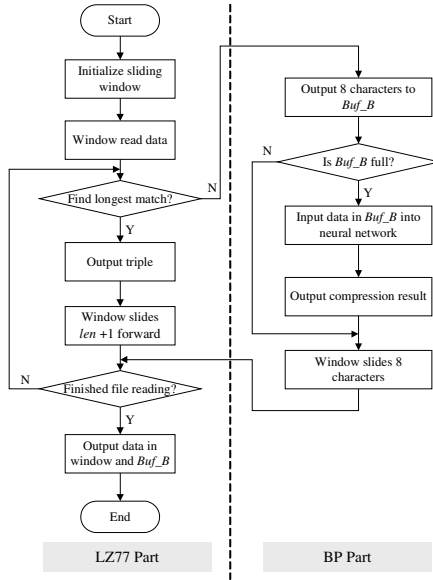


Fig. 2. Overall flow of BP-LZ77

3.3 BP Network Design

Firstly, record location information of data been inputted into BP network. Then the data is sent into buffer *Buf_B* whose size is 256Byte. When the data amount in *Buf_B* reach 256Byte, it means *Buf_B* is full. These data can be inputted into neural network for compression.

(1) Neural network layer number and neurons design

Selection of network layer number and setting of neurons weight threshold are keys to build appropriate BP neural network. As the data amount of input once is limited to 256Byte, so it only needs to build a three-layer neural network with input layer, single hidden layer and output layer.

Divide bit string constituted by 0 and 1 into lines each 8 bit according to order, we can get a matrix with 256 rows and 8 columns. Perform binary coding on row number of the matrix in order, namely each 8-character row number corresponds to 8-bit data. Then, neural network is used to simulate mapping between row number and row data. As the input and output are 8 bit, the neurons number of input layer and output layer is set 8. The row number and row data are used to train neural network. Based on mapping character to be simulated, set initial value of weight and threshold as random number in (-1, 1). Select connection weight and threshold of all neurons and conduct binary encoding in turn. Provide 16 bit code for each weight and threshold, where 1 bit is sign, 1 bit is to mark decimal point, 7bit used for encoding of latter two code and 7bit used for encoding of integer before decimal point [3].

(2) Improvement of promote ability of BP network

Promote ability is one of important indexes to check whether the neural network be powerful or not. Based on [5], Bayesian regularization method is used to improve network training performance parameters from mean square function M to M' .

$$M = \frac{1}{N} \sum_{i=1}^N e_i^2 = \frac{1}{N} \sum_{i=1}^N (\alpha_i - a_i)^2 \tag{1}$$

$$M' = \gamma M + (1 - \gamma) \frac{1}{N} \sum_{i=1}^N \omega_i^2 \tag{2}$$

Where, e_i is error; α_i is the expected output; a_i is actual output; γ is scale factor and ω_i is network connection weight. In the training process, the value of γ will automatically adjust to optimal.

With new performance index function, smaller weight can be ensured while keeping training error as small as possible. It is equivalent to reducing the network size, so as to reduce resource consumption and data redundancy and further improve compression efficiency.

4 Compression Performance Analysis

4.1 Compression Ratio

Data in compression object can be divided into two kinds, namely string with maximum match and that has not found maximum match. As to the former, the data amount after compression with BP-LZ77 and LZ77 is same. As to the latter, LZ77 can not compress, while the BP part of BP-LZ77 can play a role. The obtained compression ratio is about 7 [6].

For the file with same amount D_0 , assume the data amount after LZ77 compression algorithm is D_{11} . Compression ratio is N_1 . The data amount after BP-LZ77 compression is D_{12} and compression ratio is N_2 . As the compression ratio is the ratio of size before and after compression, so

$$N_1 = \frac{D_0}{D_{11}} = \frac{D_0}{D_1 + D_2} \quad \text{and} \quad N_2 = \frac{D_0}{D_{12}} = \frac{D_0}{D_1 + D_2} = \frac{D_0}{D_1 + \frac{D_2}{7}}$$

Where D_1 is data amount after compression with LZ77 algorithm while found the maximum match; D_2 is the directly output data amount; D_2' is data amount of D_2 after compression with BP network.

For a 500MB data object, 85% data can be compressed with LZ77 and 15% is not compressed. So the directly output data amount is $500 \times 15\% = 75MB$. Compress the data object with BP-LZ77, the remaining 75MB data will be compressed by BP part.

Assume the BP-LZ77 compression ratio improvement is ΔN compared with LZ77 algorithm, then $\Delta N = (75 - \frac{75}{7}) / 500 = 13.46\%$.

It can be seen from the above equation that the compression ratio improvement of BP-LZ77 is very significantly. If the compression object is large amount complex data as remote diagnosis data and remote sense collection data, the advantage of compression ratio will be more obvious.

4.2 Compression Time

In BP-LZ77, the principle of LZ part is same with that of LZ77, while the dictionary part is only 16KB, which can fully reflect fast characteristics of dictionary compression algorithm. The BP part uses simple three-layer structure. Input and output neuron number is only 8 and input data amount into neural network is only 256Byte, so as to avoid significantly increase of time and space complexity. In summary, BP-LZ77 algorithm inherits advantages of LZ77 in compression time.

5 Conclusion

The paper presented novel dictionary compression algorithm BP-LZ77 optimized by BP neural network. Its overall idea and specific design of each part were also provided. Based on theoretical calculation and analysis, the compression ratio of BP-LZ77 is significantly improved compared with LZ77. At the same time, it overcomes shortcomings of too long compression time based on neural network compression algorithm, which has strong practical and promotional value.

References

1. Ziv, J., Lempel, A.: A universal algorithm for sequential data compression. Trans. on Information Theory 10(3), 337–343 (1997)
2. Wang, W.-J., Song, S., Guo, X.-X.: Overview of Neural Network Adaptive Control. Computer Simulation 8, 132–136 (2005)
3. Said, A., Perlman, W.A.: A new, fast and efficient image codes based on set partitioning in hierarchical trees. IEEE Trans. on Circuits and Systems for Networks 23, 551–556 (2007)
4. Yan, F., Zhou, J.-L., Wu, Y.: Lossless Data Compression with Neural Network Based on Maximum Entropy Theory. Journal of University of Electronic Science and Technology of China 36, 101–104 (2007)
5. Jiang, L.-X., Xu, P.-P., Yang, G.-Q.: Adaptive dynamic power management model based on artificial neural network. Journal of Huazhong University of Science and Technology (Nature Science Edition) 37, 116–118 (2009)
6. Xue, H., Wu, Y., Liu, X.-S.: A general data compression scheme based on BP neural network. Control & Automation 21, 185–186 (2010)

Modeling and Optimization of Frequency Hopping Sequence Designation in Asynchronous Network

DaoCheng Lei, YongLin Yu, Gang Wang, and RunNian Ma

Telecommunication Engineering Institute, Airforce Engineering University,
Xi'an, Shaanxi, China
wg1x1@nudt.edu.cn

Abstract. The designation of Frequency Hopping (FH) sequences is the crucial factor affecting the quality of FH communication in asynchronous network (AN). According to the requirement of FH communication in AN, some definitions are put forward, the frequency similarity, the sequence similarity, as well as the general similarity, and then the similarity matrix of available FH sequences is constructed. The method of hierarchy clustering analysis (HCA) is introduced to mine the clustering structure of FH sequences in characteristic space. Based on the principle that the similarity of suitable FH sequences used in different networks should be the least, the optimized sequences can be designated from different clusters finally.

Keywords: Frequency Hopping (FH) sequence, asynchronous network (AN), hierarchy clustering analysis (HCA), designation.

1 Introduction

Asynchronous network (AN) is the main mode of FH radio network. Contrast to synchronous network, it has many advantages, such as high organizing speed, low fixed time requested. On the condition that the quality request for AN is high, or the network's frequency points are limited and the probability of sequences collision is high, FH sequences in AN must be optimized [1,2]. However, the former researches are mainly concentrated in sequence generation and the related algorithm, such as Safer+, M sequence creation algorithm and allelism algorithm [2]. From which the sequence is of outstanding randomcity and its properties are decided by the algorithm itself. However, for the case that the frequency hopping bound is fixed relatively, or FH points and FH sequence is designated in advance, the algorithms mostly is in low precision, complex operation, less efficiency and applicability.

In the paper, the research is concentrated in the case that some optional HF sequences has been given in advance and are to be optimized. That is to select some most outstanding FH sequences from the set of a serial of FH sequences. Combined with the properties and requirements of HF communication in AN, the modeling and optimization are given in detail.

The paper is organized as follows. The process of sequence similarity modeling is put forward in part two, in which some definitions are given. The optimization is

presented in third part, including the optimization theory and the process (optimization algorithm). The forth part is simulation to demonstrate the availability and good performance of the suggested modeling and optimization. For simplicity, only two HF network are considered, and the equipment power and channel attenuation in the AN are neglected.

2 System Modeling

In AN, the first and most important problem is the similarity modeling for the HF sequences. For the given HF sequences set, some definitions are to be given in array. FH similarity gives the similarity of random two frequencies in alternative HF sequence. Secondly, the sequence similarity is defined the relationship between two HF sequences in certain location. Thirdly, the definition of general similarity is given to calculate the random two HF sequences, and expressed in the form of matrix finally. After the modeling, hierarchy clustering analysis (HCA) is introduced in simulation; the final best two HF sequences are determined. The whole process is listed in Fig 1.

Denote that n is the number of available frequency points in HF radio AN, the corresponding frequencies are F_1, F_2, \dots, F_n . As well, m the length of each HF sequence, X the set of HF sequences.

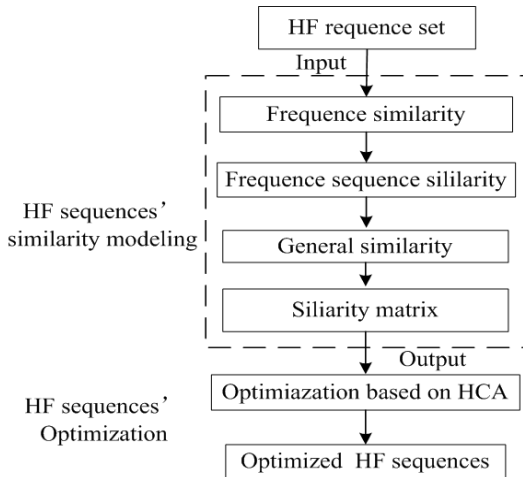


Fig. 1. FH sequences' optimization and selection

2.1 Frequency Similarity

For any $x_i, x_j \in X$, the frequency of position p in x_i is $f_{x_i p}$, and position q in x_j is $f_{x_j q}$. If $f_{x_i p}$ equals to $f_{x_j q}$, direct frequency collision will occur in the corresponding

positions. That's we can define the similarity $k(f_{x_i,p}, f_{x_j,q})$ as 1. If $f_{x_i,p}$ and $f_{x_j,q}$ are just neighbors nearby, a certain indirect frequency collision will occur. That is $k(f_{x_i,p}, f_{x_j,q})$ just between 0 and 1, denoted that $0 < k(f_i, f_j) < 1$. While, if $f_{x_i,p}$ and $f_{x_j,q}$ are in somewhat a distance, there will be not frequency collision at all, that is $k(f_{x_i,p}, f_{x_j,q})$ is zero.

Combine with the character of FH radio communication and the principle of HF programming plan, the similarity of $f_{x_i,p}$ and $f_{x_j,q}$ is as

$$k(f_{x_i,p}, f_{x_j,q}) = \begin{cases} 1, & |f_{x_i,p} - f_{x_j,q}| \leq 0.5\text{kHz} \\ 0.5, & 0.5\text{kHz} < |f_{x_i,p} - f_{x_j,q}| < 4.0\text{kHz} \\ 0, & |f_{x_i,p} - f_{x_j,q}| \geq 4.0\text{kHz} \end{cases} \quad (1)$$

in which, the narrow gap 0.5kHz is decided by the radio frequency gap and excursion requirement, and 4.0kHz the frequency band and HF point setting requirement. In different case of networks and radio performance, the values can be justified adaptively.

2.2 Sequence Similarity

Denote $\hat{\varphi}(x_i, x_j)$ the similarity of sequences x_i and x_j , in detail the sum of all the frequencies' similarity between sequences x_i and x_j .

$$\hat{\varphi}(x_i, x_j) = \sum_{p=1}^m k(f_{x_i,p}, f_{x_j,p}) \quad (2)$$

Formula (2) shows that the higher the similarity between the two sequences, the higher the frequency collision probability, and thus the higher the inner interference between the two HF networks.

2.3 General Similarity

For AN, even if each position of the HF sequences is perfectly matched with each other, for the reason that the time delay is different, sequence x_i may keep up with any one of the cycle sequences of x_j . Thus, denote general similarity of sequences x_i and x_j by $\varphi(x_i, x_j)$, which is defined the maximum value of x_i and all the cycle series of sequence x_j .

$$\varphi(x_i, x_j) = \max\{\hat{\varphi}(x_i, x_j), \hat{\varphi}(x_i, x_j^{+1}), \dots, \hat{\varphi}(x_i, x_j^{+(m-1)})\} \quad (3)$$

In which, x_j^{+r} denote the cycle series of sequence x_j while each of the frequency positions are shifted right r times. Obviously, the position of x_i can be alternatively substituted by the other, that means $\varphi(x_i, x_j) = \varphi(x_j, x_i)$.

2.4 Similarity Matrix

After HCA, sequences will be designated to different cluster. Similarity matrix just designed to expresses the general similarity of each of the sequences in cluster A to any one of other clusters, such as cluster B, cluster C. Based on the similarity matrix, according to the principle that the general sequence similarity is direct proportion to inner interference of AN, the best performance sequence couples are determined from different clusters.

From general similarity definition, for any $x_i, x_j \in X$, element D_{ij} in similarity matrix D is denoted by

$$D_{ij} = \varphi(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \tag{4}$$

in which, D_{ij} is the general similarity of x_i and x_j in two different clusters.

Obviously, the input of HF sequences selection is sequence set X , in which any element is an m -dimension, and any element is just one of n frequency points. In AN, general HF sequence similarity is corresponding to similarity matrix D , and D_{ij} the distance of x_i and x_j in a special feature space. The bigger the similarity value, the closer the two sequences pursuit in the feature space, and vice versa.

3 Optimization Based on HCA

3.1 Basic Theory of HCA

HCA is the most current method in data structure analyzing. The basic idea is dividing the hierarchy tree, and compartmentalizing some disconnecting groups according to the similarity in clusters and the difference among clusters [3]. Its advantage is that many kinds of classified situations can be given varying from coarse standards to fine ones. The typical clustering result can be displayed by a clustering tree. There are two subways in carrying out HCA, combination and division. As to combination (upward from the bottom), every sample is divided into a cluster and then reduce the number of cluster by combining different clusters. As to division (downward from the top), all the samples are divided into one cluster and then increase the number of cluster by dividing the cluster.

New interval collection among different individuals can be defined from the clustering tree produced [4]. When two different clusters are combined, the interval between them is defined as the interval including these two clusters. So, the process of finding clustering tree can be viewed as the transform from original collection d_{ij} to new collection \widehat{d}_{ij} . \widehat{d}_{ij} meets the inequality

$$\widehat{d}_{ij} \leq \max(\widehat{d}_{ik}, \widehat{d}_{jk}) \text{ to all the objects } i, j, k \tag{5}$$

That means the interval between three different clusters can be used to define an equilateral triangle or isosceles triangle. Symbol $d \rightarrow \hat{d}$ is called surpassing quantity transform.

The following process adopts combination arithmetic in hierarchy clustering analysis method. FH sequences are optimized by analyzing the data structure characters and mode.

3.2 Optimizing Process

Assume that the set of FH sequences has t elements, that means t sequences, and the similarity matrix is D . Then the process of optimizing FH sequences is organized as follows

- a) Define t sequences as t clusters, and the platform height of each cluster zero;
- b) Combine the nearest similarity of two different clusters according to similarity matrix, and make the average similarity as platform height in the clustering tree;
- c) Calculate the similarity between new cluster and current clusters. If the number of cluster is 1, turn to d), otherwise turn to b);
- d) Plot the clustering tree and confirm the number of clusters and the sequences in each cluster;
- e) Choose one sequence from the last two clusters respectively and form a FH sequence group. Constructing new similarity matrix by calculating the two sequences;
- f) Select the smallest similarity in the new similarity matrix. The corresponding sequences are the optimized ones meeting conditions.

4 Simulation Analysis

Assume that the frequency up limit and lower limit are respectively 6MHz and 12.5MHz. Select randomly 30 frequency points and construct 20 sequences with length 10, while all the sequences spanned the set of HF sequences, the corresponding number is as 1~20. Assume that the frequency up limit and lower limit are respectively 6MHz and 12.5MHz. Select randomly 30 frequency points and construct 20 sequences with length 10, while all the sequences spanned the set of HF sequences, the corresponding number is as 1~20.

Via modeling and simulation as has been formulated, the clustering tree of 20 FH sequences are listed in Fig. 2, in which sequences 5,11,4,6,17,13,18,8,15,12,16 are clustered in one set, and the others in the other set. For each of the two sequences, the more the similarity is, the small the similarity value and threshold value. The results is shown in Tab.1, in which the first column lists the number of sequences, the second the cluster that the sequences belongs to, the third the sequences in cluster A, and the forth the sequences in cluster B. In detail, the similarities of the sequences separately in two clusters are also given in Table 2.

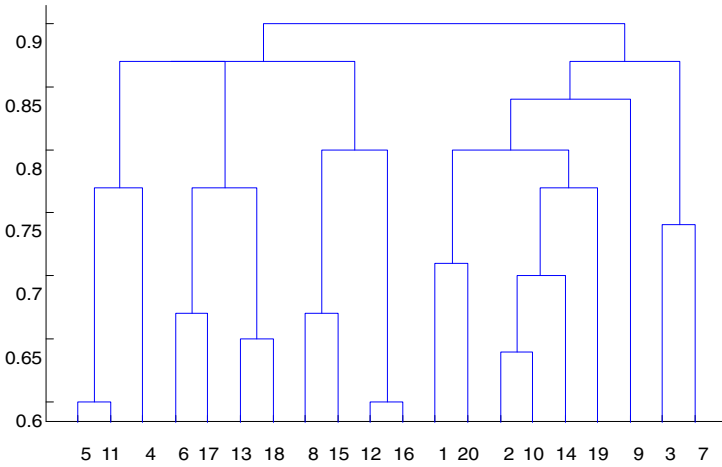


Fig. 2. Clustering tree of 20 FH sequences

Table 1. Results of clustering analysis

Number	1	2	3	4	5	6	7	8	9	10
Y	2	2	2	1	1	1	2	1	2	2
ind1	4	5	6	8	11	12	13	15	16	17
ind2	1	2	3	7	9	10	14	19	20	
Number	11	12	13	14	15	16	17	18	19	20
Y	1	1	1	2	1	1	1	1	2	2
ind1	18									
ind2										

Table 2. Similarity matrix after recombining

Number	4	5	6	8	11	12	13	15	16	17	18
1	0.19	0.16	0.16	0.23	0.26	0.20	0.23	0.20	0.20	0.13	0.26
2	0.22	0.23	0.23	0.23	0.23	0.22	0.26	0.13	0.23	0.23	0.20
3	0.16	0.36	0.23	0.20	0.20	0.20	0.20	0.20	0.16	0.20	0.20
7	0.19	0.13	0.22	0.19	0.20	0.13	0.10	0.20	0.20	0.12	0.10
9	0.10	0.16	0.13	0.13	0.13	0.36	0.23	0.20	0.20	0.23	0.16
10	0.22	0.19	0.20	0.22	0.23	0.12	0.13	0.23	0.20	0.19	0.20
14	0.20	0.30	0.20	0.23	0.20	0.16	0.15	0.13	0.32	0.10	0.26
19	0.23	0.13	0.23	0.20	0.26	0.22	0.16	0.16	0.16	0.20	0.22

From the two clusters, one sequence is selected respectively, and a couple of sequences are constructed. Thus, we can get the similarity matrix as listed in table 2. From with, we see that the smallest similarity value is 0.1, and the corresponding couples are (4, 9) (7, 13) (7, 18) and (14, 17). All these couples satisfy the optimization requirement. That is the inner interferences are low. In practical applications, in case that sequence couple is selected from the sequences set, the above suggestion may be available. Sometimes, only one sequence is designated, and another sequence is to be found, the results can also be found in table 2.

5 Conclusion

FH sequence designation is a crucial factor affecting the quality of FH communication in asynchronous networks. In the paper, modeling and optimization method is introduced. The simulation results show that the presented method is available and effective. In the coming future, more attention should be paid in its practical application, and the more complex cases.

Acknowledgement. This paper is supported by National Natural Science Foundation of China (61174162).

References

1. Kumar, P.V.: Frequency-hopping code sequence designs having large linear span. *IEEE Transactions on Information Theory* 34(2), 146–151 (1988)
2. Molisch, A.F.: *Wireless communications*. Cambridge University Press, London (2011)
3. Cormio, C., Chowdhury, K.R.: Common control channel design for cognitive radio wireless ad hoc networks using adaptive frequency hopping. *Ad Hoc Networks* 8(4), 430–438 (2010)
4. Yuan, H., Chen, H.F., Yao, D.Z., Chen, W.F.: A New Hierarchical Clustering Analysis and the Application in Localization of Brain Activities. *Chinese of Journal Electronics* 15(4), 679–681 (2006)
5. Liu, B., Wang, Y., Wang, H.: Using genetic algorithm based fuzzy adaptive resonance theory for clustering analysis. *Journal of Harbin Engineering University* 27(B07), 547–551 (2006)

Chinese Demographic Risk Early-Warning Indicator System and Empirical Research

Zeng Yongquan

College of Sociology, Huazhong Normal University, Wuhan City,
Hubei Province, P.R. China 430079

Abstract. Constructing the demographic risk early-warning indicator system (DREWIS) is an important link of Monitoring demographic risk and optimizing the demographic policy. The article designs the basic framework of DREWIS on the basis of expert consultation method. Thus it puts forward the determination of the weights of indicators through the expert consultation and the analytic hierarchy process(AHP) and the cluster analytic method, makes dimensionless transformation through the maximum and minimum law, and constructs the demographic risk comprehensive evaluation models and risk rating standard. Based on this, the article comprehensively evaluates the risk situations of the Chinese demographic in recent 10 years.

Keywords: Demographic risk Early-warning, Index system, Comprehensive evaluation.

CLC number: C91 **Document code:** A

1 Introduction

Population is the body of social life, thus it is essential to keep away from demographic risk for the society's healthy operation and sustainable development. The birth rate is decreasing dramatically under involuntary situation for family planning implementation, yet it still maybe rebound and leads to the maladjustment of population structure and population function. Strengthening early-warning of demographic risk, establishing demographic risk early-warning indicator system, as well as timely monitoring and disseminating emergency alarms, can provide scientific basis for government departments' pre-control, and correct decision making.

2 Construction of Demographic Risk Early-Warning Indicator System

Demographic risk early-warning indicator system is composed of a series of substantive individual indicators which are interrelated and can express the status and running processes of demographic risks. As a specific measurement tool and instrument, this index system must possess adequate theoretical basis and scientific method guidance.

This thesis is based on vast access to domestic and foreign related literature, specifically analyzes the advantages and disadvantages of each indicator and promotes the trade-offs, through the discussions and corrections within the research group, preliminarily studies out a set of early-warning indicators. On this basis, adopts Delphi method to pre-select the indicators. There will be a total of 12 consultants, all of these experts have associate professor's professional title and above, a number of doctoral professors are also among them, who are engaged in directions of research such as sociology, social risks, and social problems for 10 years of research age and above, and have interest, time and energy to complete the consultation work. Through the methods mentioned above, this thesis establishes a set of demographic risk early-warning indicators system as follows:

Table 1. Demographic risk early-warning indicator system

Primary indicators	Secondary indicators	Unit
Population size	Rate of population natural increase	‰
Population quality	Schooling years per capital	year
	Average expected life	year
	Mortality of children under the age of five	‰
Population structure	Population aging level	%
	Sex ratio at birth	%

In this indicator system, demographic risks include population size, population quality and population structure these three primary indicators and six secondary indicators.

3 Methods of Demographic Risk Early-Warning

The purpose of the construction of demographic risk early-warning indicator system is to evaluate and monitor risks, and to achieve the object, we must adopt scientific evaluation methods. Comprehensive evaluation method is a method being widely used currently. When a set of indicator system has been established, there is a need to determine the weights of each indicator and conduct non-dimensional transformation of indicator values and construct comprehensive evaluation models, to estimate the demographic risk values currently in China.

3.1 The Determination of Weights of Demographic Risk Early-Warning Indicators

This thesis is based on expert consultation, combined with analytic hierarchy process (AHP) and clustering methodology to empower every indicator, for the purpose of integrating qualitative analysis and quantitative analysis, through which we can not only adequately use the experts' authority knowledge and experience to play the initiative of experts, but also avoid subjectivity and haphazardry as much as possible

through certain technical means. Specifically, the determination of weights mainly includes the following steps: (1) expert consultation; (2) Calculate the similarity coefficient of expert opinions; (3) removing the matrix of the largest deviation; (4) Determine the eigenvectors of judgment matrix; (5) Determine the weights of indicators; (6) Normalization.

There will be 11 experts to participate in this expert consultation, respectively from universities and research institutions of Wuhan, Beijing, Shanghai and Chongqing, where 9 of them have positive senior professional titles, a number of doctoral professors and researchers are also among them, where two of them with the vice senior professional titles. Through expert consultation, get access to expert advice on various indicators. By calculating and programming, obtain the indicator weights at all levels. Specific analysis results are as follows:

Table 2. The determination of weights of demographic risk indicators

Experts	Similarity coefficients	Corresponding weights of experts' judgment matrix			The largest eigenvalue	CR
Expert 1	9.5210	0.2381	0.4286	0.3333	3.0000	0.0000
Expert 2	9.7752	0.2381	0.3333	0.4286	3.0000	0.0000
Expert 3	9.6678	0.3600	0.3600	0.2800	3.0000	0.0000
Expert 4	10.0573	0.2800	0.3600	0.3600	3.0000	0.0000
Expert 5	9.6919	0.2632	0.2632	0.4737	3.0001	0.0000
Expert 6	9.8171	0.3333	0.2381	0.4286	3.0000	0.0000
Expert 7	9.8171	0.3333	0.2381	0.4286	3.0000	0.0000
Expert 8	10.0529	0.3333	0.3333	0.3333	3.0000	0.0000
Expert 9	10.0529	0.3333	0.3333	0.3333	3.0000	0.0000
Expert 10	9.4715	0.3913	0.2174	0.3913	3.0001	0.0000
Expert 11	10.0573	0.2800	0.3600	0.3600	3.0000	0.0000
Final weights		0.3076	0.3153	0.3771		

Through the above study, we can draw the weights of demographic risk early-warning indicator system at all levels, as follows:

Table 3. The determination of weights of demographic risk early-warning indicator system

Indicators	Primary indicators	Secondary indicators
Demographic risk I	Population size 0.3076	Rate of population natural increase 1
		Schooling years per capital 0.3352
	Population quality 0.3153	Average expected life 0.2938
		Mortality of children under the age of five 0.3710
		Population aging level 0.5178
	Population structure 0.3771	Sex ratio at birth 0.4822

3.2 The Non-dimensional Approach

In order to achieve comprehensive assessments of demographic risk, the dimensions of each indicator must be unified, that is to conduct dimensionless treatment of various indicators different units represent. According to the meanings and data collection's different difficulty degree of various indicators, this thesis will use two dimensionless approaches to treat every indicator data. Schooling years per capital,

average expected life, Mortality of children under the age of five and population aging level, the dimensions of these four indicators is unified by minimax method, whereas, the dimensions of the rate of population natural increase and sex ratio at birth is unified by the treatment of interval.

Maximum and minimum value and interval value of this thesis are defined mostly basing on two standards: (1) Regard international practice as reference standard. Such as society population aging, sex ratio at birth, average expected life, the death rate of baby and so on all are defined as universal standards. (2) Use the all-round well-off society indicator and harmonious society indicator as reference standard, such as the rate of population natural increase and schooling years per capital.

3.3 Demographic Risk Assessment Model

After acquiring weight values and evaluation scores of various indicators at all levels, we can use multi-level linear weighted comprehensive assessment method to calculate its evaluation value required for demographic risk evaluation. The formula is

$$v_t = \sum_{i=1}^3 w_i \sum_{j=1}^m w_{ij}^{(1)} v_{jt}^{(2)}$$

Where, v_t denotes the comprehensive assessment value of demographic risk system in time period t , $v_{jt}^{(2)}$ represents the j secondary sub-index risk value in time period t , w_i represents the weight of the i primary sub-index corresponding to demographic risk system, $w_{ij}^{(1)}$ represents the weight of the j secondary sub-index corresponding to the i primary sub-system.

3.4 The Levels of Demographic Risk Early-Warning

After calculating the assessment value of demographic risk, determine the levels of social risk according to the following table, and identify with the appropriate early-warning signals.

Table 4. Demographic risk early-warning assessment form

Risk value	0-2	2-4	4-6	6-8	8-10
Warning level	No warnings	Slight warnings	Medium warnings	Serious warnings	Emergency warnings
Signals	Green light	Blue light	Yellow light	Orange light	Red light

In the table, the green light means the development of the population system is stable. Demographic conditions are safe, the possibility of risk occurrence is very low; blue light means that the population development is in a transition period towards stable or turbulent, the demographic conditions are relatively safe; yellow light indicates the population development begins to appear small turmoil, the risks may occur at any time; orange light indicates the population development is in disarray, the risks have affected social stability; red light means the demographic condition is in a state of extreme confusion, large-scale risk events can occur at any time.

4 Empirical Study of Demographic Risk Comprehensive Assessment in China

The data collection is a key element of comprehensive evaluation. The data collected according to demographic risk early-warning indicator system is as follows:

Table 5. Data sheet of demographic risk early-warning indicators

Specific indicators	Units	In 2000	In 2001	In 2002	In 2003	In 2004	In 2005	In 2006	In 2007	In 2008	In 2009
Rate of population natural increase	‰	7.58	6.95	6.45	6.01	5.87	5.89	5.28	5.17	5.08	5.05
Schooling years per capita	year	7.79	7.84	7.88	7.93	7.97	8.02	8.25	8.40	8.48	
Average expected life	year	71.4	71.71	72.02	72.33	72.64	72.95	73.26	73.57	73.88	
Mortality of children under the age of five	‰	39.70	35.90	34.90	29.90	25.00	22.50	20.60	18.10	18.50	17.2
Population aging level	%	6.96	7.1	7.3	7.5	7.6	7.7	7.9	8.03	8.3	8.5
Sex ratio at birth	% Female=100		115.6 5	119.8 5	117.5 4	121.1 8	118.8 8	119.5 6	125.4 8	120.56	119.45

The data is mainly derived from “China Statistical Yearbook”, “Statistical Yearbook of Chinese Society”. Schooling years per capital and average expected life based on the theory of China builds a well-off society in an all-round way Process Tracing developed by the research team of build a well-off society in an all-round way of the State Statistics Bureau.

By calculating, the risk values from 2000 to 2009 are as follows:

Table 6. Demographic risk values from 2000 to 2009

Time	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Risk value	4.5302	4.3102	4.6902	4.2137	4.3675	4.1652	4.0449	3.9914	3.9757	3.9035

The results can also be depicted by the chart as follows:

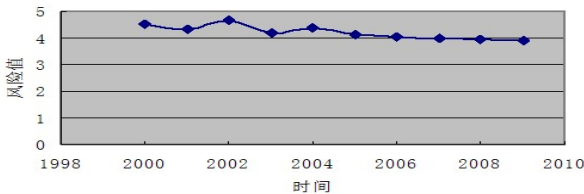


Fig. 1. Demographic risk values from 2000 to 2009

From the above chart we can see that the demographic risk of the past decade in China has been decreasing in general trend. According to the risk warning level evaluation sheet, the warning level of demographic risk of the past decade has been decreased from medium warning to slight warning step by step. This mainly thanks to the long-term planning family policy and the improvement of education and medical and healthy system. It is noticeable that population aging is getting severely and the proportion of the elderly people is rising, which is a tremendous challenge for our country that is not affluent but infirm. Additionally, we must take sex ratio question seriously. Currently, the China's sex ratio at birth exceed heavily the normal range between 102 and 107, and it undoubtedly will cause long-standing disadvantage influence to population development, marriage and family, and sexual relationship.

References

1. Li, Z.: Population Economics. Tsinghua University press (2006)
2. Li, J.: Contemporary Western Population Theorie. Fudan University Press (2004)
3. Huang, R.: The Impact of Population Aging on Economics. Economic Research (4) (2010)
4. Brander, J.A., Dowrick, S.: The Role of Fertility and Population in Economic Growth: Empirical Results from Aggregate Cross-National Data. Journal of Population Economics (7), 100–109 (1994)
5. Werding, M.: Aging and Productivity Growth: Are There Macro-Level Cohort Effects of Human Capital? CESIFO Working Paper No.2207, Category 5: Fiscal Policy. Macroeconomics and Growth, p. 18 (2008)
6. Johnson, R.W., Kaminsk, J.: Older Adults' Labor Force Participation since 1993. A Decade and a Half of Growth, 3–4 (2010)
7. UN, World Economic and Social Survey 2007: Development in an Ageing World, Department of Economic and Social Affairs, p. 75 (2007)
8. Jiang, Y., Fan, Y.: The Decision Theory and Method Based on Judgment Matrix. Science Publishing Company (12), 18 (2008)

A Sparse Recovery Method of DOA Estimation for Uncorrelated Sources with Multi-snapshot

MaoWei Yin^{1,2}, DongYang Xu², and ZhongFu Ye^{2,*}

¹ National Defence School of Sci. & Tech., Southwest Univ. of Sci. & Tech., Fucheng DC
Qinglong Avenue Middle 59, 621010 Mianyang, P.R. China

² Institute of Statistical Signal Processing, University of Sci. & Tech. of China,
Huangshan Road 443, 230027 Hefei, P.R. China

{ymwxk, yezf}@ustc.edu.cn, xudong198602@163.com

Abstract. In this paper, a novel sparse recovery method is proposed for direction of arrival (DOA) estimation on the assumption that source signals are uncorrelated and the number of snapshots is enough. We demonstrate in this paper how to solve DOA estimation problem with multiple measurement vectors (MMV) from the perspective of the correlation rather than from the view of data which appears in almost all sparse signal recovery methods. By using this method we remove the effect of noise and attain both DOA estimation and signal power. Theoretical analysis and experimental results demonstrated that our approach has a number of advantages compared other source localization schemes, including increased resolution, lower bias and variance, unknowing the number of sources, better performance in low SNR, decreased computational complexity.

Keywords: Direction of arrival (DOA), multiple measurement vectors (MMV), correlation, uncorrelated sources, sparse representation.

1 Introduction

Direction of arrival (DOA) estimation has always been an active research area. Some methods have been developed over the years to provide high-resolution and high-accuracy in estimating DOA [1], which can be divided into three broad categories: 1) Beamforming algorithms or data-adaptive spatial filtering algorithms like Capon's method; 2) Subspace algorithms like MUSIC and ESPRIT; 3) Parametric methods like the maximum likelihood (ML) and stochastic ML (SML). The topic of sparse signal representation [2] and compressed sensing (CS) [3] have evolved very rapidly in the last decade. There also has potential application in the context of DOA estimation [4]-[7]. Recursive weighted least-squares algorithm called FOCUSS [4] and sparse bayesian learning (SBL) algorithm [5] can be used to estimate DOA for the single snapshot case, and the extensions of FOCUSS and SBL to multiple measurement vectors (MMV) are M-FOCUSS [6] and M-SBL [7], respectively. However, these methods aforementioned will suffer from some properties degradation

* Corresponding author.

while the number of sources is unknown, and the methods for multi-snapshot case are almost all based on the perspective of the data to solve a sparse-joint recovery problem and do not take good use of the statistical properties of the data.

In this paper, we address the issue of recover sparse signals from multi-snapshot via constructing a rather different framework, where the number of sources is unknown. We start from the traditional model and then extend it to the sparse representation model. At last, we consider l_1 -norm penalty for sparsity which leads to a convex optimization problem. Theoretical analysis and experimental results given subsequently demonstrate that our algorithm works well.

2 Algorithm Scheme

Consider K zero-mean and uncorrelated signals $u_k(t)$, $k \in \{1, \dots, K\}$, impinging on an uniform linear array (ULA) of M omnidirectional sensor corrupted by additive Gaussian noise $n_m(t)$ resulting in sensor outputs $x_m(t)$, $m \in \{1, \dots, M\}$. Let $\{\theta_1, \dots, \theta_K\}$ be the location parameters and θ_k is the direction of the k th signal. In the narrowband case, the array output vector can be represented as

$$\mathbf{x}(t) = \mathbf{A}(\theta)\mathbf{u}(t) + \mathbf{n}(t), t \in \{t_1, \dots, t_T\} . \tag{1}$$

Where $\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)]$ is a $M \times K$ matrix, $\mathbf{u}(t) = [\mathbf{u}(t_1), \dots, \mathbf{u}(t_K)]^T$ is the source waveform vector at time t . The steering vector corresponding to the k th source for a far-field ULA is given by

$$\mathbf{a}(\theta_k) = [1, \nu_k, \nu_k^2, \dots, \nu_k^{M-1}] . \tag{2}$$

Where $\nu_k = e^{-j \frac{2\pi d}{\lambda} \sin(\theta_k)}$, λ is the signal wavelength and d is sensors interspace.

We consider a multi-snapshot case. Let $\mathbf{X} = [\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_T)]$, and define \mathbf{U} and \mathbf{N} similarly. Then, (1) becomes

$$\mathbf{X} = \mathbf{A}\mathbf{U} + \mathbf{N} . \tag{3}$$

Where \mathbf{X} is data matrix and (3) is equivalent to (1). Now, we try to turn out (3) into a vector form. Define

$$r_m = E\{x_m x_1^*\} = E\left\{ \left[\sum_{k=1}^K \nu_k^{m-1} u_k(t) + n_m(t) \right] \cdot \left[\sum_{k=1}^K u_k^*(t) + n_1^*(t) \right] \right\} . \tag{4}$$

Assume signals and noise are both Gaussian distribution and zero-mean, and the power of k th signal is P_{u_k} . According to the assumption of correlation, we can get

$$r_m = [\nu_1^{m-1}, \nu_2^{m-1}, \dots, \nu_K^{m-1}] \cdot \begin{bmatrix} P_{u_1} \\ P_{u_2} \\ \vdots \\ P_{u_K} \end{bmatrix} . \tag{5}$$

Then we can define a vector of the correlation as follow:

$$\mathbf{r} = \begin{bmatrix} r_2 \\ r_3 \\ \vdots \\ r_M \end{bmatrix} = \begin{bmatrix} \nu_1 & \nu_2 & \cdots & \nu_K \\ \nu_1^2 & \nu_2^2 & \cdots & \nu_K^2 \\ \vdots & \vdots & \ddots & \vdots \\ \nu_1^{M-1} & \nu_2^{M-1} & \cdots & \nu_K^{M-1} \end{bmatrix} \cdot \begin{bmatrix} P_{u_1} \\ P_{u_2} \\ \vdots \\ P_{u_K} \end{bmatrix} = \mathbf{B}\mathbf{P} . \quad (6)$$

In the model above, DOA and the power of each signal are concerned, but we can not determine \mathbf{B} and \mathbf{P} in (6). To solve this problem, we extend (6) into a sparse signal representation model with an overcomplete set Φ and a sparse vector \mathbf{s} .

Let $\{\theta_1, \theta_2, \dots, \theta_{N_\theta}\}$ be a sampling grid of all potential of arrival. Then we can represent (6) by

$$\mathbf{r} = \Phi \mathbf{s} . \quad (7)$$

Where $\Phi = [\tilde{\mathbf{a}}(\theta_1), \tilde{\mathbf{a}}(\theta_2), \dots, \tilde{\mathbf{a}}(\theta_{N_\theta})]$, $\tilde{\mathbf{a}} = [\nu_n, \nu_n^2, \dots, \nu_n^{M-1}]$, $n \in \{1, 2, \dots, N_\theta\}$. The n th element $s(n)$ is non-zero and equal to P_{u_k} if k th source comes from θ_n .

Note that Φ differs from \mathbf{B} that it contains steering vector $\tilde{\mathbf{a}}$ for all potential DOAs, rather than only the source signals DOAs. Hence, Φ is known and does not depend on the actual DOAs.

Now we consider DOA estimation as a convex optimization problem, which can be solved by using CVX [8]. Through this method, we can not only reserve the information of multi-snapshot, but also remove the effect of the noise. Figuring out this convex optimization, we can get both source localization and the power of every signals. We define the optimization criterion as follow:

$$\min \|\mathbf{r} - \Phi \mathbf{s}\|_2^2 \quad s.t. \quad \|\mathbf{s}\|_1 \leq \alpha P_{total} . \quad (8)$$

Where P_{total} is the total power of all signals, and α is a performance parameter set here an exponential value of the range from 0.3 to 0.9.

Next, we consider how to estimate the total power of the signals. Assuming that the power of the noise is known as P_n , the total power of all signals takes the form

$$P_{total} = \frac{1}{M} \sum_{i=1}^M E\{x_i^H x_i\} - P_n . \quad (9)$$

We only need a rough estimate for P_{total} due to α has a value of wide range. Also, it means that this algorithm has good robust to noise, especially under low SNR case, which would be discussed in the part of simulation result.

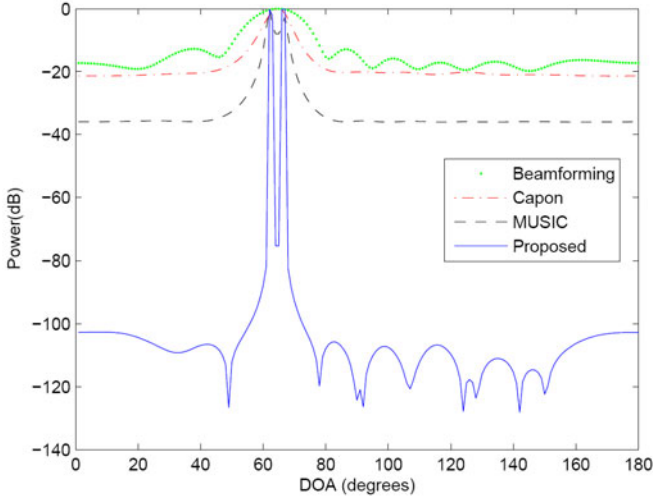
After we get the sparse solution \mathbf{s} , if some element $s(n)$ is bigger than the detective threshold, we can determine the presence of one signal and the index corresponding to $s(n)$ is the DOA of this signal. When we get the estimate of the location parameter, \mathbf{B} can be reconstructed by $\hat{\mathbf{B}}$ with $\{\hat{\theta}_1, \dots, \hat{\theta}_K\}$. Then we can get the least squares (LS) solution by

$$\hat{\mathbf{P}} = (\hat{\mathbf{B}}^H \hat{\mathbf{B}})^{-1} \hat{\mathbf{B}}^H \mathbf{r} . \quad (10)$$

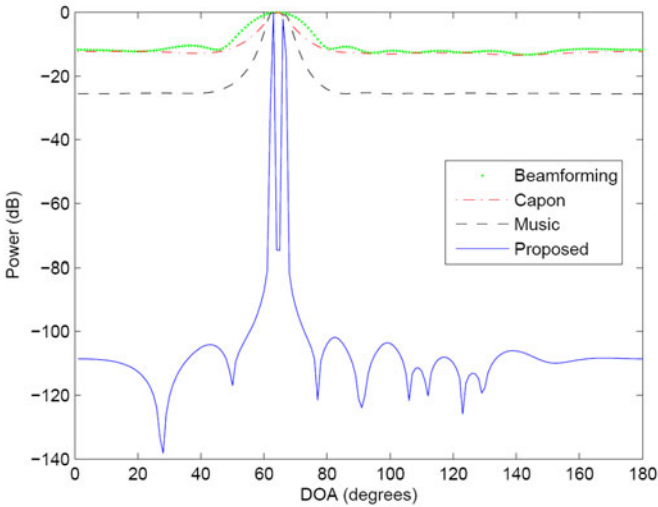
Using this method, we can precisely estimate the power of each signal.

3 Simulation Results

An ULA of $M = 8$ sensors separated by half a wavelength of the actual narrowband source signals is considered. We consider two zero-mean uncorrelated narrowband signals in the far-field impinging upon this array. The number of snapshots is $T = 200$, and the grid is uniform with 1° sampling $N_\theta = 180$. We compare in Fig. 1



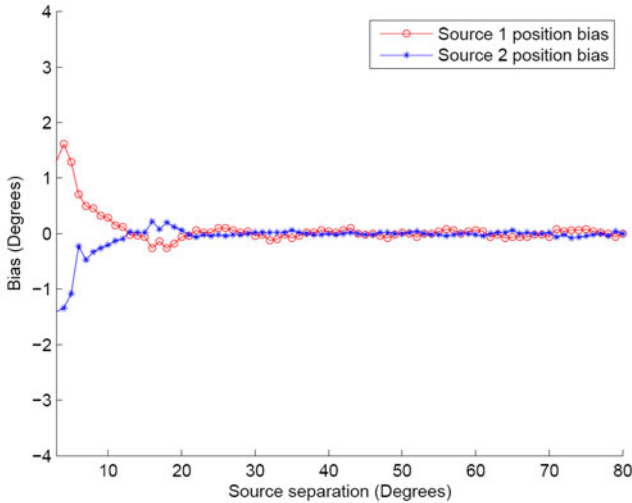
(a)



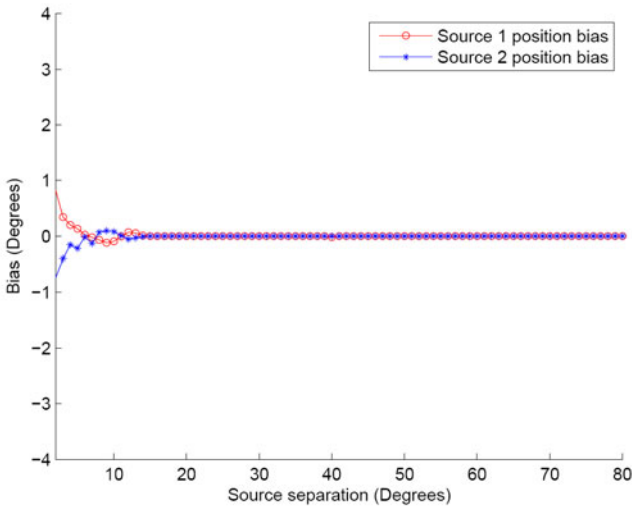
(b)

Fig. 1. (a) and (b). Spatial spectra for beamforming, Capon’s method, MUSIC, and the proposed method for mean-zero and uncorrelated sources. DOAs: 62° and 67° . Top: $SNR = 10$ dB. Bottom: $SNR = 1$ dB.

the spectrum obtained using ours method with those of beamforming, Capon’s method, and MUSIC. SNR is set as 10 dB and 1 dB in the top and bottom plots, respectively. And the sources are closely spaced (5° separation). In the top one, our algorithm and MUSIC can resolve the two sources, while Capon’s method and beamforming method merge the two peaks. When SNR decrease to 1 dB, and only our algorithm can resolve the two sources.



(a)



(b)

Fig. 2. (a) and (b). Bias of the proposed method in localizing two sources as a function of separation between the two sources SNR = 10 dB. Top: $M = 8$, Bottom: $M = 16$.

Next, we consider the bias of this algorithm in the closely spaced sources case. The reason for the bias is that in this case, the steer vectors will be similar and the correlation of these steer vectors will be enhanced. It is obvious that improving the SNR can enhance the estimation accuracy. However, increasing the sensors number can also achieve this goal, for in this instance, the correlation of these steer vectors will decline greatly and the accuracy of DOA estimation will be improved greatly. Now we explore bias more closely via varying the angular separation to estimate DOAs of two sources with different sensors numbers. The snapshots are $T = 200$, and SNR is 10 dB. Assuming that one source locate 50° , the other one changes the location. We get the bias plot shown in Fig. 2, where the sensors numbers are $M = 8$ and $M = 16$ in the top and bottom plots, respectively. The plot shows that the bias decreases when the sensors number increases. And the bias only exists when the SNR is low and disappears when sources are separated respectively by more than 20° .

4 Conclusion

In this paper, we propose a novel method for DOA estimation when the signals are zero-mean and uncorrelated. We start from a traditional model and then derive the sparse representation model. With this preprocessing, we successfully transform the joint-sparse problem into a single sparse vector inverse problem under the multi-snapshot case. Then the computational complexity decreases greatly. Due to the process of correlation and removing the effect of the noise, the robust to noise increases. The following work is to solve the DOA estimate problem that the sources are coherent, near-field and the non-linear array geometry.

References

1. Krim, H., Viberg, M.: Two Decades of Array Signal Processing Research: The Parametric Approach. *IEEE Signal Process. Mag.* 13, 67–94 (1996)
2. Donoho, D.L., Elad, M., Temlyakvo, V.: Stable Recovery of Sparse Overcomplete Representations in the Presence of Noise. *IEEE Trans. Inf. Theory* 52, 6–18 (2006)
3. Candés, E.J., Wakin, M.B.: An Introduction to Compressive Sampling. *IEEE Signal Process. Mag.* 25, 21–30 (2008)
4. Gorodnitsky, I.F., Rao, B.D.: Sparse Signal Reconstruction from Limited Data Using FOCUSS: a Re-weighted Minimum Norm Algorithm. *IEEE Trans. Signal Process.* 45, 600–616 (1997)
5. Tipping, M.E.: Sparse Bayesian Learning and the Relevance Vector Machine. *J. Mach. Learn. Res.* 1, 211–244 (2001)
6. Cotter, S.F., Rao, B.D., Kjersti, E., Kreutz-Delgado, K.: Sparse Solutions to Linear Inverse Problems with Multiple Measurement Vectors. *IEEE Trans. Signal Process.* 53, 2477–2488 (2005)
7. Wipf, D.P., Rao, B.D.: An Empirical Bayesian Strategy for Solving the Simultaneous Sparse Approximation Problems. *IEEE Trans. Signal Process.* 55, 3704–3716 (2007)
8. CVX: Matlab Software for Disciplined Convex Programming, <http://cvxr.com/cvx>

Compensation of Double Closed Loop Control Systems with Time-Delay and Data Packet Losses

HuiSheng Gao, XiaoRui Li, and Jun Wang

Department of Electronics and Communication Engineering,
North China Electric Power University, Baoding Hebei 071003, China
000hwd@163.com

Abstract. With the network-induced time-delay in Networked Control Systems NCS the performance of, control systems degrades even destabilizes. Also in practice the states of some systems are unmeasured. In this paper two state observers are designed to estimate the states and when data packet losses the state observers can compensate it. Then exponential stability theory is used to analyze the stability of the close loop system and the method of how to calculate the control parameter, is given. At last a motor is used to simulate.

Keywords: Networked Control Systems (NCS), state observer; data, double closed loop control.

1 Introduction

Introducing the network into the control system can expand the region of control system and meanwhile guarantee the low link cost of the system, and make the system easy to maintain and diagnose, therefore, the Network Control System (NCS) has been paid more and more attention, and constantly more and more popular. For example, in the reference [1], the position AC SERVO system with the Three-Loop Control Structure can make the speed and position loop do the network control to satisfy the application requirement of the network servo system in industry.

Based on the two outputs, this paper designed the double output control system of the state observer which can compensate for the random time-delay in order to adapt to the quality requirements of the servo system and the double output control system in the similar position.

2 The Description of the Problems

The diagram of the network control system is shown in figure 1. From the figure 1 it can figure out the system has two outputs: the sensors in the system inner ring and outer ring control. In the system, the sensors encapsulate the output outputs of the inner and outer loops, and transfer to the controller through the CAN general line, then the controller calculates the control signal of the inner and outer loops and encapsulates to a packet to send the actuator. The timer samples the controlled objects and the controller drives for the event.

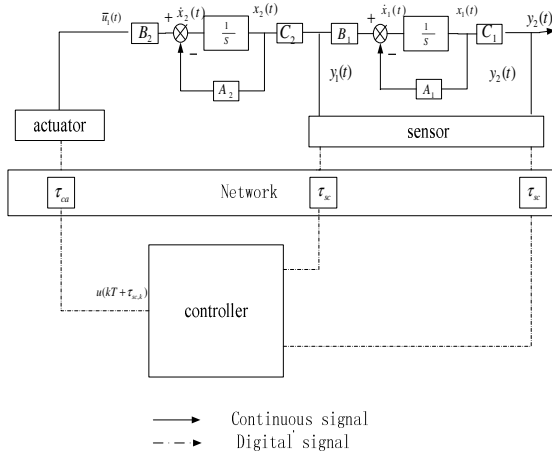


Fig. 1. The structure model diagram of the network control system

In the network control system there are upper bounds from the sensor to the network time delay of the controller τ_{sc} , from the controller to the network time delay of the implementation τ_{ca} . $\tau = \tau_{sc} + \tau_{ca} < T$. According to the time delay statistics the most appropriate and the biggest $\tau_{sc,max}$ was designed. Take the k cycle for example, if a packet arrives in the $t \in [kT, kT + \tau_{sc,max}]$, the controller will calculate the control amount using the time of arrival.

Due to the two outputs in the system, and in the process the state can not be measured directly, therefore, in the following paragraphs the output feedback is used to construct the two observations to reconstruct the state of the controlled objects.

3 The Design of the State Observer

If in the $\hat{x}_i(kT)$ $t \in [kT, kT + \tau_{sc,max}]$ pack arrival is S1, the probability of S1 occurrence will be λ_1 , First, the previous circle calculates the value of the state observation $\hat{x}_i(kT)$, and using the value it can fix the state value of the moment k --- $\tilde{x}_i(kT)$, then the value of the state observation in the following circle--- $\hat{x}_i((k+1)T)$, is estimated. $x_i(kT)$ is the corresponding state value of $y_i(kT)$ and L_i is the output feedback array of the state observer.

1) Use the output to modify respectively the state observation in the k moment:

$$\tilde{x}_i(kT) = \hat{x}_i(kT) + L_i(y_i(kT) - C_i\hat{x}_i(kT)) \tag{1}$$

- 2) Calculate the state value of the state observer and the input signal in the moment $kT + \tau_{sc,k}$. Here. $u_2(kT) = y_1(kT) = C_1 x_1(kT)$:

$$\hat{x}_i(kT + \tau_{sc,k}) = e^{A_i \tau_{sc,k}} \tilde{x}_1(kT) + \int_{kT}^{kT + \tau_{sc,k}} e^{A_i(kT + \tau_{sc,k} - t)} B_i dt u_i(kT) \quad (2)$$

$$u_1(kT + \tau_{sc,k}) = K_1 \hat{x}_1(kT + \tau_{sc,k}) + K_2 \hat{x}_2(kT + \tau_{sc,k}) \quad (3)$$

- 3) Calculate the state value of the state observer- $\hat{x}_i[(k+1)T]$ in the moment $(k+1)T$:

$$\hat{x}_i[(k+1)T] = e^{A_i(T - \tau_{sc,k})} \hat{x}_i(kT + \tau_{sc,k}) + \int_{kT + \tau_{sc,k}}^{(k+1)T} e^{A_i((k+1)T - t)} B_i dt u_i(kT) \quad (4)$$

Because $\hat{x}_i((k+1)T) = \hat{x}_i(kT + \tau_{sc,k})$ and $y_i[(k+1)T] = y_i[kT + \tau_{sc,k}]$ in the period $t \in [kT, (k+1)T]$, from the formula (3) $\tilde{x}_i[(k+1)T]$ can be got:

$$\begin{aligned} \tilde{x}_i[(k+1)T] &= \hat{x}_i((k+1)T) + L_i \left\{ y_i[(k+1)T] - C_i \hat{x}_i[(k+1)T] \right\} \\ &= \hat{x}_i((k+1)T) + L_i \left\{ y_i[kT + \tau_{sc,k}] - C_i \hat{x}_i[kT + \tau_{sc,k}] \right\} \\ &= e^{A_i(T - \tau_{sc,k})} \hat{x}_i(kT + \tau_{sc,k}) + \int_{kT + \tau_{sc,k}}^{(k+1)T} e^{A_i((k+1)T - t)} B_i u_i(t) dt \\ &\quad + L_i C_i (x_i(kT + \tau_{sc,k}) - \hat{x}_i(kT + \tau_{sc,k})) \end{aligned} \quad (5)$$

- 4) From the formula (4) $\hat{x}_i((k+1)T + \tau_{sc,k+1})$ can be got:

$$\begin{aligned} \hat{x}_i((k+1)T + \tau_{sc,k+1}) &= e^{A_i(T + \tau_{sc,k+1} - \tau_{sc,k})} \tilde{x}_i[(k+1)T] + \int_0^{T + \tau_{sc,k+1} - \tau_{sc,k}} e^{A_i t} B_i dt u_i[(k+1)T] \\ &= e^{A_i(T + \tau_{sc,k+1} - \tau_{sc,k})} \hat{x}_i(kT + \tau_{sc,k}) + \int_0^{T + \tau_{sc,k+1} - \tau_{sc,k}} e^{A_i t} B_i dt u_i(kT + \tau_{sc,k}) \\ &\quad + L_i C_i e^{A_i(T + \tau_{sc,k+1} - \tau_{sc,k})} (x_i(kT + \tau_{sc,k}) - \hat{x}_i(kT + \tau_{sc,k})) \\ &= \Phi_i(\xi_k) \hat{x}_i(kT + \tau_{sc,k}) + \Gamma_i(\xi_k) u_i(kT + \tau_{sc,k}) \\ &\quad + L_i C_i \Phi_i(\xi_k) (x_i(kT + \tau_{sc,k}) - \hat{x}_i(kT + \tau_{sc,k})) \end{aligned} \quad (6)$$

Here $\xi_k = T + \tau_{sc,k+1} - \tau_{sc,k}$, $\Phi_i(\xi_k) = e^{A_i \xi_k}$, $\Gamma_i(\xi_k) = \int_0^{\xi_k} e^{A_i t} B_i dt$.

4 The Analysis of the Closed Loop System

4.1 The Error Analysis

Define the error variable $e_i(k)$: $e_i(k) = x_i(k) - \hat{x}_i(k)$.

Define the state augmented variable $z(k)$:

$$z(k) = (x_1(k), x_2(k), e_1(k), e_2(k), \bar{u}_1(k))$$

From the above derivation we can get $z((k+1)T + \tau_{k+1}) = \Lambda_i z(kT + \tau_k)$, in which $i = 1, 2$ and they respectively respond to the defined T1、 T2 above, there into:

$$\Lambda_1 = \begin{bmatrix} \bar{\Phi}_1 + \bar{\Gamma}_1 K_1 & \bar{\Gamma}_2 K_2 & -\bar{\Gamma}_1 K_1 & -\bar{\Gamma}_2 K_2 & 0 \\ \bar{\Gamma}_2 C_1 & \bar{\Phi}_2 & 0 & 0 & 0 \\ 0 & 0 & \bar{\Phi}_1 - L_1 C_1 \bar{\Phi}_1 & 0 & 0 \\ 0 & 0 & 0 & \bar{\Phi}_2 - L_2 C_2 \bar{\Phi}_2 & 0 \\ K_1 & K_2 & -K_1 & -K_2 & 0 \end{bmatrix}$$

$$\Lambda_3 = \begin{bmatrix} \bar{\Phi}_1 & 0 & 0 & 0 & \bar{\Gamma}_1 \\ \bar{\Gamma}_2 C_1 & \bar{\Phi}_2 & 0 & 0 & 0 \\ -\bar{\Gamma}_1 K_1 & -\bar{\Gamma}_1 K_2 & \bar{\Phi}_1 - L_1 C_1 \bar{\Phi}_1 + \bar{\Gamma}_1 K_1 & \bar{\Gamma}_1 K_2 & \bar{\Gamma}_1 \\ 0 & 0 & 0 & \bar{\Phi}_2 - L_2 C_2 \bar{\Phi}_2 & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}$$

4.2 The Stability Analysis

Theorem 1: For asynchronous dynamic system T1, T2, T3, T4 has determined, if the scalar a exists, $a_i > 0$, ($i = 1, 2, 3, 4$) and there also exists positive definite matrix $X_1 \in R^{n \times n}$,

$$X_2 \in R^{n \times n}, Y_1 \in R^{n \times n}, Y_2 \in R^{n \times n}, Z \in R^{m \times m}, \Delta = \begin{bmatrix} X_1 & 0 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 & 0 \\ 0 & 0 & Y_1 & 0 & 0 \\ 0 & 0 & 0 & Y_2 & 0 \\ 0 & 0 & 0 & 0 & Z \end{bmatrix},$$

the conditions will be satisfied:

$$a_1^{r_1} a_2^{r_2} a_3^{r_3} a_4^{r_4} > a > 1$$

$$\Omega_i = \begin{bmatrix} -a_i^{-2} \Delta & \Delta \Lambda_i^T \\ \Lambda_i \Delta & -\Delta \end{bmatrix} < 0$$

Then the system is stable and the attenuation rate is $\gamma = a_1^{r_1} a_2^{r_2} a_3^{r_3} a_4^{r_4}$.

5 The Simulation Example

Apply three-phase permanent magnet synchronous motor model to do simulation study, and make speed loop and current loop do network control. Make use of CAN general line network with network bandwidth $1.60 \times 10^6 \text{ bit/s}$.

According to the literature [1] we can get: $A_1 = \begin{bmatrix} 0 & 1 \\ 0 & -1.66 \times 10^3 \end{bmatrix}$, $B_1 = \begin{bmatrix} 0 \\ 1.91 \times 10^5 \end{bmatrix}$, $C_1 = [1 \ 0]$, $A_2 = [0]$, $B_2 = [1]$, $C_2 = [6.67 \times 10^{-1}]$, and set the attenuation to meet the feedback parameters of the attenuation rate $\gamma = 1.23$. The feedback parameters are $K_1 = [-2.16, -3.21 \times 10^{-3}]$ and $K_2 = [-7.85 \times 10^1]$. The sampling period is $T = 5.00 \times 10^{-4} s$, the maximum delay is $\tau_{\max} = 3.00 \times 10^{-4} s$ and the network delay is any value less than a sampling period. The rate of the network packet from the sensor to the controller and from the controller to the sensor is set to 20%. Take $\alpha_1 = 1.05$, $\alpha_2 = 0.92$, $\alpha_3 = 0.88$, $\alpha_4 = 0.85$, then calculate decay rate $\gamma = 1.23$. Using the LMI toolbox in matlab, the solution is:

$$X_1 = \begin{bmatrix} -2.63 & -1.88 \times 10^3 \\ -1.88 \times 10^3 & 2.30 \times 10^6 \end{bmatrix}, Y_1 = \begin{bmatrix} 4.73 & -3.36 \times 10^3 \\ -3.36 \times 10^3 & 3.98 \times 10^6 \end{bmatrix}, X_2 = [3.45 \times 10^{-2}],$$

$$Y_2 = [3.45 \times 10^{-2}], Z = [3.16 \times 10^1], L_1 = [1.68; -6.12 \times 10^2], L_2 = [1.00].$$

Give the initial state $(x_1(0), x_2(0), x_3(0))^T = (2, 2, 2)^T$ and make the simulation in the MATLAB's "truetime" simulation platform. The figure shows that the system can reach very good stable state under the adjustment of the controller.

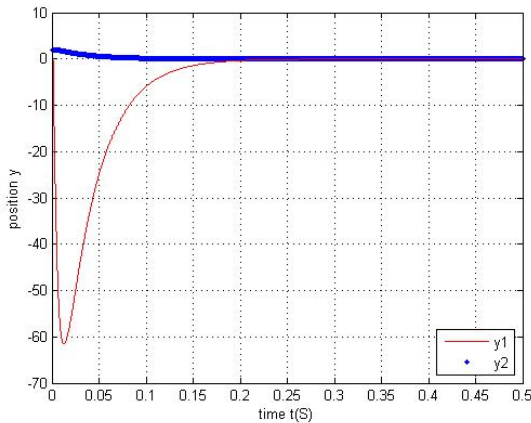


Fig. 2. The output y curve

6 Conclusion

According to two unmeasured output control systems, two state observers were designed for state reconstruction. And according to the conditions of the delay in the

network control system. a system model is established. Besides, this paper uses the exponential stability theory to analyze the conditions of the system stability, uses the simulation way to verify the results, and proves the effectiveness of its methods.

References

1. Wu, Q., Li, Y., Qin, Y.: Study on Design Approach of Reconfiguration Controller in AC Servo Systems. *Computer Emulation* 12, 327–331 (2006)
2. Nilsson, J.: Real-time control systems with delays. Dept. of Automatic Control. Lund Institute of Technology, Lund (1998)
3. Ma, Y., Chen, W., Wang, B.: Compensation of networked control systems with time-delay. *Computer Engineering and Applications* 46(14), 75–78 (2010)
4. Li, S.B., Wang, Z., Sun, Y.X.: Delay-dependent controller design for networked control systems with long time delays: an iterative LMI method. In: Proc. of the 5th World Congress on Intelligent Control and Automation, Hangzhou, pp. 1338–1342 (2004)
5. Kristiansen, R., Nicklasson, P.J.: Spacecraft formation flying: a review and new results on state feedback control. *Acta Astronautica* 65(11-12), 1537–1552 (2009)

Optimization Design of Improving Signal Integrity for Enhancing Differential-Mode Transmission Lines

Wen-Jong Chen and Ming-Huang Wu

Department of Industrial Education and Technology, Changhua University of Education,
No.2, Shi-Da Road, Changhua City 500, Taiwan
wjong@cc.ncue.edu.tw

Abstract. This study proposes the optimization design to improve the integrity of the voltage signals of a conventional differential-mode structure. In this paper, we employ the electronic packages simulation software SPEED 2000 to perform an analysis of signal integrity. Simulation results show that the enhanced differential-mode transmission lines have less EMI than the conventional differential-mode lines at both of the near-end and far-end voltages. Accordingly, this study provides an enhanced signal method by increasing the number of transmission lines that can efficiently diminish EMI of the conventional differential-mode transmission lines and further increase signal integrity.

Keywords: Differential-mode, Common-mode, Signal integrity, EMI.

1 Introduction

Since the CPU transmission circuits are much shorter and smaller than those in PCBs, CPU can sustain higher pulses than a PCB. Short transmission lines result in high signal integrity because they reduce energy consumption. EMI phenomenon must be reduced to improve the transmission efficiency. Designing transmission lines with low EMI is therefore an important issue in high-speed digital technology. Because both capacitance and inductance effects occur between parallel transmission lines as they near each other, these two effects can result in cross-talk, problems in signal integrity, and limitations in clock speed [1-2]. Shi, W. *et al* [3] and Novak I. [4] have mathematically modeled cross-talk in parallel lines, and many previous studies [5-7] discuss this topic as well. Conventional differential-mode lines suffer less signal energy loss while still maintaining reasonable signal integrity. This study presents enhanced differential-mode transmission lines for improving the signal integrity of conventional differential-mode transmission lines.

2 Interference Analysis of Common-Mode and Conventional Differential-Mode Lines

In this section, we describe the equivalent circuit models of common-mode and conventional differential-mode transmission lines. The common-mode structure is

characterized by an in-phase signal (see Fig. 1(a)), while the conventional differential-mode structure has an out-of-phase signal between two transmission lines (see Fig. 1(b)). V_1 and V_2 are voltage of 1st and 2nd parallel lines, respectively. Figures 2(a) and 2(b) show the equivalent circuits of capacitance and inductance. In Fig. 2 (a) and (b), I_1 and I_2 are the current of 1st and 2nd parallel lines respectively, C is the ground capacitance of 1st and 2nd lines, C_m is the mutual capacitance of two parallel lines, L is the self-inductance of 1st and 2nd lines, and L_m is the mutual inductance of two parallel lines.

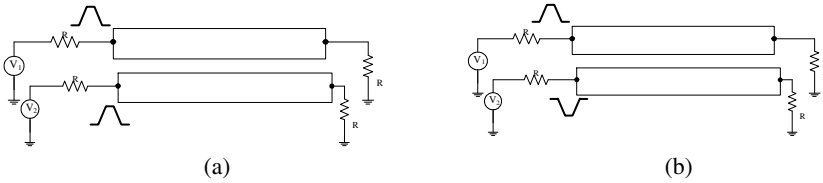


Fig. 1. The equivalent circuit models: (a) common-mode lines, (b) differential-mode lines



Fig. 2. The equivalent circuits: (a) the capacitance effect (b) the inductance effect

2.1 Capacitance and Inductance for Common-Mode Lines

As the system supports an even mode with voltage $V_1 = V_2$, the equation of the analysis of the capacitance effect for common-mode lines is derived as follows,

$$I_i = C \frac{dV_i}{dt} + C_m \frac{d(V_i - V_i)}{dt}, \text{ where } i = 1, 2 \tag{1}$$

In Eq. (1), the 1st term is the currents of the lines for the ground and the 2nd term is that of the mutual capacitance. As the current conditions is $I_1 = I_2$, the analysis of the inductance effect for common-mode lines is as follows,

$$V_i = L \frac{dI_i}{dt} + L_m \frac{dI_i}{dt}, \text{ where } i = 1, 2 \tag{2}$$

The common-mode inductance effect L_c from Eq. (2) equals $L + L_m$. In Eq. (2), the 1st terms are the voltages of the lines for the ground and the 2nd terms are that of the mutual inductance.

2.2 Capacitance and Inductance for Conventional Differential - Mode Lines

As this system supports an odd mode with voltage $V_1 = -V_2$, the analysis of capacitance effect for conventional differential-mode lines is written by:

$$I_i = C \frac{dV_i}{dt} + C_m \frac{d(V_i + V_j)}{dt}, \text{ where } i = 1, 2 \tag{3}$$

In Eq. (3), the meanings of the 1st and 2nd terms are same as Eq.(1). The capacitance effect equals $C+2C_m$ from Eq. (3). In the odd mode, the currents relationship between Line 1 and Line 2 is $I_1 = I_2$. The analysis of the inductance effect for conventional differential-mode lines is as follows,

$$V_i = L \frac{dI_i}{dt} + L_m \frac{d(-I_i)}{dt}, \text{ where } i = 1, 2 \tag{4}$$

The conventional differential-mode inductance L_d from Eq. (4) is given by $L-L_m$.

2.3 Interference for Common-Mode and Differential-Mode Structures

Table 1 compares the results of interference from each mode. The interference from the capacitance effect does not change in the common-mode lines, while the inductance increases L_m . In the differential-mode lines, the capacitance effect raises C_m , but the inductance effects lowers L_m .

Table 1. Comparison of capacitance and inductance of common-mode and differential-mode

Mode	Capacitance effect	Inductance effect
Common-mode lines	Constant	$L + L_m$ (increase)
Differential-mode lines	$C + 2C_m$ (increase)	$L - L_m$ (decrease)

3 Structure and Analysis of the Enhanced Differential-Mode

Fig. 3 depicts the structure of the enhanced differential-mode lines. Lines 1–2 represent signal lines, and lines 3 – 4 represent auxiliary lines. The phenomenon of interference that Line 2 disturbs Line 1 indicates the enhanced differential-mode effect, and Line 3 with disturbing Line 1 indicates the common-mode effect.

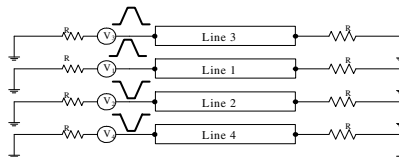


Fig. 3. The structure of the enhancing differential-mode transmission lines

3.1 Analysis of Capacitance Effects

We consider that the current of Line 3 interferes with Line 1 for common-mode lines Fig. 3. The total induced current on Line 1 is as follows,

$$I_1 = C \frac{dV_1}{dt} + (C + C_m) \frac{dV_1}{dt} = (2C + C_m) \frac{dV_1}{dt} \quad (5)$$

From Eq. (5), the effect of induced capacitance on Line 1 increases from C to $2C + C_m$.

3.2 Analysis of Inductance Effects

The inductance effect on Line 1 that is interfered by Line 3 and Line 2 does not increase, as revealed in Eq. (6). The induced interference resulting from the capacitance and inductance effects, which imply that Line 2 is disturbed by Lines 1 and 4, are the same as those of the interference on Line 1 according to the above-mentioned method. The inductance effect induces greater interference than the capacitance effect in transmission lines, but the enhanced structure proposed in this study can therefore counteract a very high level of EMI (inductance effect).

$$V_1 = (L + L_m) \frac{dI_1}{dt} + (L - L_m) \frac{dI_1}{dt} = L \frac{dI_1}{dt} \quad (6)$$

4 Simulation

This study uses SPEED 2000 software to simulate the different modes with an unfixed length. The simulation was performed with the following parameters: a layer of 1.5 mm thickness and a layer medium of 4.3 mm thickness, a line of 0.035 mm diameter and 60 mm length, sinesquare wave for input signal with 2 V (highest voltage) and 0V (lowest voltage), 10 ps ascending and descending times, and 1020 ps period. Fig. 4 (a) illustrates the near-end voltage wave of the conventional differential-mode lines, where V_{ne1} and V_{ne2} denotes the near-end voltage for the first and the second line, respectively. V_{hd} and V_{ld} are the absolute difference which is between the undisturbed and disturbed voltages at the high and low levels, respectively. The simulation results in Fig. 4(a) are $V_{hd} = 0.032V$ and $V_{ld} = 0.033V$. Fig. 4(b) shows the far-end voltage wave V_{fe1} and V_{fe2} with the 1st and 2nd lines of the enhanced differential mode, where $V_{hd} = 0.092V$ and $V_{ld} = 0.064V$. Fig. 5(a) illustrates the deviation voltage of high and low level near-end waves, where $V_{hd} = 0.021V$ and $V_{ld} = 0.020V$. Fig. 5(b) illustrates the far-end voltage wave of the enhanced differential modes, where $V_{hd} = 0.035V$ and $V_{ld} = 0.025V$.

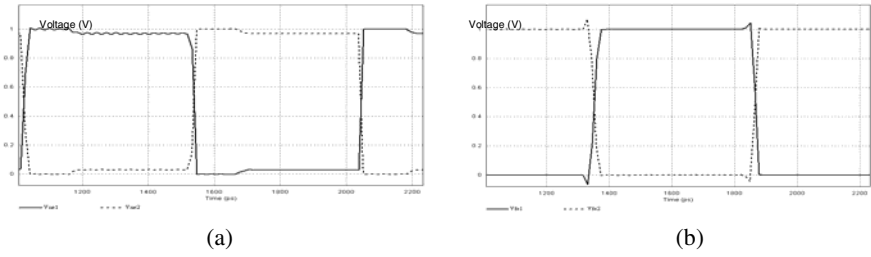


Fig. 4. The signal wave of the conventional differential mode (a) Near-end, (b) Far-end

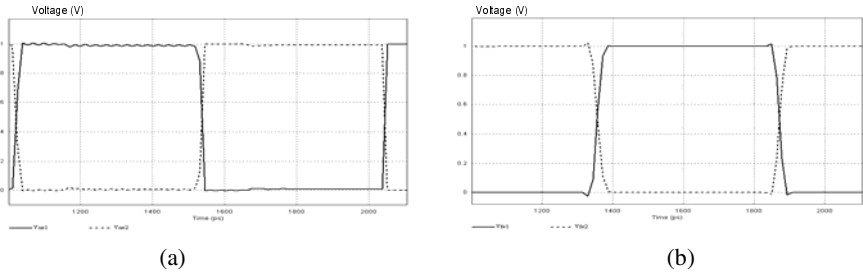


Fig. 5. The signal wave of the enhanced differential mode. (a) Near-end , (b) Far-end

Table 2. V_{hd} and V_{ld} of near-end in different modes (unit of V_{hd} and V_{ld} are voltage)

Rise time (<i>ps</i>)	Conventional differential Mode		Enhanced differential mode	
	V_{hd}	V_{ld}	V_{hd}	V_{ld}
10	0.032	0.033	0.021	0.020
20	0.033	0.030	0.014	0.014
30	0.032	0.032	0.015	0.014
40	0.030	0.030	0.011	0.012
50	0.030	0.030	0.012	0.011
60	0.030	0.031	0.012	0.013
70	0.030	0.030	0.011	0.012
80	0.030	0.030	0.010	0.011
90	0.030	0.030	0.010	0.010
100	0.030	0.030	0.010	0.010

These simulation results in which the V_{hd} and V_{ld} of the near-end and those of far-end voltages occur at different ascending times are described in Tables 2 and 3. These results demonstrate that the enhanced differential mode has less interference than the conventional differential mode at both the near and far ends. Additionally, the enhanced differential modes offer superior signal integrity between these both modes.

Table 3. V_{hd} and V_{ld} of far- end in different modes (unit of V_{hd} and V_{ld} are voltage)

Rise Time (ps)	Conventional Differential Mode		Enhanced Differential Mode	
	V_{hd}	V_{ld}	V_{hd}	V_{ld}
10	0.092	0.064	0.035	0.025
20	0.081	0.079	0.030	0.030
30	0.040	0.021	0.014	0.008
40	0.011	0.030	0.003	0.011
50	0.016	0.014	0.004	0.005
60	0.015	0.010	0.005	0.004
70	0.007	0.011	0.001	0.001
80	0.005	0.003	0.001	0.001
90	0.007	0.006	0.001	0.002
100	0.004	0.004	0.000	0.002

5 Conclusions

The enhanced differential mode offers much better anti-interference and signal integrity than the conventional differential mode. The difference of anti-interference and signal integrity between these two different differential modes gradually increases with transmission speed increases. That is, a higher speed circumstance leads to better anti-interference in enhanced differential mode. In summary, this study presents an optimization method of transmission lines that offers better transmission quality while maintaining good signal integrity.

References

1. Walker, C.S.: Capacitance, Inductance and Crosstalk Analysis. Artech House Publishers (1990)
2. Castillo, S., Mitra, R.: A Study of Crosstalk and Distortion of High-Speed Pulse in Digital Circuits: Research Report R-1033. University of Illinois-Urbana, Illinois (1985)
3. Shi, W., Fang, J.: Evaluation of Closed-Form Crosstalk Models of Coupled Transmission Lines. IEEE Transactions on Advanced Packaging 22, 174–181 (1999)
4. Novak, I.: Modeling, Simulation, and Measurement Considerations of High-Speed Digital Buses. In: Proceedings of the Instrumentation and Measurement Technology Conference, pp. 921–925 (1994)
5. Gilb, J.P., Balanis, C.A.: Pulse Distortion on Multilayer Coupled Microstrip Lines. IEEE Transactions on Microwave Theory and Techniques 37, 1620–1627 (1989)
6. Homo, M., Marques, R.: Coupled Microstrips on Double Anisotropic Layers. IEEE Transactions on Microwave Theory and Techniques 32, 467–470 (1984)
7. Talgat, R.G.: Far-End Crosstalk Reduction in Double-Layered Dielectric Interconnects. IEEE Transactions on Electromagnetic Compatibility 43(4), 566–571 (2001)

Coordinate Transformations in Satellite Navigation Systems

Pengfei Zhang¹, Chengdong Xu¹, Chunsheng Hu¹, and Ye Chen²

¹ School of Aerospace Engineering, Beijing Institute of Technology,
No.5 South Zhongguancun Street, Haidian District, Beijing, 100081, China

² School of Information and Communication Engineering, North University of China,
No.3 Xueyuan Road, Taiyuan, Shanxi, 030051, China
successful.2008@163.com

Abstract. The methods of coordinate transformations in the same satellite navigation system and among different satellite navigation systems were summarized. The former includes a transformation between Cartesian coordinate and geodetic coordinate in Earth-Centered Earth-Fixed (ECEF) coordinate system, transformations among Cartesian coordinates with common point of origin, and a transformation between ECEF Cartesian coordinate system and local geographical coordinate system with uncommon point of origin. For the latter, seven-parameters Bursa model which is one of the most commonly used transformation methods was introduced. Finally, the transformation process of common coordinate systems in satellite navigation systems was concluded so that the data of satellite navigation and communication system could be processed conveniently.

Keywords: Satellite navigation, coordinate system, transformation process.

1 Introduction

The positioning function of satellite navigation system is realized by measuring the distance between receiver and satellites[1], so determining the position of receiver is a basic task in satellite navigation system. However, coordinate of the receiver position needs to be described in the selected reference coordinate system in satellite navigation system. Firstly, different satellite navigation systems have different reference coordinate systems. Secondly, selections of the reference coordinate system are different from each other in the same satellite navigation system according to different navigation tasks. Finally, there are different representation methods in the same coordinate system. Therefore, it is very important to research the transformation between different coordinate systems and representations for the processing of navigation data and result comparison. Transformations among different coordinate systems were summarized in this paper, including transformations in the same satellite navigation system and among different satellite navigation systems. At the end of the paper, transformation process of common coordinate system in satellite navigation system was concluded so that the user could get the different forms of coordinate as quickly as they want.

2 Coordinate Transformations in the Same Satellite Navigation System

2.1 Transformation between Cartesian Coordinate and Geodetic Coordinate in ECEF Coordinate System

In Earth-Centered Earth-Fixed (ECEF) coordinate system, the origin is located in the center of the earth, XY plane coincides with the equatorial plane, X-axis points to the orientation of 0° longitude, Y-axis points to the orientation of 90° east longitude, Z axis is perpendicular to the equatorial plane and points to the geographical Arctic. Transformation between Cartesian coordinate and geodetic coordinate (latitude, longitude, height) is needed in geodesy and navigation data processing[2]. As shown in Fig. 1, point P (X, Y, Z) which corresponds to the geodetic coordinate is (B, L, H). Positive latitude indicates north latitude, positive longitude indicates east longitude.

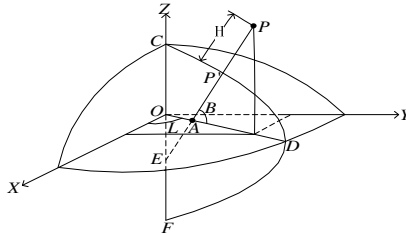


Fig. 1. Cartesian Coordinate and Geodetic Coordinate

(1) BLH ==> XYZ

$$\begin{cases} X = (N + H) \cos B \cos L \\ Y = (N + H) \cos B \sin L \\ Z = [N(1 - e^2) + H] \sin B \end{cases} \quad (1)$$

Where:

$$N = P'E = \frac{a^2}{(a^2 \cos^2 B + b^2 \sin^2 B)^{1/2}} = \frac{a}{(1 - e^2 \sin^2 B)^{1/2}}, \quad a = |OD|, \quad b = |OC|.$$

a represents the semi-major axis of reference ellipsoid, *e* represents the eccentricity of reference ellipsoid. Reference ellipsoid model and parameters of common satellite navigation systems are shown in Table 1.

Table 1. Reference ellipsoid model and parameters of common satellite navigation systems

Navigation system	Reference ellipsoid	Coordinate system	Semi-major axis <i>a</i> (m)	Oblateness reciprocal <i>1/f</i>
GPS	WGS1984	WGS-84	6378137.00	298.257223563
GLONASS	PE1990	PZ-90(PE-90)	6378137.00	298.257839303
GALILEO	GTRF	ITRF-96	6387136.49	298.256450000
BD-1	Krasophuskii1940	BJ-54	6378245.00	298.300000000
BD-2	GRS80	CGCS2000	6378137.00	298.257222101

The relationship between Ellipsoid oblateness f and eccentricity e is:

$$e = \sqrt{2f - f^2} .$$

(2) XYZ ==> BLH

$$\begin{cases} B = \arctan \frac{Z}{\sqrt{X^2 + Y^2}} \left(1 - e^2 \frac{N}{N + H} \right)^{-1} & (\text{When } Z < 0, B = -|B|) \\ L = \begin{cases} \arctan(Y / X) & X < 0 \\ \arctan(Y / X) - \pi & X < 0, Y < 0 \\ \arctan(Y / X) + \pi & X < 0, Y > 0 \end{cases} \\ H = \frac{\sqrt{X^2 + Y^2}}{\cos B} - N \end{cases} . \tag{2}$$

It can be seen that latitude B and altitude H affect each other from the above equation, so B and H should be calculated through iterative methods. When the absolute value of the gap between the last two calculated results is less than a very small value, the iteration ends.

At high latitudes area or high altitudes area, the iterative formula above will be unstable[1], so B and H should be calculated following the formulas below:

$$\begin{cases} \cot B = \frac{\sqrt{X^2 + Y^2}}{Z + \Delta Z} \\ H = \sqrt{X^2 + Y^2 + (Z + \Delta Z)^2} - N \end{cases} . \tag{3}$$

Where:

$$\Delta Z = OE = |AE| \sin B = e^2 N \sin B = \frac{ae^2 \sin B}{\sqrt{1 - e^2 \sin^2 B}} .$$

2.2 Transformations among Cartesian Coordinates with Common Point of Origin

As shown in Fig. 2, there are two Cartesian coordinates, respectively $O - X_a Y_a Z_a$ and $O - X_b Y_b Z_b$. They have common point of origin, so the one of them which only needs to rotate successively at most three times can coincide with the another[2][4].

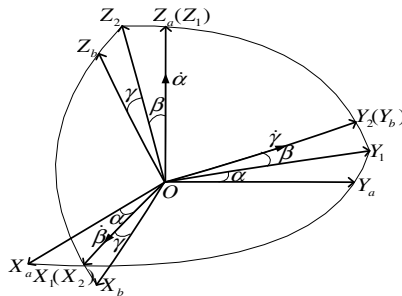


Fig. 2. Cartesian Coordinates with Common Point of Origin

Assuming that the transformation is from coordinate system $O-X_aY_aZ_a$ to $O-X_bY_bZ_b$. Firstly, rotating α anticlockwise around Z_a -axis (from the positive direction of the axis to the original point of view) can get $X_1Y_1Z_1$, and then rotating β anticlockwise around X_1 -axis can get $X_2Y_2Z_2$, finally, rotating γ anticlockwise around Y_2 -axis can get $X_bY_bZ_b$. Then

$$\begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix} = \begin{bmatrix} \cos \gamma & 0 & -\sin \gamma \\ 0 & 1 & 0 \\ \sin \gamma & 0 & \cos \gamma \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \beta & \sin \beta \\ 0 & -\sin \beta & \cos \beta \end{bmatrix} \begin{bmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_a \\ Y_a \\ Z_a \end{bmatrix}. \tag{4}$$

There are many Cartesian coordinate transformations with common point of origin in satellite navigation system. Such as transformation between local geographical coordinate system and carrier coordinate system, transformation between satellite orbit coordinate system and ECEF Cartesian coordinate system. The most commonly used transform angles of the former are ψ : yaw angle, θ : pitch angle and γ : roll angle, and the most commonly used transform angles of the latter are ω : argument of perigee, i : orbital inclination and Ω : right ascension of ascending node.

2.3 Transformation between ECEF Cartesian Coordinate and Local Geographical Coordinate with Uncommon Point of Origin

The local geographical coordinate is used to describe the north, east and vertical direction of a carrier’s position. According to the differences of the axis positive point orientation, there are some different selections of the coordinate system such as north-zenith-east (NZE), east-north-zenith (ENZ) and north-west-zenith (NWZ). This paper takes NZE coordinate system for example below. The establishment of local geographical coordinate system is changing with the movement of a carrier, so the real-time position of the carrier should be known before coordinate transformation. As shown in Fig. 3, $P(B, L, H)$ is the real-time position of the carrier, then the local geographical coordinate system can be established like the Figure shows.

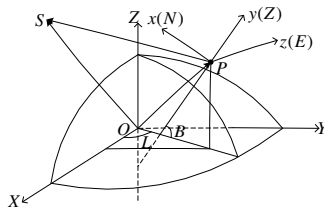


Fig. 3. ECEF Cartesian Coordinate System and Local Geographical Coordinate System

There are two different kinds of transformation commonly used in this situation.

(1) Transformation of Velocity

Because velocity is a vector, translation can’t change the coordinate. Assuming velocity of the carrier in ECEF coordinate system is $\vec{V}_{ECEF} = [V_X, V_Y, V_Z]^T$, $\vec{V}_{NZE} = [V_N, V_Z, V_E]^T$ is the corresponding velocity in NZE coordinate system. Then

$$\begin{bmatrix} V_N \\ V_Z \\ V_E \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos B & \sin B \\ 0 & -\sin B & \cos B \end{bmatrix} \begin{bmatrix} \sin L & -\cos L & 0 \\ \cos L & \sin L & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix}. \tag{5}$$

(2) Transformation of Position

In this condition, translation of the origin should be considered[2]. The ECEF Cartesian coordinate can be calculated according to Formula (1), assuming it is $P(X_p, Y_p, Z_p)$. Supposing that the position of a point S in ECEF is $S(X_s, Y_s, Z_s)$, $S(X_{sN}, Y_{sZ}, Z_{sE})$ is the corresponding position in NZE coordinate system. Then

$$\begin{bmatrix} X_{sN} \\ Y_{sZ} \\ Z_{sE} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos B & \sin B \\ 0 & -\sin B & \cos B \end{bmatrix} \begin{bmatrix} \sin L & -\cos L & 0 \\ \cos L & \sin L & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_s - X_p \\ Y_s - Y_p \\ Z_s - Z_p \end{bmatrix}. \tag{6}$$

3 Coordinates Transformation among Different Satellite Navigation Systems

Different navigation systems have different earth reference models respectively, so they differ from each other in the origin, orientation of coordinate axis and scale[5][6]. One of the most commonly used methods of transformation among different navigation systems is seven-parameters Bursa model which is described below:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} dX_0 \\ dY_0 \\ dZ_0 \end{bmatrix} + (1+dm) \begin{bmatrix} 1 & \beta_z & -\beta_y \\ -\beta_z & 1 & \beta_x \\ \beta_y & -\beta_x & 1 \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix}. \tag{7}$$

Where:

$[X \ Y \ Z]^T$ and $[X' \ Y' \ Z']^T$ represent the coordinate of a same point in the first system and the second system respectively; $[dX_0 \ dY_0 \ dZ_0]^T$ represents the origin’s projection coordinate of the second system in the first system; $\beta_x, \beta_y, \beta_z$ represents the transformation parameters of the two different coordinate systems; dm represents scale factor.

The accuracy of seven parameters selection determines the accuracy of the transformation, so it is very important to determine how to calculate the seven parameters. It can be seen that at least three exact points coordinates in the two different systems respectively should be known before using least square method to calculate seven parameters. After a simple linear process and ignoring the micro amount, Bursa model can be expressed as follows[3]:

$$\begin{bmatrix} X - X' \\ Y - Y' \\ Z - Z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X' & 0 & -Z' & Y' \\ 0 & 1 & 0 & Y' & Z' & 0 & -X' \\ 0 & 0 & 1 & Z' & -Y' & X' & 0 \end{bmatrix} \begin{bmatrix} dX_0 \\ dY_0 \\ dZ_0 \\ dm \\ \beta_x \\ \beta_y \\ \beta_z \end{bmatrix} \quad (8)$$

Supposing that (X_i, Y_i, Z_i) and (X_i', Y_i', Z_i') , $i = 1, 2, 3, \dots, n (n \geq 3)$ are corresponding coordinates of n points in the two different coordinate systems. Then

$$\underbrace{\begin{bmatrix} X_1 - X_1' \\ Y_1 - Y_1' \\ Z_1 - Z_1' \\ X_2 - X_2' \\ Y_2 - Y_2' \\ Z_2 - Z_2' \\ \dots \\ X_n - X_n' \\ Y_n - Y_n' \\ Z_n - Z_n' \end{bmatrix}}_P = \underbrace{\begin{bmatrix} 1 & 0 & 0 & X_1' & 0 & -Z_1' & Y_1' \\ 0 & 1 & 0 & Y_1' & Z_1' & 0 & -X_1' \\ 0 & 0 & 1 & Z_1' & -Y_1' & X_1' & 0 \\ 1 & 0 & 0 & X_2' & 0 & -Z_2' & -Y_2' \\ 0 & 1 & 0 & Y_2' & Z_2' & 0 & -X_2' \\ 0 & 0 & 1 & Z_2' & -Y_2' & X_2' & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & X_n' & 0 & -Z_n' & Y_n' \\ 0 & 1 & 0 & Y_n' & Z_n' & 0 & -X_n' \\ 0 & 0 & 1 & Z_n' & -Y_n' & X_n' & 0 \end{bmatrix}}_H \underbrace{\begin{bmatrix} dX_0 \\ dY_0 \\ dZ_0 \\ dm \\ \beta_x \\ \beta_y \\ \beta_z \end{bmatrix}}_D \quad (9)$$

That is $P = HD$, this is an overdetermined equations. The following cited theorem is picked up from a classical material: *If $A \in C^{m \times n}$, $b \in C^{m \times 1}$, then $x = A^+b$ is the best least squares solution of $Ax = b$. Where A^+ is the pseudo-inverse matrix of A , $A^+ = (A^T A)^{-1} A^T$ [7]. So:*

$$D = (H^T H)^{-1} H^T P \quad (10)$$

is the best least squares solution for formula (9).



The research institutes in United States, Russia and Germany got the transformation parameters between PZ-90 coordinate system and WGS-84 coordinate system by experiment in local range, however, because of the differences among experimental sites and the observation accuracy, the parameters which they got have a bit different. Russian MCC (Russian Mission Control Center) got the transformation parameters between PZ-90 and IRTF-94 which they thought is the most exact in 1997. International Earth Rotation Service agency gave the transformation parameters between WGS-84 and ITRF-94 in annual report in 1995. The gap between ITRF-96 coordinate system which GALELEO system uses and ITRF-94 is centimeter or less. So the gap could be negligible for the vast majority of users. The detailed transformation parameters are shown in Table 2.

Table 2. Transformation Parameters among Different Navigation Systems

Transformation Parameters	Transformation systems		
	PZ-90==>WGS-84	PZ-90==>ITRF-94	WGS-84==>ITRF-94
dX_0 (m)	-0.47	-0.49	-0.02
dY_0 (m)	-0.51	-0.50	0.01
dZ_0 (m)	-1.56	-1.57	-0.01
dm	22×10^{-9}	31×10^{-9}	-17×10^{-9}
β_x	0.076×10^{-6}	0.091×10^{-6}	0.015×10^{-6}
β_y	0.017×10^{-6}	0.020×10^{-6}	0.003×10^{-6}
β_z	-1.728×10^{-6}	-1.745×10^{-6}	0

There are still no information which shows that relevant organizations have measured the other transformation parameters among the different navigation systems, however, the parameters can be measured with the methods which were introduced in above paragraphs.

4 Transformation Process among Coordinate Systems

If the position and velocity of a carrier in a satellite navigation system have been known, other relevant information in this system and the corresponding information in another navigation system can be solved quickly by the methods which are introduced in above paragraphs. The transformation process of common coordinate systems in satellite navigation system is shown in Fig. 4 in which the content of  is input parameters, the content of  is output parameters.

Firstly, at least three points whose geodetic coordinates in two satellite navigation systems should be measured, and then the Cartesian coordinate of each point can be calculated according to Formula (1), we should notice that different navigation systems should select the corresponding reference ellipsoid parameters respectively when using Formula (1). Then Formula (10) can be used to calculate the best least squares solution of seven transformation parameters between the two navigation systems. If the location and velocity of a carrier in the first navigation system have been known, the same information in the second navigation system can be solved by Formula (7) combined with the seven parameters which have been solved. In addition, the geodetic coordinate of the carrier in the second navigation system can be found by Formula (2) or (3) based on the actual situation. After the position of carrier is determined, local geographic coordinate system can be set up, then the coordinate of a point in ECEF Cartesian coordinate system and local geographic coordinate system can transform to each other by Formula (6). If pitch angle, yaw angle and roll angle of carrier are known, carrier coordinate system can be established, and then coordinate of the point in carrier coordinate system can be solved by Formula (4). Meanwhile, the velocity of carrier can convert to each other between the two coordinate systems by Formula (5). If the orbit information (perigee,

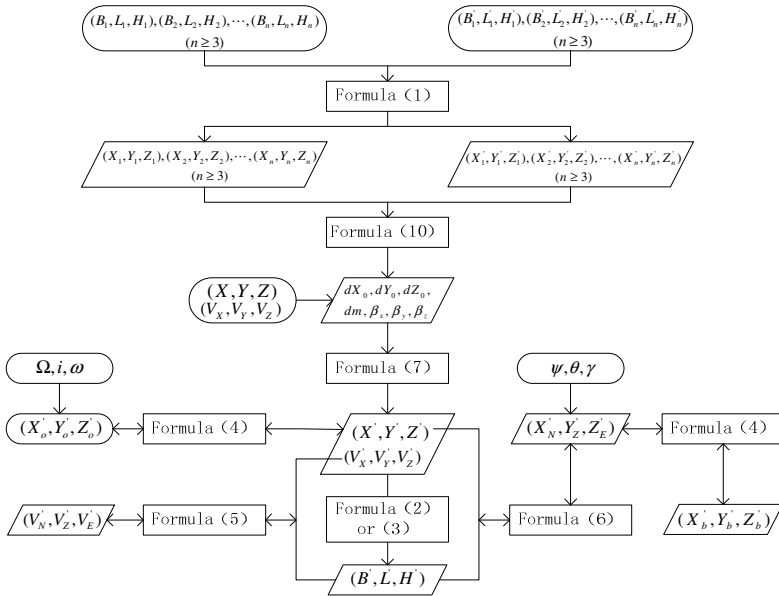


Fig. 4. Transformation Process among Common Coordinate Systems

inclination and right ascension of ascending node) of satellite are known, the orbit coordinate can convert to ECEF coordinate by Formula (4). Thus, the navigation data can be processed conveniently according to the transformation process which is shown in Fig. 4.

5 Conclusion

Common coordinate transformations in satellite navigation system were summarized by collecting and putting the relevant materials in order, and by the actual calculation and derivation. Different coordinate systems in the same satellite navigation system and among different satellite navigation systems were concluded in detail. The former includes transformation between Cartesian coordinate and geodetic coordinate in ECEF coordinate system, Cartesian coordinates which have common point of origin, and ECEF Cartesian coordinate system and local geographical coordinate system which have uncommon point of origin. For the latter, seven-parameters Bursa model which is one of the most commonly used transformation methods was introduced. Finally, the transformation process of common coordinate system in satellite navigation system was concluded so that the user can get the different forms of coordinate which they want conveniently. For most users, the transformation accuracy of the formulas can meet their requirements.

Acknowledgments. This work was supported by the National High-Tech. R&D Program, China (No.2011AA120505) and the National Natural Science Foundation, China (No.61173077).

References

1. Wang, A.: GNSS Measurement Data Processing. China University of Mining and Technology Press (2010)
2. Misra, P., Enge, P.: Global Positioning System, Signals, Measurements, and Performance, 2nd edn. Artech House Publisher (2006)
3. Liu, Q., et al.: Time Transformation and Coordinates Transformation among GPS, GLONASS and GALILEO. *Journal of Science of Surveying and Mapping* 33(5), 13–15 (2008)
4. Kaplan, E.D., Hegarty, C.J.: Understanding GPS: Principles and Applications, 2nd edn. Artech House Publisher (2006)
5. Lei, W., Xu, K., Zheng, H.: Research on the Coordinate Transformation Methods between National System and Local Independent System. In: Fourth International Conference on Intelligent Computation Technology and Automation, pp. 895–898 (2011)
6. Liu, X., Song, Y., Li, H.: Research on Design Methods of Coordinate Transformation and Navigation in Virtual Maintains System. In: Sixth International Conference on Natural Computation, pp. 4224–4228 (2011)
7. Shi, R., Wei, F.: Matrix Analysis, 2nd edn. Beijing Institute of Technology, Beijing (2005)

The Verification Method Research of Mining Right Boundary Based on GIS Technology in Shanxi Province

Bei Jia, Huijuan Li, and Baoping Wang

Taiyuan University of Technology, 030024 Taiyuan, China

1023409935@qq.com

Abstract. The national field verification of mining rights realized the unified within a system of the inflection point coordinate data mining, which is in 1980 xi'an coordinate system and 1985 national elevation benchmarks, comprehensively checked the status of the national mining right, rammed the data base of a mine management politics. In this paper, combining with experiences of field verification of mining rights in Shanxi province, To the having-drawn graphics, the graphic visual comparison, coordinate difference value comparison method and area comparison were used to compare and analysis, verified the mining right boundary, checked whether the ore bounded inflection point coordinates measurement and the provisions of the state coordinates were conformed, and evaluated these methods.

Keywords: mining rights ore bounded coordinate difference value comparison method area comparison.

1 Introduction

The national mining right field verification is an important investigation of mineral resource field, Its purpose is to check our country mining right status, Grasp the actual range of mining right person's exploration and development, understand the mining right's distribution state and rules, correct the problems found in check, update registration database of exploration and mining, make the mining right management level get a bigger ascension[1].The cross-border of Ore bounded include the cross of plane and elevation, This check is mainly aimed at flat cross-border situation, that is analysed and compared to check whether some areas of basic data mining right exist a deviation or error, whether part of the mining right's actual space scope and registration position is consistent, whether there is crossing and overlapping phenomena between the mining rights.

2 The General Situation of Research Area

ShanXi Province located in the east coast of middle Yellow River, east of loess plateau, geographic coordinates range is $34^{\circ} 34' N - 40^{\circ} N$, $110^{\circ} 43' 14'' E - 114^{\circ} 33' E$, the total area is 15.6 million square kilometers, covers 1.6% of the national total area, state profile were slightly oblique southwest of the northeast parallelogram. Shanxi Province

is an important energy base in China, has rich mineral resources, Its distribution from north to south is Datong、 Ningwu、 Xishan、 Qinshui、 Huoxi、 Hedong, called six coalfield muddy source, and other coal origin such as Hunyuan and Wutai,the coal-bearing area is 6.2 million square kilometers, covers 39.6% of the province total area. Many important deposit geological structure and hydrogeology conditions are relatively simple, appropriate open cast.

3 Data Sources and Processing

3.1 Data Sources

From shanxi SMC Geographical information center purchase C + GPS control points meeting the requirements of surveyed area and state of triangle dots remember above third-class. According to dots remember field reconnoiter to look for the point of control, sort out the national points appropriating GPS control network layout, and add control points in necessary area according to the general layout nets request, to ensure that the overall control nets are rationality scientific.

Field reconnoiter all layout net's points, spotting well-preserved and firm, and ensure that can use. Each point is collected system results of 1954 Beijing coordinate system and 1980 xi'an coordinate system[2], the plane head control network counted data is completed with the change after conversion, use GPS fitting method to determine the elevation control, fitting counted are no less than four points.

3.2 Data Consolidated

Coordinate transformation is mainly convert Beijing 1954 coordinates system to Xi'an 1980 coordinate system[3]. This conversion is a kind of ellipsoid parameter conversions, due to the difference between ellipsoid benchmark, use seven parameters Boolean Sally model to perform the relatively strict transformation. Transformation method is to select 3 ~ 5 uniform basic points in range 500 km² ~ 1000 km², collect two sets coordinates of these points, use HDS2003 data processing software decoding out coordinate transformation seven parameters between two reference ellipsoid, and use the seven calculated parameters directly transform inflection point coordinate of mineral rights. If the area is not big, use four parameters method for conversion.

Most coals take the division conversion method in Shanxi coal mining right field verification, that is one area correspond a set of transformation parameters, using seven parameters with two-dimensional coordinate transformation conversion model to realize. There are also some areas which don't use conversion parameters or use two sets of parameters, duo ro different transformation parameters of each area, and some ores in a area the parameters are different. so, only by using correct switching model and conversion parameters, to ensure the accuracy of converted coordinates and database achievements precision.

After all point coordinates are transformed into 1980 xi'an coordinates, draw on graphics, the rendering steps are as follows:

- (1) According to the exploration license approved exploration blocks Ed figure range, determ to exhibit inflexion point coordinates and use line connection to form exploration boundary in the picture.

- (2) The geographical elements exhibition. Main geographical factors draw according to the test result show, paint by way of turn drawing for large scale topographic maps.
- (3) Exhibit measured mineral exploration right exploration engineering position in the picture, use symbols or connect point into line with physical display, entity fill within designs or color according to illustrations' requirements.
- (4) According to the map's range, turn draw neighboring exploration, geographical elements and other related contents.
- (5) Edit these elements, such as TuMing, scale, exploration Numbers, coordinates nets, illustrations, compass, drawing frame, responsibility column.

Plotted AutoCAD figure, through the middle data format DXF realize MapGis data conversion, After completion, open conversion of dot, line files in the input editing system, If some lines are not showing, probably because the cad-map.clr document does not give corresponding color number data in the color of AutoCAD and MapGIS system, when convert ,system defaults it into number9 color, change the system into other colors, it can be displayed. Another reason is that, Some layers themselves have problems, transformation are not finished, When transforming, remove some layers can complete graphical transformation. After generated MapGIS diagram, use ministry data acceptance procedures to inspect it, submit results, specific operation flow as shown in figure 1.

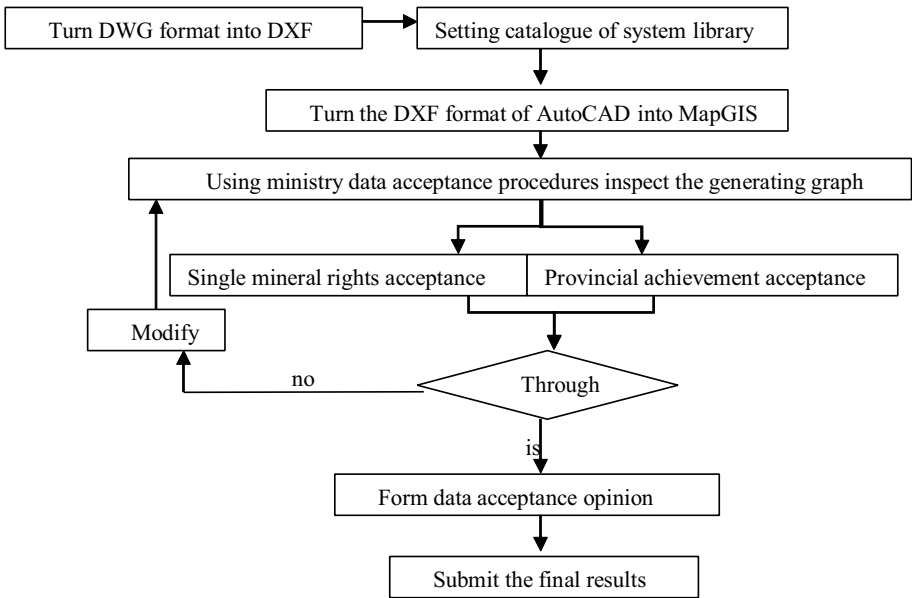


Fig. 1. Graphics operation flow chart

4 The Ore Bounded Comparison in Inspectors

This paper discusses 3 methods to compare ore bounded border, including the direct comparison through graphics;through calculate the difference coordinates in value,comparison the precise data;through calculate area, account the area.

4.1 Graphical Comparison

Take the known measured point coordinates and state regulations point coordinates into ArcGIS respectively in the form of the coordinates of the text, automatically achieve the transformation from point to area, the generation areas are in two different layers, display in two different color. Figure 2 is an inflection point coordinate offset situation of an open mining, through comparison of graphic interface, you can clearly see the beyond situation of the borders, there are some differences between the provisions of the state mining right boundary and the meseased ore bounded. Three reasons lead to appear this kind of circumstance,first, the scientific method is not the delimitation of homework,second, Grassroots land management department lack of necessary technical conditions,third, Ore administration management system exist problems.

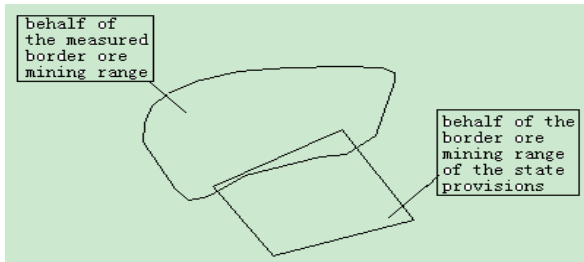


Fig. 2. The beyond situation of Ore bounded boundary

4.2 Coordinates Comparison

In ArcGIS, take point coordinates into ArcMAP with Excel format,and reexport it, shapfile format file is generated,add XX,YY to one of the two attributes, XX is used to store X coordinate's offset, YY is used to store Y coordinate's offset, use join field to connect two tables,after joining, use Field Calculator function to directly subtract two groups of point coordinates, using this method know the beyond situation of the border, realized the comparison of two graphic coordinate offset situation,namely, compared with the original graphic X, Y, every feature points' offset, as shown in table 1.

Table 1. The contrasting table of Ore bounded coordinates

	X1(m)	Y1 (m)	X2(m)	Y2(m)	$\Delta X(m)$	$\Delta Y(m)$
1	xxx102.132	xxx8590.843	xxx193.349	xxx8565.105	91.217	-25.738
	xxx259.142	xxx8651.869	xxx257.607	xxx8508.251	-1.535	-143.618
	xxx224.447	xxx8504.912	xxx241.408	xxx8557.595	16.961	52.683
	xxx120.744	xxx8484.206	xxx202.822	xxx8511.744	82.078	27.538
2	xxx567.538	xxx7799.503	xxx545.659	xxx7816.069	-21.879	16.566
	xxx586.565	xxx7773.956	xxx575.491	xxx7842.387	-11.074	68.431
	xxx631.013	xxx7787.080	xxx625.061	xxx7838.999	-5.952	51.919
	xxx608.053	xxx7821.106	xxx618.971	xxx7818.808	10.918	-2.298
3	xxx291.151	xxx0134.286	xxx303.152	xxx0134.285	12.001	-0.001
	xxx413.248	xxx0041.366	xxx413.249	xxx0041.365	0.001	-0.001
	xxx330.817	xxx9965.954	xxx330.817	xxx9965.954	0	0
	xxx229.008	xxx0036.220	xxx229.008	xxx0036.220	0	0
4	xxx796.867	xxx3087.195	xxx867.134	xxx3114.069	70.267	26.874
	xxx792.443	xxx3105.343	xxx856.942	xxx3111.279	64.499	5.936
	xxx894.906	xxx3150.101	xxx875.784	xxx3165.393	-19.122	15.292
	xxx924.838	xxx3116.136	xxx873.216	xxx3128.305	-51.622	12.169
5	xxx941.926	xxx2829.097	xxx765.993	xxx2876.556	-175.933	47.459
	xxx056.978	xxx2766.412	xxx712.308	xxx2848.027	-344.67	81.615
	xxx032.497	xxx2720.172	xxx765.788	xxx2870.340	-266.709	150.168
	xxx917.445	xxx2782.857	xxx747.879	xxx2882.074	-169.566	99.217

Among them, X_i , Y_i behalf of the state of control points coordinates, X_2 , Y_2 behalf of the actual measurement of the control points of coordinates, ΔX , ΔY behalf of coordinates deviation.

According to the measured precision control provisions of the state[4], some open-pit mine deviation in county are in permitted range, do not belong to mistake, such deviation can be acceptable. Other mines' coordinates are seriously out of range, that is the act of cross-border exploitation.

4.3 Area Accounting

Calculating the beyond area, Each point of a polygon coordinates is known, according to the area formulas to account area. Through the method of calculating border area to compare with accorded degree.

$$\text{Area} = \frac{1}{2} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ \cdot & \cdot \\ \cdot & \cdot \\ x_n & y_n \\ x_1 & y_1 \end{vmatrix} = \frac{1}{2} [(x_1y_2 + x_2y_3 + x_3y_4 + \dots + x_ny_1) - (y_1x_2 + y_2x_3 + y_3x_4 + \dots + y_nx_1)] \quad (1)$$

Take number 3 open-air ore as an example, ore sector as prescribed by the state area calculation process are as follows.

$$A_1 = \frac{1}{2} [(XXX\ 291.151 \times XXX\ 0041.366 + XXX\ 413.248 \times XXX\ 9965.954 + XXX\ 330.817 \times XXX\ 0036.220 + XXX\ 229.008 \times XXX\ 0134.286) - (XXX\ 0134.286 \times XXX\ 413.248 + XXX\ 0041.366 \times XXX\ 330.817 + XXX\ 9965.954 \times XXX\ 229.008 + XXX\ 0036.220 \times XXX\ 291.151)] = 15608.804$$

Similarly, other regions' area can be calculated. Table 2 is through the coordinate calculation to compare the area difference in value.

Table 2. The areas contrasting table of Ore bounded

	$A_1(m^2)$	$A_2(m^2)$	$\Delta A(m^2)$
1	16200.258	3341.334	12858.924
2	1629.853	2645.408	-1015.554
3	15608.804	15577.904	30.9
4	3635.055	1401.402	2233.653
5	6854.509	1738.886	5115.623

Among them, A_1 behalf of the calculation graphics area by the country Control coordinate, A_2 behalf of calculation graphics area by the actual measurement control point coordinates, ΔA behalf of Area deviation.

Through the above area comparison ,we can see that, according to the demand of precision, the absolute value of some ore area deviation do not exceed limit,therefore, Although the measured data of the border area and regulations ore mineral bounded area is not completely the same, these ores are not cross mining. But there is still part of the open pit mining have the behavior of non-compliance, it is need to be warned, fined, or even ban or other penalties.

5 Conclusion and Forward

Through comparison of three kinds of methods above, we can see that the three kinds of comparative method each have advantages and disadvantages. Graphical

comparison are intuitive, but persuasive weakly; Coordinates comparison through data precise compared, and the comparison with the accuracy of provisions, they are persuasive, but look more obscure, lack of integrity; Comparative area can be clearly glanced, clearly described area, but do not see specific offset circumstances, such as specifically a point what direction is to slant. Combined these three methods can achieve very good comparative effect. Through the above methods compare in Shanxi Province mining right boundary, that provides some methods for boundary comparison. Of course, there are other methods to compare the border, for example, the distance can be compared and analyzed to understand clearly whether overrun the length; Through angle comparison, can realize the purpose of the azimuth comparison.

References

1. Tan, Y., Chang, Y.-G.: The national mining right on-the-spot verification work instructions and technical requirements, p. 1. China Continent Press, Beijing (2008)
2. Homeland ZiYuanTing in Shanxi Province. Shanxi mining right field inspectors implementation plan. Shanxi Coal Geology Exploration Surveying and Mapping Court, 17 (2009) (in Chinese)
3. Chen, R.-C., Wang, F.-J.: Primary Study on Key Technologies Field Verification of Mining Rights. Shandong Land and Resources 27(1), 1 (2011) (in Chinese)
4. The national mining right field review program office. The national mining right on-the-spot verification technical requirements supplement and the detailed rules for the implementation of the data standardization arrangement, p. 2 (2009)

Research on Evaluation Model of Digital Library Service Quality

YanHui Liu, DaLi Xie, and XiaoBo Wang

School of Information Management, Heilongjiang University, Harbin 150080, China
liuyanhu1987000@126.com

Abstract. Based on the fuzzy concept with service quality evaluation of the digital library, considering its fuzzy nature, multi-level, and many indexes, the paper uses AHP to identify weights, and uses fuzzy comprehensive evaluation method to develop a set of digital library service quality evaluation model. According to the principle of evaluation, starting from the resources of digital library, the paper analyzed its impact factors, established system model and combined with examples to show the model's practicality.

Keywords: digital library, the service quality, Fuzzy evaluation method, AHP.

In the age of information explosion, the quality of service has become the core of all walks of life, while digital library as a facility to improve information service also has its advantages and disadvantages. This article starts with the construction of digital library resources, establishing an evaluation model to analyze service quality of digital Library.

1 The Index System of Digital Library

LibQUAL⁺ is a new method[1] proposed and applied by ARL (Association of Research Libraries, abbr. ARL) in 1999 to evaluate service quality of digital library. Based on LibQUAL⁺ evaluation model, the paper examine from the construction of digital library resources, in which resources are divided into collections of resources and building facilities and resources; collection of resources are divided into information content and interface; facilities resources are divided into equipment and human resources.

1.1 Establishing Hierarchical Model of Digital Libraries

Based on the above discussions on the factors impacting the digital library services, combined with AHP, this digital library service quality evaluation model is divided into three levels (Figure 1), the top hierarchical level as the destination layer: Digital Library Museum resources and the environment; the second layer as the criterion level: broken down into four indicators including content, interfaces, electronic devices, and human resources; the third level as an indicator layer: the impact of specific elements, or a total of nine evaluation indicators.

1.1.1 Establishing AHP Model to Determine the Matrix Structure

According to the established hierarchy model (Figure 1), by comparing the importance of various factors, a judgment matrix is constructed. Any judgment matrix should meet, $X_{ii} = 1, X_{ij} = 1/X_{ji}$ ($i, j = 1, 2, \dots, n$). Here the author use ratio scale method by ALSarry to determine the index weight, it is a statistical method proposed in AHP.

1.1.2 Evaluation Indicator to Determine the Weight

Based on the above established digital library evaluation index system model, and according to the importance of each factor, a judgment matrix was established to calculate levels, and the weight of each resulted level relative to its upper level is showed in Table 1, Table 2, Table 3, Table 4, and table 5:

Table 1. Four-leveled resources and environment judgment matrix A-C and its relative weight

indicator	Information content	Interface	Electronic device	Human resources	Weight W_c
Information content	1	3	2	8	0.516
Interface	1/3	1	2	2	0.210
Electronic device	1/2	1/2	1	5	0.207
Human resources	1/8	1/2	1/5	1	0.065

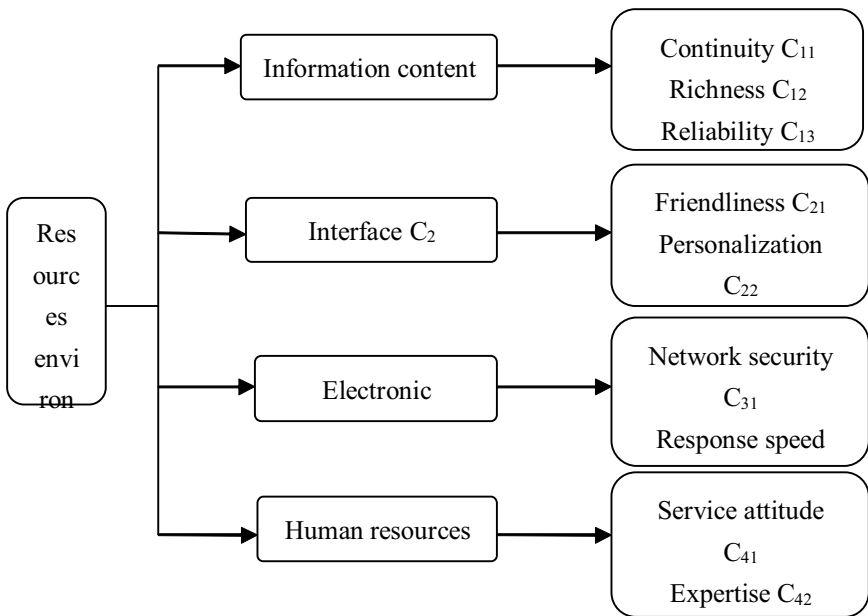


Fig. 1. Digital Library Evaluation System Model

Table 2. Leveled information content judgment matrix C-C1 and its relative weight vector

Indicator	continuity	Richness	Reliability	Weight Wc1
continuity	1	3	5	0.636
richness	1/3	1	3	0.258
Reliability	1/5	1/3	1	0.104

Table 3. Leveled interface judgment matrix C—C2 and its relative weight vector

Indicator	Friendliness	Personalization	Convenience	Weight Wc2
friendliness	1	2	3	0.539
Personalization	1/2	1	2	0.296
convenience	1/3	1/2	1	0.163

Table 4. Leveled electronic device judgment matrix C—C3 and its relative weight vector

Indicator	Network security	Response speed	System stability	weight Wc3
Network security	1	5	7	0.730
Response speed	1/5	1	3	0.188
System stability	1/7	1/3	1	0.080

Table 5. Leveled human resources judgment matrix C-C4 and its relative weight vector

Indicator	Service attitude	Expertise	Work efficiency	Weight Wc4
Service attitude	1	1/7	1/2	0.090
Expertise	7	1	6	0.758
Work efficiency	2	1/6	1	0.151

1.1.3 Consistency Test to the Judgment Matrix

In order to test the consistency of the judgment matrix, its consistency index CI was compared with the average random consistency index RI, table 6 showed the value of the average random consistency index RI.

Table 6. The value of the average random consistency index RI

Matrix dimension	1	2	3	4	5
RI	0	0	0.52	0.89	1.12
Matrix dimension	6	7	8	9	10
RI	1.25	1.35	1.42	1.46	1.49

The ratio of CI and RI is referred to as the consistency ratio CR, the $CR = CI / RI$, when $CR < 0.1$, then the judgment matrix has satisfactory consistency, otherwise the matrix would need to be adjusted. The consistency ratio CR of each judgment matrix obtained in this paper was less than 0.1 (see Table 7), met consistency requirements; hence shows that the logical relationship between the various elements in the matrix is rational, and the results by its calculation can be used for service quality evaluation.

Table 7. The consistency test results of the judgment matrix

The judgment matrix level	λ_{max}	CI	RI	CR
A as resources environment	4.176	0.058	0.89	0.066
C1 as information content	3.038	0.019	0.52	0.037
C2 as interface	3.009	0.004	0.52	0.008
C3 as electronic device	3.064	0.032	0.52	0.062
C4 as human resources	3.032	0.016	0.52	0.031

2 Case Study on the Digital Library Service Quality Fuzzy Comprehensive Evaluation

This paper applied fuzzy comprehensive evaluation method into the digital library service quality assessment. This method comes from Chad an American, who created in 1965 a method to describe the fuzzy phenomenon in math - fuzzy set theory [2]. The results from approaching fuzzy problem with fuzzy means will make the evaluation more realistic and accurate.

2.1 The Establishment of Indicator System

The set factor U to determine the quality of digital library services has two levels of factors, the first level of factors, $U = \{U_1, U_2, U_3, U_4\}$, in it the first-level factor U_i is composed of several second-level factor U_j .

2.2 The Establishment of Evaluation Set of Each Element

In this model, the evaluation structure of each element has four levels, namely excellent, good, fair and poor. Evaluation set characterizes the results from a comprehensive evaluation in the form of membership grade. So excellent, good, fair and poor are showed as:

$$V = \{0.05, 0.04, 0.03, 0.02, 0.01\}$$

2.3 Constructing Judgment Matrix

This paper applied the expert assessment method to determine the evaluation index weights; the judgment matrix was showed as follows:

$$R_{C1} = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.2 \end{bmatrix} \quad R_{C2} = \begin{bmatrix} 0.2 & 0.3 & 0.3 \\ 0.4 & 0.3 & 0.3 \\ 0.3 & 0.2 & 0.4 \end{bmatrix}$$

$$R_{C3} = \begin{bmatrix} 0.4 & 0.2 & 0.2 \\ 0.4 & 0.3 & 0.3 \\ 0.4 & 0.2 & 0.3 \end{bmatrix} \quad R_{C4} = \begin{bmatrix} 0.3 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.3 \\ 0.3 & 0.2 & 0.3 \end{bmatrix}$$

2.4 The Matrix to Calculate Evaluation Results

Two evaluation results have to be through the form of fuzzy transformation for synthetic operations.

$$C_1 = W_{C1} \circ R_{C1}$$

In it, \circ signifies fuzzy synthetic operation, generally use the max min algorithm, but this synthetic method has many problems [3-4], this paper applied (\bullet, \otimes) , a fuzzy relation synthetic method [5], relative to the ordinary matrix product..

$$C_1 = W_{C1} \circ R_{C1} = \{0.636 \quad 0.258 \quad 0.104\} \times \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.2 \end{bmatrix}$$

$$= \{0.064 \quad 0.129 \quad 0.072\},$$

in the same way,

$$C_2 = \{0.097, 0.060, 0.120\}$$

$$C_3 = \{0.109, 0.036, 0.054\}$$

$$C_4 = \{0.139, 0.061, 0.117\}$$

The judgment matrix R_A for the objective level A can be composed of C_1, C_2, C_3, C_4 , in this way,

$$A = W_A \circ R_A = \{0.516 \quad 0.210 \quad 0.207 \quad 0.065\} \\ \times \begin{bmatrix} 0.064 & 0.129 & 0.072 \\ 0.097 & 0.060 & 0.120 \\ 0.109 & 0.036 & 0.054 \\ 0.139 & 0.061 & 0.117 \end{bmatrix} \\ = \{0.013 \quad 0.045 \quad 0.009\}$$

In accordance with the principle of maximum membership degree, we can see the level of 0.045 corresponds to the excellent and good in between, that is, the digital library service quality assessment rates good.

3 Conclusion

This paper applied multi-level fuzzy comprehensive evaluation method. This method is based on the effective combination of qualitative and quantitative these two points of view, to make the digital library service quality assessment. The consistency test proves the evaluation model established in this article is objective, practical, and through it some weaknesses in the construction of digital libraries can be found.

Acknowledgments. National social science fund project funding (10 CGL076), heilongjiang natural science fund project assistance (G201020).

References

1. Fan, X.-H., Yuan, Y., Li, Q.-R.: The library service quality evaluation system 88, 46–54 (2009)
2. He, Z.-X.: The fuzzy mathematics and its application. Science and Technology Press, Tianjin (1983)
3. Lu, H.-Q., Wang, N.-S., Shen, F.-H.: About fuzzy comprehensive evaluation take big in small the discussion of the algorithm 114, 124–128 (2001)
4. Pang, Y.-J., Liu, K.-D., Liu, J.: In fuzzy mathematics "take big in small" operation caused problem. The System Engineering Theory and Practice, 98–100 (2001)
5. Cao, X.-S., Cai, L.-G., Li, P.-G.: Many process program second order fuzzy comprehensive evaluation. The System Engineering Theory and Practice (2000)

Bayesian Model Averaging for Lasso Using Regularization Path

Mei Wang¹ and Erlong Yang²

¹ School of Computer and Information Technology,
Northeast Petroleum University, 163318 Daqing, China

² Key Laboratory of Enhanced Oil Recovery in Ministry of Education,
Northeast Petroleum University, 163318 Daqing, China
meiwang.nepu@gmail.com

Abstract. We consider the problem of accounting for model uncertainty in LASSO models. Bayesian model averaging (BMA) provides a coherent mechanism for accounting for this model uncertainty. In this paper, we propose a method for implementing Bayesian model averaging for LASSO based on regularization path. First, we construct the initial model set using the regularization path, whose inherent piecewise linearity makes the construction easy and effective. Then, we elaborately select the models for BMA from the initial model set through the Occam's Window method. Finally, we carry out the BMA on the selected models. Experimental results show that BMA has significant advantage over the model selection method based on Bayesian information criterion (BIC).

Keywords: LASSO, Bayesian model averaging, Regularization Path, Occam's Window.

1 Introduction

LASSO (least absolute shrinkage and selection operator) was introduced by [1] and has since been an attractive technique for regularization and variable selection and well studied [2, 3, 4]. Some reasons for the popularity might be that the entire regularization path of the LASSO can be computed efficiently [2], that LASSO is able to handle more predictor variables than samples and produces sparse models which are easy to interpret.

Consider the usual linear regression setting with data obtained as independently identically distributed (i.i.d.) pairs of feature vectors $\{\bar{x}_i\}_{i=1}^n$ and corresponding targets $\{y_i\}_{i=1}^n$, where the $\bar{x}_i = \{x_{i,1}, \dots, x_{i,p}\}$ is the regressor with p fixed predictor variables and y_i the response for the i th observation. In this situation, ordinary least squares regression finds the linear combination of the x s that minimises the residual sum of squares. However, if p is large or if the regressor variables are highly correlated, then the variances of the least-squares coefficient estimates may be

unacceptably high. [1] proposed the “least absolute shrinkage and selection operator” (LASSO), which minimises the residual sum of squares under a constraint on the L_1 - norm of coefficient vector. Thus the LASSO estimator solves the optimisation problem

$$\min_{\beta_1, \dots, \beta_p} \frac{1}{2} \sum_{i=1}^n (y_i - \sum_{j=1}^p x_{ij} \beta_j)^2 \text{ s.t. } \sum_{j=1}^p |\beta_j| \leq t, \tag{1}$$

for some $t > 0$. For smaller values of t , the LASSO shrinks the estimated coefficient vector towards the origin (in the L_1 sense), typically setting some of the coefficients equal to zero. Thus the LASSO combines characteristics of ridge regression and subset selection and promises to be a useful tool for variable selection.

In order to obtain good generalization performance of LASSO, it is necessary to choose a sufficiently good model. However, model selection intrinsically has the risk of uncertainty in real-world situations where every model has limitations and tends to make errors. Bayesian model averaging (BMA) is an alternative to overcome the limitations of model selection, which can integrate all useful information from the candidate models into the final hypothesis to guarantee good generalization performance.

In this paper, we focus the Bayesian model averaging for LASSO on regularization path. First, we construct the candidate model set according to the regularization path, whose inherent piecewise linearity makes the construction easy and effective. All possible models are involved in the initial model set. Then, we elaborately select the models for combination from the initial model set through the improved Occam’s Window method and the input-dependent strategy. Finally, we carry out the combination on the selected models using the Bayesian model averaging, in which model posterior probabilities are estimated by Bayesian information criterion (BIC) approximation.

2 LASSO Model Set for BMA

2.1 Model Set Based on Regularization Path

In these years, researchers have developed many algorithms for solving LASSO [2, 4]. We will follow [2] and get the all possible LASSO models for BMA.

The formulation (1) for LASSO can be written in a *loss + penalty* form [5]:

$$\min_{\beta_1, \dots, \beta_p} \frac{1}{2} \sum_{i=1}^n (y_i - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j|, \tag{2}$$

where $0 \leq \lambda \leq \infty$, for a given λ there exists a $t \geq 0$ such that the two problems share the same solution, and vice versa.

We write the model as

$$f = X\beta + \varepsilon.$$

The method proceed iteratively in a series of steps. We make the following definitions, and these are corresponding the model M_l at step l , $l = 1, \dots, k$:

- β_l : the vector of coefficients at (the end of) step l .
- A_l : the active set of predictors during step l .
- X_l : the columns of X corresponding to A_l .

Initially all coefficients β_0 are zero, and A_0 is empty. At the l th step, the active set is updated, and a new set of coefficients β_j are obtained by determining a step length λ_l along a direction d_l from β_{l-1} :

$$\beta_l \leftarrow \beta_{l-1} + \lambda_l d_l.$$

Then we can get all possible models which make the model set $\Upsilon = \{M_1, \dots, M_k\}$.

2.2 Occam’s Window

All possible LASSO models are involved in the model set Υ , including good performance ones and poor performance ones. We apply Occam’s Window [6] method to eliminate the poor performance models. In this method, for give dataset T , posterior model probability $\text{pr}(M | T)$ is used as a metric to guide model selection. There are two basic principles underlying this approach.

First, if a model M predicts the data far less well than the model which provides the best prediction, then it has been discredited and should no longer be considered. Thus models not belonging to

$$\Upsilon' = \{M_j : \frac{\max_l \{\text{pr}(M_l | T)\}}{\text{pr}(M_j | T)} \leq W\} \tag{3}$$

should be excluded from the combination candidate model set, where posterior probability ratio W is chosen by the data analyst and $\max_l \{\text{pr}(M_l | T)\}$ denotes the model in initial candidate model set Υ with the highest posterior model probability.

Second, appealing to Occam’s razor, we exclude models which receive less support from the data than any of their simpler submodels. Here we give the definition of submodel. If M_i is a submodel of M_j , we mean that all the support vectors involved in M_i are also in M_j . Thus we also exclude models belonging to

$$\Upsilon'' = \{M_j : \exists M_l \in \Upsilon, M_l \subset M_j, \frac{\text{pr}(M_l | T)}{\text{pr}(M_j | T)} > 1\} \tag{4}$$

then we obtain the model set

$$\Upsilon_{ao} = \Upsilon \setminus \Upsilon'' \subseteq \Upsilon.$$

The posterior probability ratio W is usually a constant as in [6]. However, the statistical results show that only few of LASSO models in Υ have strongly peak posterior probabilities. So, we apply a query-dependent method to determine the ratio W . Starting from $W = k/20$, we double it for every iteration and examine the number of models in set Υ'' . Once the number changes dramatically, we terminate the iteration process and use the last W value. In the experiments below, this would be the case with the model set size $|\Upsilon'|$ increasing more than 4. Here, if the model set size enlarges dramatically, it means that many models with low posterior probabilities enter the model set Υ' .

3 BMA Prediction

For a given dataset T , The Bayesian model averaging over Υ_{ao} has the form

$$\hat{f}_{bma}(x) = \sum_{j=1}^{k_{ao}} \hat{f}_j(x) \text{pr}(M_j | T) \tag{5}$$

where $\text{pr}(M_j | T)$ is the posterior probability of model M_j , $\hat{f}_j(x)$ is the prediction M_j of on x , $j = 1, \dots, k_{ao}$. Then this is the process of estimating the prediction under each model M_j and then averaging the estimates according to how likely each model is [7].

The posterior probability of model M_j in equation (5) can be estimated using BIC. Given the model set Υ_{ao} , we compute the BIC criterion for each model $M_j \in \Upsilon$, denoted by BIC_j , and estimate the posterior probability of each model M_j as

$$\widehat{\text{pr}}(M_j | T) = \frac{e^{-\frac{1}{2} \cdot \text{BIC}_j}}{\sum_{M_m \in \Upsilon_{ao}} e^{-\frac{1}{2} \cdot \text{BIC}_m}}. \tag{6}$$

4 Experiments

The In this section we compare the performance of our BMA with BIC-based model selection on simulation data. Simulations are based on the following three functions found in [8]:

(D1) One-dimensional $\sin(\bullet)$ function:

$$f(x) = \frac{\sin(\pi x)}{\pi x} + r_1,$$

where x is distributed as uniform $(-2,2)$, and the errors $r_1 = 0.19$.

(D2) Five-dimensional additive model:

$$f(x) = 0.1e^{4x_1} + \frac{4}{1 + e^{-20(x_2 - 0.5)}} + 3x_3 + 2x_4 + x_5 + r_2,$$

where x_1, \dots, x_5 are distributed as uniform $(0,1)$, and the errors $r_2 = 1$.

(D3) Alternating current series circuit:

a. Impedance

$$f(R, \omega, L, C) = (R^2 + (\omega L - \frac{1}{\omega C})^2)^{1/2} + r_3;$$

b. Phase angle

$$f(R, \omega, L, C) = \tan^{-1}(\frac{\omega L - \frac{1}{\omega C}}{R}) + r_4,$$

where (R, ω, L, C) are distributed uniformly in $(0.100) \times (40\pi, 560\pi) \times (0,1) \times (1,11)$, and the errors $r_3 = 218.5$ and $r_4 = 0.18$.

We generated 1000 training observations from each function along with 500 test observations. We repeat this process 50 times and compute the average prediction errors and their corresponding standard errors. Since the usual goal of regression analysis is to minimize the predicted squared-error loss, the prediction error is defined as

$$PE = \frac{1}{m} \sum_{i=1}^m (y_i - f(\bar{x}_i))^2, \tag{7}$$

where m is the number of the test data. We calculate the prediction error with each test data for each method.

Table 1. Prediction Error on two methods

DataSet	BIC	BMA	BMA _{ao}
D1	0.0554 (0.0082)	0.0571 (0.0078)	0.0514 (0.0063)
D2	0.8306 (0.1202)	0.8682 (0.1187)	0.7564 (0.1120)
D3(a)	81291 (7712)	85281 (7668)	78282 (7257)
D3(b)	0.0489 (0.0021)	0.0529 (0.0026)	0.0468 (0.0018)

The results are summarized in Table 1. In the table, *BIC* denotes using the model selected by BIC, *BMA* denote Bayesian model averaging on the initial model set \mathcal{Y} , and *BMA_{ao}* denote Bayesian model averaging on the refined model set \mathcal{Y}_{ao} . From the tables we find that the BMA on the initial model set have higher prediction error and lower standard errors than BIC-based model selection, while BMA on the refined model set have low prediction error and lower standard errors than BIC-based model selection.

5 Conclusion

This paper proposes a Bayesian model averaging method for LASSO, inspired by the the regularization path algorithm rising in recent years. One advantage of the regularization path of LASSO is piecewise linear, which facilitates model selection. We can obtain all possible models according to the regularization path, including the models that have good performance on training set and the ones that have poor performance. Applying the Occam's Window method, we greatly reduce the number of candidate models and improve the BMA prediction performance on test data.

Acknowledgments. This work is supported in part by Youth Science Foundation of China under Grant No. 51104030 and Heilongjiang Education Department of China under Grant No. 12511010.

References

1. Tibshirant, R.: Regression shrinkage and selection via the lasso. *J. R. Statist. Soc. B* 58(1), 267–288 (1996)
2. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R.: Least angle regression. *Annals of Statistics* 32, 407–499 (2004)
3. Zou, H., Hastie, T., Tibshirani, R.: On the "degrees of freedom" of the lasso. *Annals of Statistics* 35(5), 2173–2192 (2007)
4. Fraley, C., Hesterberg, T.: Least angle regression and LASSO for large datasets. *Statistical Analysis and Data Mining* 1(4), 251–259 (2009)
5. Liu, H., Zhang, J.: On the L1-Lq regularized regression. Technical report, Technical Report arXiv: 0802.1517 v1, Carnegie Mellon University (2008)
6. Madigan, D., Raftery, A.E.: Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association* 89(428), 1535–1546 (1994)
7. Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer (2008)
8. Friedman, J.H.: Multivariate adaptive regression splines. *The Annals of Statistics* 19(1), 1–67 (1991)

Optimization Design of Parameters with Hybrid Particle Swarm Optimization Algorithm in Multi-hole Extrusion Process

Wen-Jong Chen, Dyi-Cheng Chen, Wen-Cheng Su, and Fung-Ling Nian

Department of Industrial Education and Technology,
Changhua University of Education, No.2, Shi-Da Road,
Changhua City 500, Taiwan Wen-Jong Chen,
wjong@cc.ncue.edu.tw

Abstract. In this paper, we present a Hybrid Particle Swarm Optimization (HPSO) algorithm that combines the Particle Swarm Optimization (PSO) algorithm with the Genetic Algorithm (GA) method. The algorithm is used for investigating the plastic deformation behavior of titanium alloy (Ti-6Al-4V) in a multi-hole extrusion process. The simulation used rigid-plastic finite element (FE) DEFORMTM-3D software to obtain the minimum mandrel bias angle and exit tube bending angle. Results of the simulation indicate that these two angles were significantly less than 0.3 degrees, suggesting that metaheuristic algorithms based on HPSO and FE analysis could be used for general multi-hole extrusion processes.

Keywords: Multi-hole extrusion, Particle Swarm Optimization, Finite element, Genetic Algorithm.

1 Introduction

In recent years, extruded titanium has been widely used in various engineering fields. While the titanium alloy (Ti-6Al-4V) is a typical type of two-phase titanium alloy, it does possess characteristics of high economic value, such as its light weight, high strength, corrosion resistance, and high-temperature oxidation resistance [1]. As market demand and quality requirements increased, most of the industry began to favor the economically efficient multi-hole mold that can help improve production efficiency in the extrusion process. However, in the multi-hole extrusion molding process, billet plastic deformation in complex multi-hole molding processes can lead to excessive extrusion pressure on high-strength alloys, which can generate non-symmetrical material flow in the mandrel. The most important of these factors are hole location, billet temperature, extrusion speed, and friction. The objective of this study is to use the extrusion process to seek out the optimal parameter combination that could generate minimum bias in the molding process.

Numerous past studies have discussed extrusion generated bias in molding. For example: Peng and Sheppard [2] used finite element analysis in their research on multi-hole extrusion process on the mold. They analyzed the amount and distribution

of holes on the mold and their impact on extrusion load and material flow. Chen et al. [3] used DEFORMTM-3D software to analyze the impact of process parameters, such as billet temperature, extrusion speed, billet size, and hole location on the extrusion process in a single, porous, aluminum alloy. Their results indicated that hole location makes a significant impact on the exit tube bending angle generated by the extrusion process. Most previous research findings were based on traditional methods of analyzing multi-hole extrusion process; only a few were based on evolutionary algorithms designed to analyze parameters from the extrusion process. Among evolutionary algorithms, PSO is a technique commonly used to solve optimization problems [4-5]. Relative to GA, PSO does not use operators such as GA selection, crossover, and mutation. Thus, PSO takes less computing time [6] and widely used in many different fields [7]. Though the advantage of PSO is in its fast convergence, it is rather easy to fall into the trap of local optimal solution at later stages of calculation [8]. To overcome problems with local optimal solution and early convergence, other evolutionary algorithms were adopted in PSO calculation in order to move away from the local optimal solution trap [9-11]. In this paper, HPSO was design to investigate mandrel eccentric angle and the exit tube bending angle in multi-hole seamless tube extrusion under various extrusion parameter conditions.

2 Research Method

2.1 Application of FE Modeling

This study used rigid-plastic FE simulations with DEFORMTM-3D software for its research into the multi-hole tube extrusion molding process. DEFORMTM-3D is designed to simulate the distribution of metallic materials in the mold, such as ductile failure value, post deformation temperature, plastic flow speed, and stress and strain. The software is composed of a combination of multiple modules. Its main structure is divided into the pre-processing, post-processing, and multi-function modules, and the simulation engine.

2.2 Optimize Parameter for Multi-hole Extrusion Processes

This study used the HPSO algorithm with FE modeling technique to analyze the effect of various process characteristics on mandrel bias angle and exit tube bending angle in the extrusion molding process. These characteristics include hole position, billet temperature, extrusion speed, and friction coefficient. The process simulation and characteristics are shown in Fig. 1(a) and (b)

2.3 HPSO Algorithm

The search space in PSO is multi-dimensional. As such, the particle group can refer to the optimum experience of an individual or a group in selecting its corrective approach. It can select an objective optimization approach to calculate the fitness of

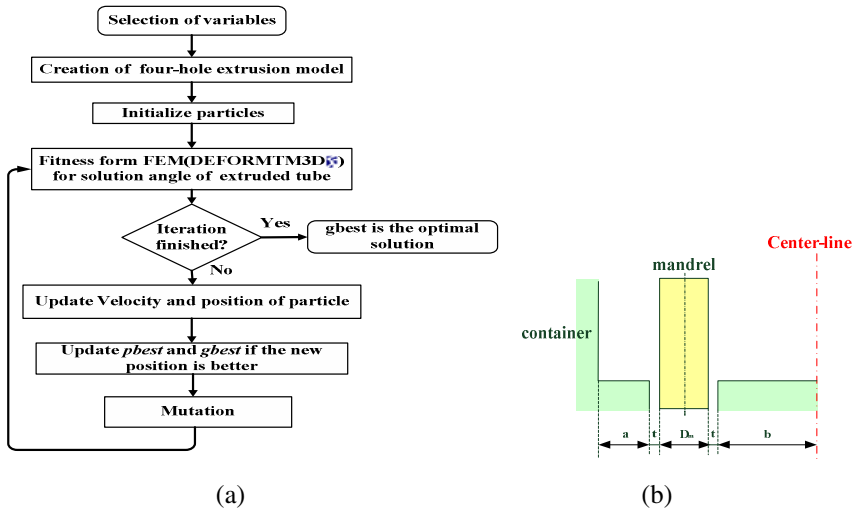


Fig. 1. (a) The process simulation. (b) Process characteristics.

each particle value. During the iteration process, the speed of each particle is calculated as shown in Eq. 1:

$$v_i(t+1) = \omega v_i + c_1 r_1 (p_{id} - x_i(t)) + c_2 r_2 (p_{gd} - x_i(t)) \tag{1}$$

Where, t is the current iteration number, ω is inertia weight, and c_1 and c_2 are learning factors. In this study, ω was set to 0.9, C_1 was set to 2, and C_2 was set to 1.5, while r_1 and r_2 were random values in the range of [0,1]. x_i is the current position of the particle, p_{id} is the optimal solution for each individual particle, and p_{gd} is the optimal solution for all particle groups. The position of each particle is calculated using Eq. 2.

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{2}$$

The main objective of mutation operator in GA is to prevent particles from falling into the trap of local optimal solution. Since excessive mutations can destroy the original particle position, the rate of mutation occurrence in GA is therefore set rather low. In this study, the mutation ratio was set to 10%. L is the lower limit of the parameter, U is the upper limit of the process parameter, and r_3 is a random value in the range of [0, 1]. The HPSO algorithm is designed to always update the variables specified in these equations until termination conditions are reached.

$$x_i = L + r_3(U - L) \tag{3}$$

2.4 Extrusion Parameter Setting

The plastic deformation behavior in the multi-hole extrusion molding process and exit tube bending angle is critically important. These factors include hole location, billet temperature, extrusion speed, and friction. In this study, the billet diameter was set to 100 mm, mandrel length was set to 50 mm, billet thickness was set to 50 mm, and the billet temperature range was set to 150 to 900 °C. The grid was divided into 40,000 elements. The mold was assumed to be rigid. The range of the extrusion molding speed was set to 0.3 mm/s to 1.2 mm/s, and the friction coefficient range of the billet and mold interface was set to between 0.1 and 0.5. Among these, hole position had a huge impact on the multi-hole extrusion process, with its range set to 0.35 to 0.65. As indicated in Fig. 1(b), a is the minimum distance between hole and container, b is the minimum distance between the hole and the center of mold, D_m is the mandrel diameter, and t is tube thickness. Hole position e is defined as in Eq. 1 (a) and Fig. 1 (b).

$$e = \frac{b}{a + b} ; 0 \leq e \leq 1 \tag{4}$$

2.5 Constructing a Multi-hole Extrusion Model

This study used Solid Work 2010 for component modeling of multi-hole extrusion. The multi-hole seamless tube extrusion process was set with the hole number as 4, the container diameter as 100 mm, the hole diameter as 13 mm, and billet length as 50 mm. The choices for tube selection (mandrel diameter of 9 mm and tube thickness of 2 mm) and hole position in the multi-hole extrusion process was complied based on the center of symmetry principle, with a quarter of the model intercepts entered into the FE software for modeling analysis, which is shown in Fig. 2 (a).

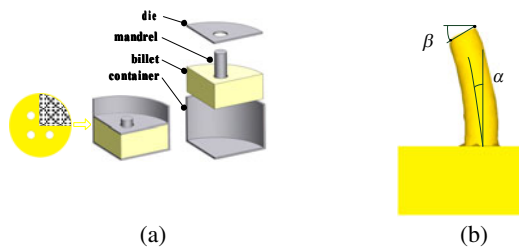


Fig. 2. (a) Three-dimensional FE modeling of multi-hole extrusion model (b) Exit tube bending angles α and β

2.5 Initialization and Fitness Assessment of PSO

In this algorithm, 20 particles were randomly selected at the start, with each particle consisting of a combination of four parameters. Then FE software DEFORMTM-3D was employed to assess the modeling of the multi-hole seamless tube extrusion model

by looking at the fitness of its eccentric angle α and exit tube bending angle β (Fig.2(b)). Smaller α and β angles indicate a better fit.

3 Simulation Results

This study used HPSO algorithm in combination with FE modeling techniques to analyze the impact of various process parameters on the mandrel and exit tube bending angles during the extrusion molding process. The results are shown in Fig. 3(a), 3 (b) and Fig.4.

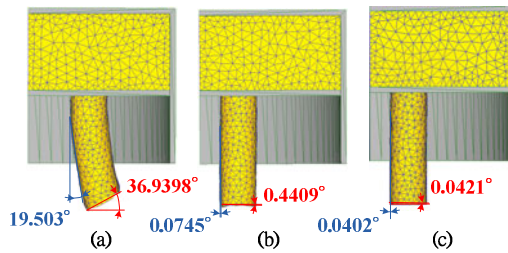


Fig. 3. (a) Results from parameter modeling in parent generation (b) Results of modeling using GA (c) Results of modeling using HPSO

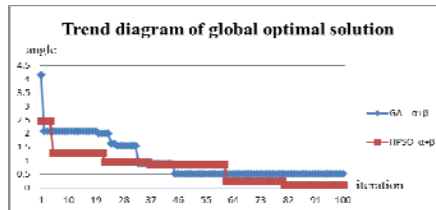


Fig. 4. Convergence comparison of average value in each generation ($\alpha + \beta$) using GA and HPSO

As can be seen in Fig. 3 (a), 3(b) and 3(c), the mandrel bending angle was reduced from 19.5030 ° to 0.0402 °, while the exit tube bending angle was reduced from 36.9398 ° (initial generation) to 0.0421 ° (after 100 iterations). The resulting optimal solutions were quite good in that they are better than those generated from GA. It is revealed from Fig. 4 that 100 iterations of optimization calculation using HPSO were able to generate a solution with the minimum parameter combination, based on mandrel eccentric angle (α) and exit tube bending value (β) for the multi-hole seamless tube extrusion molding process. As can be seen in Table 2, the hole location (e) was 0.4, billet temperature (T) was 658 °C, extrusion speed (V) was 0.57, the friction coefficient (F) was 0.16, the minimum angle (α) was equal to 0.0421°, and the minimum angle (β) equaled 0.0402°.

4 Discussions

The results from the modeling demonstrate that during the extrusion process, the ability of HPSO to find an optimal solution for the extrusion parameter is better than that of GA, especially in terms of accuracy. HPSO is indeed fit to use as a solution generator algorithm for addressing problems arising from the extrusion process. HPSO can be effectively utilized to improve the exit tube bending and mandrel eccentric angles during the extrusion molding process.

References

1. Li, L.X., Rao, K.P., Lou, Y., Peng, D.S.: A Study on Hot Extrusion of Ti–6Al–4V using Simulations and Experiments. *International Journal of Mechanical Sciences* 44, 2415–2425 (2002)
2. Peng, Z., Sheppard, T.: Simulation of Multi-Hole Die Extrusion. *Materials Science and Engineering A* 367, 329–342 (2004)
3. Chen, F.K., Chuang, W.C., Shan, T.: Finite Element Analysis of Multi-Hole Extrusion of Aluminum-Alloy Tubes. *Journal of Materials Processing Technology* 201, 150–155 (2008)
4. Kennedy, J., Eberhart, R.C.: Particle Swarm Optimization. In: *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948. IEEE, Perth (1995)
5. Bergh, F.V., Engelbrecht, A.P.: A Study of Particle Swarm Optimization Particle Trajectories. *Information Sciences* 176, 937–971 (2006)
6. Kennedy, J., Eberhart, R.C.: The Particle Swarm: Social Adaptation in Informal-Processing Systems. In: *New Ideas in Optimization*, pp. 379–387. McGraw-Hill, Maidenhead (1999)
7. Eberhart, R.C., Shi, Y.H.: Particle Swarm Optimization: Developments, Applications and Resources. In: *IEEE Proceedings of the Congress on Evolutionary Computation*, pp. 81–86. IEEE, Seoul (2001)
8. Gaing, Z.L.: A Particle Swarm Optimization Approach for Optimumdesign of PID Controller in AVR System. *IEEE Transactions on Energy Conversion* 19, 384–391 (2004)
9. Andrews, P.S.: An Investigation Into Mutation Operators for Particle Swarm Optimization. In: *Proceedings of the IEEE Congress on Evolutionary Computation*, Vancouver, pp. 1044–1051. IEEE, Canada (2006)
10. Higasshi, N., Iba, H.: Particle swarm optimization with Gaussian mutation. In: *Proceedings Swarm Intelligence Symposium*, pp. 72–79. IEEE, Washington, D.C., USA (2003)
11. Stacey, A., Jancic, M., Grundy, I.: Particle swarm optimization with mutation. In: *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 1425–1430. IEEE, Washington, D.C., USA (2003)

Decision-Making Model of Business Continuity Management

Gang Chen

Shanghai University of Finance and Economic Shanghai P.R. China, 200433
School of Economic and Management, Tongji University Shanghai

Abstract. Business Continuity Management (BCM) becomes the most important assurance of E-learning. We should frame decision-making model of BCM to improve the quality of BCM. Decision-making model of BCM takes advantage of changes to improve competitiveness by interaction between members of BCM in dynamic environment. Description of decision-making model of BCM involves three elements: property of agents, changes in environment and rights and obligations of agents. Decision-making model of BCM consists of four parts: analysis and description of dynamic environment, BCM capability measurement and assessment, BCM strategy and hierarchical structure of decision-making.

Keywords: Business Continuity Management, Decision-making model, Description of decision-making model, Structure of decision-making.

1 Introduction

In economic globalization and development of information technology, business continuity management has to face challenges of changing dynamic environment. Decision-making of business continuity management is capabilities to make effective adjustments. It should take more consideration to environmental changes. Decision-making model of business continuity management takes advantage of changes to improve competitiveness by interaction between members of BCM in dynamic environment.

2 Description of Decision-Making Model of Business Continuity Management

Description of decision-making model of business continuity management involves three elements: First element is property of agent s. Second element is changes in environment (including information, behavior and abilities for respond to changes). Third element is rights and obligations of agents (It is knowledge and decision of agents according to changing environment). In decision-making model, actors make decision according to its own properties and changes of environment. Model = (agents identity attributes, environmental change order, corresponding causal relationship).

Agent s' identity has attributes in model. Each attribute includes co-principal rule, policy set and policy influence composition of interaction. The function is Agent = (co-rule, policy set, sets of interaction effects).

Definition 1. $O = \{A_1, A_2, A_3, \dots \dots A_n\}$ is collection of BCM decision-making body, A_i is the decision-making body. O is BCM organization. $B = \{B_1, B_2, \dots \dots B_n\}$ are behavior set of A_j .

Definition 2. $C = K_1 \cap K_2 \cap \dots \dots \cap K_n$, K_i is the amount of information that decision-making body has. C is common information of all decision-making body.

Definition 3. $G = B_1 \times B_2 \dots \dots \times B_n$, A_j is interaction between main effects.

Environment in the model includes external environment and internal environment. Changes in environment impact by space and time.

Definition 4. $D = \{d_1(e^t), d_2(e^t), \dots \dots d_n(e^t)\}$, d_i is the i -th decision-making agent of BCM reflects ability to changes of environment while time is t . D is ability of BCM respond to environment. $e^t \in E^T$, E^T is environment sets while time is T .

Definition 5. $X = \mu(D, C, E^T)$, X is appropriate response to external effects under certain circumstances E^T when X has a common message of system. Utility function used here means a decision to represent system in time and size of environmental change.

Definition 6. $e^t \in E^T, e^{t'} \in E^T$, $e^{t'}$ is successor of changing environment after e^t . Effect of decision-making at this time would be the ability for respond to environment of the members of BCM while $t \rightarrow t'$. When $D \rightarrow D'$ then $X \rightarrow X'$.

As environment changes, overall effectiveness of BCM will continue to change. These changes can be controlled by controlling agents property and let them change to the most favorable toward of BCM. Dynamic programming method can be used to overall effectiveness of optimization problem. Just like

$$X_{t'} = \max(D, C, E^T) + \alpha E \mu(D', C', E^{T'})$$

α is the discount factor. We can gain cooperation strategy for optimal management of business continuity services by using variables input parameters of this function.

3 Structure of Decision-Making Model of Business Continuity Management

Structure of decision-making model of BCM consists of four parts: analysis and description of dynamic environment, BCM capability measurement and assessment, BCM strategy and hierarchical structure of decision-making, shown in Figure1.

3.1 Dynamic Environmental Analysis and Description

Environment of Business Continuity Management can be divided into three dimensions. The first dimension can be divided into internal environment and external environment. The second dimension can be divided into controllable environment and uncontrollable environment. The third dimension can be divided into general environment and task environment.

3.2 BCM Capability Measurement and Assessment

BCM capability refers to organization changes automatically and rapidly in dynamic environment. BCM capabilities include three aspects such as high available capability, continuous operations capability and disaster recovery capability.

Figure1 Structure of Decision-making model of Business Continuity Management

High available capability is ability to continue business operations when system is failure. System failure includes business processes failures or physical facilities failures. High available capability ensures quality of BCM. The important indicators of availability are MTTF, MTBF and MTBSI. MTTR (mean time to repair) means average time between failures to recovery. It is summation of detection time and resolution time. It is on behalf of organization's response to events and processing capacity. MTBF (mean time between failures) is average time from recovery time of last event to the time of next event occurred. It is the normal running time. The higher of the value the stronger stability of representing maintain service. MTBSI (mean time between system incidents) is average time between adjacent interval events. MTBSI equals to summation of mean time to repair (MTTR) and mean time to failure (MTBF).

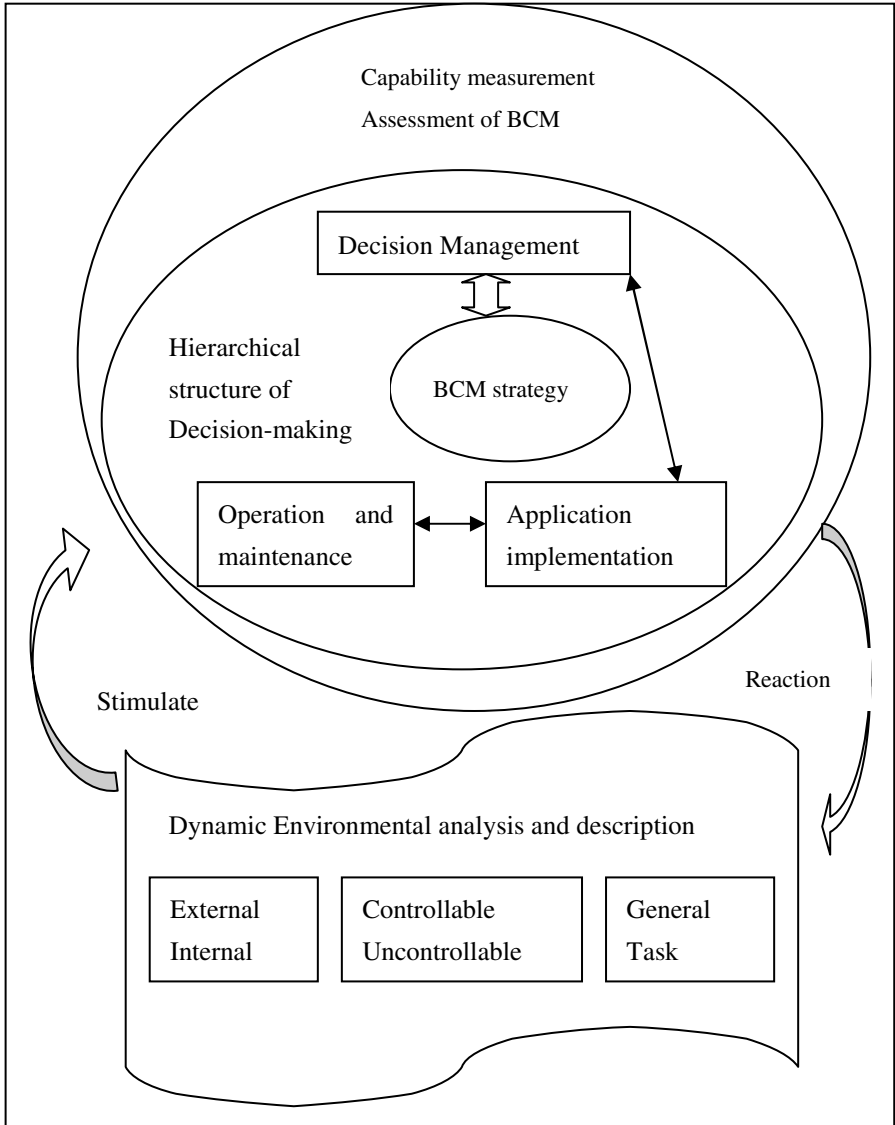
Continuous operations capability is capabilities to maintain business continuity when all equipments are trouble-free running. It means that users do not to stop business operations when they back-up or maintenance system. When failure occurs or business changes, business system can provide continuous uninterrupted service.

Disaster recovery capability means ability to restore business operations in different locations, when business systems are destruction.

3.3 Business Continuity Management Strategy

BCM includes transparent strategy, compromise strategy and redundancy strategy. Transparent strategy means to improve visibility of BCM as business information is passed quickly. Compromise strategy can find balance point between cost control and

high level of compromise. With internal environment and external environment increasing uncertainty, organization will become more willing to implement BCM strategy to reduce risk of emergency. Redundancy strategy is the most commonly method for reduce business continuity management risk.



3.4 Hierarchical Structure of Decision-Making

Hierarchical structure of BCM can be divided into three levels: decision management, application implementation and operation and maintenance.

Decision management of BCM involves establishment of mechanisms including business continuity incentive, restraint mechanisms and establishment coordination mechanisms of Redundancy. Decision management of BCM considers strategic transformation of coupling efficiency and switching costs among members of organization at application implementation level and operational and maintenance level. Decision management should demand project portfolio management, processes, tools issues and business integration. It also should determine necessary processes and tools, real-time understanding of implementation of business process management, impact on business, shown value of BCM.

Application implementation of BCM is to achieve overall BCM by designing organizational structure. Organizational structure design includes design of service agreement, organizational system, service system, technical system, and process design. It should ensure the quality of Business continuity management meanwhile reducing cost of business continuity management.

Objective of operation and maintenance is to improve system operating effectively. Procedures of operation and maintenance are data security, data storage, application architecture design, backup and recovery and infrastructure protection design.

4 Conclusion

Decision-making of business continuity management is capabilities to make effective adjustments in dynamic environment. It improves the quality of BCM and implies necessary processes and tools to control and manage budgets.

Acknowledgement. This work is supported by Leading Academic Discipline Program 211 Project for Shanghai University of Finance and Economics (the 3rd phase).

References

1. Le Van, C., Morhaim, L., Dimaria, C.-H.: The discrete time version of the Romer model. *Economic Theory*, 1148–1158 (2002)
2. Chen, G.: A web service model based on information dimensionality. In: *Proceedings of 2008 IEEE International Conference on Service Operations and Logistics, and Informatics, IEEE/SOLI 2008*, pp. 1084–1088 (2008)
3. Gang, C.: Risk Evaluation of Business Continuity Management by Using Green Technology. In: Zaman, M., Liang, Y., Siddiqui, S.M., Wang, T., Liu, V., Lu, C. (eds.) *CETS 2010. CCIS*, vol. 113, pp. 86–92. Springer, Heidelberg (2010)
4. Gang, C.: IT service system design and management of educational information system. In: *2009 1st International Conference on Information Science and Engineering*, pp. 307–310 (2009)

Research on Energy Supply Modes of Electric Vehicle

Changxian Cheng and Jianhua Yan

School of Information and Mechatronic Engineering,
Beijing Institute of Graphic Communications, Beijing, P.R. China
New Energy Dynamic Company Limited,
No. 6 Haidian North Second Street, Beijing, P.R. China
cx-cheng@163.com

Abstract. This paper studies the development and application of electric vehicle (EV) at home and abroad. Through the research of the application and energy supply of EV, analysis and summary the two viable energy supply modes in current technical situation. One model is the vehicle charge; other is the quick-replacement battery mode, also known as electric changing mode. Both modes of energy supply were compared and analyzed. The technology of vehicle charging mode is achieved with lower degree of difficulty and its application is relatively easy in the early development stages of EV, but requires large-scale construction of charging pillars. When EV becomes more popular, it has greater pressure on the grid and would not be conducive to battery protection. Electric changing mode can quickly replace the battery, so that to provide EV users with the convenience and make the energy supplies for EV as easy as refueling. But it is difficult to achieve a unified standard technical and the cost for electric changing equipment is larger. The reference observations of future EV electric changing equipment are made in the summary.

Keywords: electric vehicle, power battery, quick change, electrical changing equipment, vehicle charging.

I Introduction [1, 2]

With the intensifying global energy crisis and environmental degradation, the development of electric vehicles has become one of the imperatives or the inevitable trends for the world's energy security and environmental protection. As you can see, China ranks now the second globally as for the carbon dioxide emissions. The automobile emission's pollution has become an important factor in urban air pollution and the pressure of some degree of crisis caused by energy imports is increasing. On the other hand, China has the advantages of manufacturing cost and the huge market space, and the development of electric vehicles is the requirement of Chinese automobile industry to develop by leaps and bounds and with sustainability. Our research is dedicated to the difference between EV and conventional fuel vehicles and pays more attention to the practical problems such as shorter EV driving mileage and inconvenient energy supply in the early development stage of EV. The corresponding studies are made on battery changing mode that more suitable to the current trends, and its specific methods and equipments.

2 The Development and Application of Electric Vehicles [2, 3, 4]

A. The development of electric vehicles at home and abroad

Electric vehicle industry was brewing in 1970s, and started to speed up in 1990s. In 1996, the United States and Japan have begun the small batch production, sales and application stage of EV. For the last more than ten years, because of the world-wide oil resource constraints and the increased air pollution, energy saving EVs are brought into focus much more than ever before. Western countries are putting in a lot of investment and manpower, and developed a series of preferential policies to develop EV. With the progress of major R&D projects of EV in China, energy-saving and new energy vehicle demonstration pilots are set up in 13 cities such as Beijing, Shanghai, Chongqing, Changchun, Dalian, Hangzhou, Jinan, Wuhan, Shenzhen, Hefei, Changsha, Kunming, Nanchang, to encourage pilot cities to take the lead to promote using of EV in public services such as transportation, taxi, official business, sanitation and postal services etc.. At the same time, the R&D and industrialization of lithium power battery and relating materials have made great progress.

B. The forecast of development trend of EV in China

1) The development of EV by 2015

EV is at the initial stage of which the main features are that battery technology is at a low level with longer charging time and shorter driving mileage; pure, or battery electric city buses and the plug-in charging buses begin to demonstrate and apply in a small scale; total annual output of all kinds of pure electric passenger vehicles reaches 10 percent of any passenger vehicles. The retain volume of electric passenger vehicles of all kinds is expected to reach 2.25 million. The retain volume of EV passenger buses will be about 51,000, of which 80% being public buses, 20% other passenger buses such as electric shuttle bus etc.. The retain volume of special EV is about 250,000, of which 20% being electric sanitation trucks, the rest 80% other special vehicles such as electric postal vehicles etc.. The retain volume of electric taxis would be about 100,000.

2) The development of EV by 2020

EV enters into the development period of which the main features are that battery technology will improve greatly, the charging time will be shortened, and the driving mileage of EV raises greatly; all kinds of dedicated short-range electric vehicles will begin to popular; The annual production of all kinds of electric passenger vehicles will reach 20% of all kinds of passenger vehicles; the retain volume of passenger vehicles will be expected to 15.37 million. The retain volume of EV passenger buses will be about 200,000, of which 80% being electric buses, 20% other passenger vehicles such as electric shuttle bus etc.. The retain volume of special EV is about 1 million, of which 20% being electric sanitation trucks, the rest 80% other special vehicles such as electric postal vehicles etc.. The retain volume of electric taxis would be about 200,000.

3) The development of EV by 2030

EV would enter into the popularity period of which the main features are that all kinds of electric passenger vehicles will begin to be made; New generation of batteries will be used; the battery charging time will be greatly shortened; driving mileage of vehicles will be over 400 kilometers; the performance, cost, output of the EV will equal the level of oil fuel cars. EV will have a mass production and wider application. The retain volume of passenger vehicles will be expected to 61.07 million. The retain

volume of EV passenger buses will be about 1,720,000, of which 80% being electric buses, 20% other passenger vehicles such as electric shuttle bus etc.. The retain volume of special EV is about 5 million, of which 20% being electric sanitation trucks, the rest 80% other special vehicles such as electric postal vehicles etc.. The retain volume of electric taxis would be about 800,000.

3 The Research on Energy Supply Modes of EV [4, 5]

A. Problems in the present development of EV

The main problems in the energy supply of EV batteries at present stage are: First, due to the EV technology is still in the early stages of development and its technical imperfections, the charging time of EV battery is long, which is less convenient than refueling of conventional fuel vehicles and hence difficult to meet the users' requirement of convenience. Secondly, there are still not better forms of energy supply both at home and abroad; and the charging stations for EV are not like the gas stations which support the development of fuel vehicles. EV's users worry about the problems of energy supply, resulting in that EV does not sell, the market is not active, and its technology lacks strong momentum.

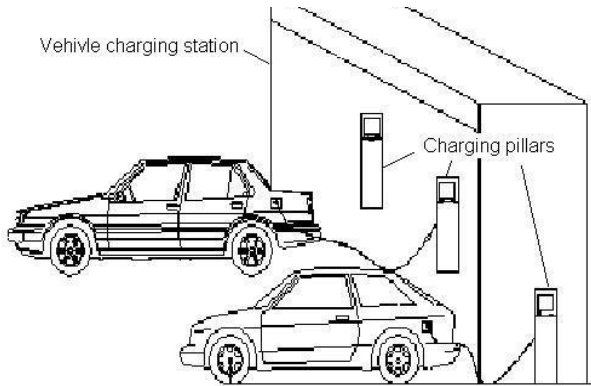


Fig. 1. Vehicle charging of EV

B. The study on energy supply modes of vehicle charging

At present, battery energy supply for pure electric passenger vehicles, both at home and abroad, is still using the vehicle charging mode (Figure 1). The passenger vehicles with batteries are charged by an external battery charging device. However, based on the current developments, the vehicle charging has also the undeniable advantages and the inevitable shortcomings.

1) Advantages:

- i . The equipments are simple, low cost, and suitable for initial development;
- ii. A lower requirement for standardization, as long as the charging plug and charging protocol standardization, charging equipments are compatible with various vehicle models;

- iii. For the planning of city construction, various types of charging piles can be used in the parking areas, increase another function for the parking areas;
- iv. Private cars can choose their own charging way, so that to save fuel costs and thus stimulate the market.

2) *Disadvantages:*

- i. At present, the driving mileage of battery is short. The users need more frequent charging;
- ii. The charging time is longer. It's inconvenient for the users;
- iii. The users' own batteries are not conducive to the safety protection of battery and it's easy to reduce the cycle life;
- iv. For large-scale development, it is difficult to unify users' charge manner and will generate pressure on power grids;
- v. For the urban residents, it is unlikely to charge under the current electricity equipped for residential buildings.

C. The study on energy supply mode of electric changing

As the limited battery performance in this stage, the average mileage of a single charge of electric passenger vehicles is about 160 km. The demand for running distance of EV during a day is far longer than the existing driving mileage of a single charge. If using energy supply mode of quick-change battery (see Figure 2), the public battery quick-change networks for EVs can be constructed in terms of the fact of EV's free driving directions. EV can drive to near battery changing stations before battery exhausts. While the empty battery is removed by battery replacing equipment, a fully charged battery is replaced.

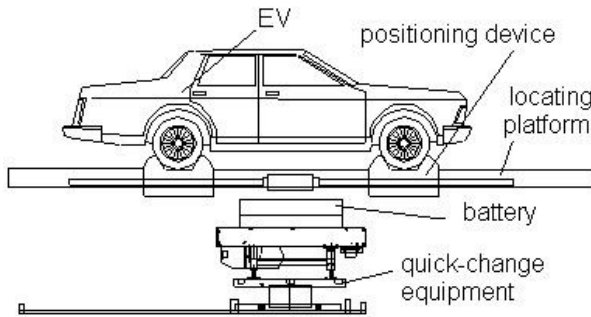


Fig. 2. EV in battery changing

Comparing with energy supply mode for electric charging, energy supply mode for battery changing has many advantages: first of all, the electric changing time is short, as easy as refueling, and suitable to today's needs of modern life; Second, since the electric changing batteries can be charged centrally, the requirements on the site and power are lower, so that to facilitate control of charging time and to play a role of cutting peak to fill the valley of the grid; Thirdly, concentration of electric charge in electric changing

can unify the battery management, and hence has the advantages of the security protection to the batteries used and the extended battery life.

The energy supply mode for battery changing also has the following inherent drawbacks. Firstly, battery changing equipment is a complete automation equipment with more complex technology and higher precision and higher cost; Secondly, The energy supply mode for battery changing needs a unified standard so that to facilitate the exchange among battery changing equipments, batteries and EVs.

The following will discuss on battery changing process in a more detailed way.

4 The Process and Methods of Quick Replacement of EV Battery [6, 7]

The battery changing mode for EV is relatively complex. When the driver drives the vehicle into battery changing station, he punches the card in a pos machine, drives the car into the vehicle location device by following the signs, and then parks the vehicle with the help of parking positioning slot. The driver presses the button of battery replacement when a wireless communication system at the front windshield of the vehicle establishes wireless communication with battery changing equipment on ground and transfers in real-time the status of battery changing to the ground equipment. The front signal lights flash yellow and green alternately. The battery changing process starts now (see Figure 3) and the battery replacement equipment is running. At the same time, Information Scheduling Management System will read the information in the card, deploy the correspondent type of battery to the car, and manipulate battery handling equipment to retrieve the battery from the battery storage rack and transfer the battery to the battery transfer equipment. At the same time, the vehicle positioning and dragging device positions the vehicle longitudinally and

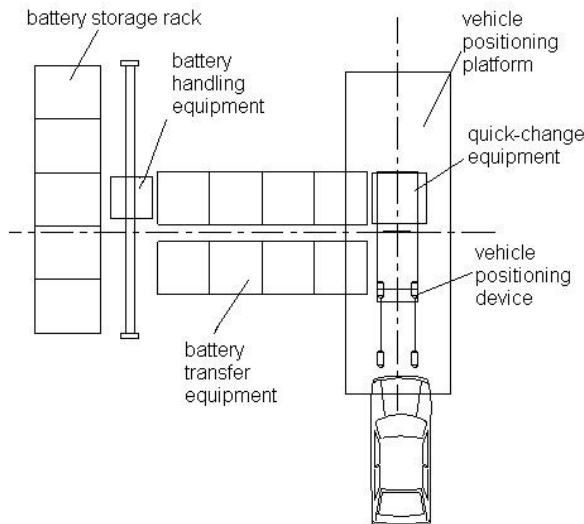


Fig. 3. Schematic diagram of EV's battery changing

laterally according to vehicle position information. After the quick-change equipment (see Figure 2) positions the vehicle along its vertical and horizontal positioning adjustable rails, the lifting platform is lifted up to the bottom of the vehicle. When the lifting platform is in full access to the battery compartment installed in the bottom of vehicle, the locking mechanism in the battery compartment unlocks, the lifting platform falls down, and the battery box is removed from the vehicle and transported to the background position. Meanwhile, a fully charged battery has been in place in the waiting area. The conveyor system will transport a fully charged battery to the lifting platform which lifts up the battery box to contact with the vehicle completely; the locking system will then lock the battery box in the vehicle. The lifting platform falls down, the light turns green, and the battery replacement process is completed.

5 Conclusion

The study on EV and battery changing equipment is the only game which China stands on the same starting line with the developed country in the world. We have to fight the initiative. Through the above analysis and research, we believe that EV is a star in the new energy industry and solving the EV's energy supply model is the key issue of solving problems of EV's sustainable development. Both the battery changing mode and vehicle charging mode are undoubtedly the best choices in today's current technologies. It needs not discard any kind of EV's energy supply models which can promote EV's development. We can develop and apply both modes in a parallel way and the market will determine which technology is the final development rout.

References

- [1] Teng, L.: Electric vehicle charging machine (station) design, p. 21. China Electric Power Press (2009)
- [2] Zhang, T., Jia, Y.: Electric vehicle technology revolution, p. 150. Machinery Industry Press (2010)
- [3] Liu, X., Qi, W.: Analysis of Electric Vehicle and Its Developing Trends. *Hunan Electric Power*, 59–62 (October 2010)
- [4] Nan, J.: Study of Energy Management System of Electric Vehicle. *Transactions of Beijing Institute of Technology*, 384–386 (May 2005)
- [5] Liu, B.: EV Battery Management System. *Electrical Automation*, 60–61 (January 2010)
- [6] Shen, Y.: *Theory of Machines and Mechanics*, 2nd edn. Higher Education Press, Beijing (2005)
- [7] Cheng, C.: New Trends of Ergonomics and Its Importance in Modern Industrial Design. In: *Proceedings of ICEIS 2011*, pp. 543–547 (2011)

Why Does Firm Choose Internet Marketing Channel? A Firm Perspective and Contingent Framework

Maohong Liu¹, Xiaoliang Feng², and Minxue Huang²

¹ Management School, Wuhan University of Science and Technology,
Wuhan City, P.R. China
maohongliu@sohu.com

² School of Economics and Management, Wuhan University,
Wuhan City, P.R. China
ebusiness@whu.edu.cn

Abstract. The emergence and diffusion of the Internet motivates the firm to change and take the Internet as a kind of marketing channel. Extensive literatures focus on the consumer's perspective but seldom literatures research from firms' perspective. This paper aims to analyze the determinants of applying or switching to the Internet direct marketing channel from firms' perspective. First, the paper reviews the relevant literatures of marketing channel structure and transaction cost theory; second, the paper summarizes the possible determinants such as transaction cost and their effective condition; last, the paper proposes a synthesized framework and eight propositions.

Keywords: Internet marketing channel, Transaction cost economics, Diffusion of the Internet.

1 Introduction

Prior literature proves that application of information technology can improve efficiency of transaction and lead to lower transaction cost (Malone et. al. 1989). Internet as one of the most important IT, its diffusion can decrease transaction cost, especially in the external market. According to the transaction cost theory originated from Coase and followed by Williamson (1979), the company may choose to outsource their internal functions/department (e.g. OEM) for lower external market cost. Nowadays, Internet marketing channel is becoming a new channel for firm.

Literatures (Malone et. al. 1987; Afual 2003; Stapleton 2001) try to resolve this question, but they focus on the discussion of whether the intermediaries should exist or not. In fact, the constructing of channel is a complicated strategic issue and cannot be simplified to answer the question of choosing intermediaries or not. Generally, if the current marketing channel mode doesn't have the cost advantage after adopting Internet, the manufacturers will not switch the current channel mode to a new one; Another neglected concern is the switching cost of reconstructing channel which is an evolutionary process and needs cost (Coughlan et al., 2001).

Our study aims to filling these research gaps and provide some practical instructions. We take the adoption of Internet Channel as an evolutionary process of

firm's reconstructing channel in Internet era, combining theory of transaction cost economics and classical marketing channel theory to instruct how firm chooses suitable strategy by analyzing the relevant determinants of transaction cost.

2 Theoretical Backgrounds

2.1 The Choice of Marketing Channel and Transaction Cost Economics

Lots of marketing researchers employed theory of transaction cost economics (TCE) to explain why firms choose different marketing channel modes. Leading research was conducted by Anderson (1985; 1996), who concerned outsourcing issue of sales function and employed TCE to explain how firms make choice. Generally, TCE has three important dimensions as the following:

Asset specificity is defined as kind of non-redeployable physical and human investments that are specialized and unique to a task by Williamson (1979). Anderson (1985; 1996) finds that the specificity of the intangible assets is significantly related with the integration of marketing channel.

Uncertainty is a core assumption in transaction costs theory and means "unknowable" character of future events and outcomes. Its effect on integration of channel is ambiguous. According to Williamson(1979), external uncertainty can lead the firm to requiring for a greater integration while the external uncertainty may lead a firm to vertical integration. But related literatures show controversial viewpoints about the effect of uncertainty.

Frequency means the frequency of transaction.

2.2 The Effect of the Internet Diffusion on Transaction Cost

Transaction cost is defined as a series of cost of handling information during the process of collaboration and communication and performance by Malone et al. (1989).They suggest that the development of information technology can improve the efficiency of the firms' transaction, and reduce the collaboration cost within a firm.

Furthermore, Afuah (2003) suggests that the internet can reduce an Asset specificity by (1) making more information about the asset and potential uses available to more owners at a low cost, (2) reducing the need to move relationship-specific assets, and (3) replacing specific information technology assets.

3 A Contingent Framework from Firm's Perspective

3.1 A Model of Internet Diffusion Effecting on Marketing Channel Choice

Based on the marketing channel researches in TCE framework,we propose a new research model (Figure 1) .The model implicates that on the condition of internet diffusion in a firm, the information technology of the units would moderate the relationship between the internet diffusion and the determinants of transaction cost; And the organizational technology of the units is a moderator of the relationship

between the internet diffusion and the opportunism and asymmetry. The determinants of transaction cost, information technology, and organizational technology have direct or indirect impacts on internal and external transaction cost. In addition, for the firm, which has an established channel, we, furthermore, include the effect of various switching cost on choosing the marketing channel.

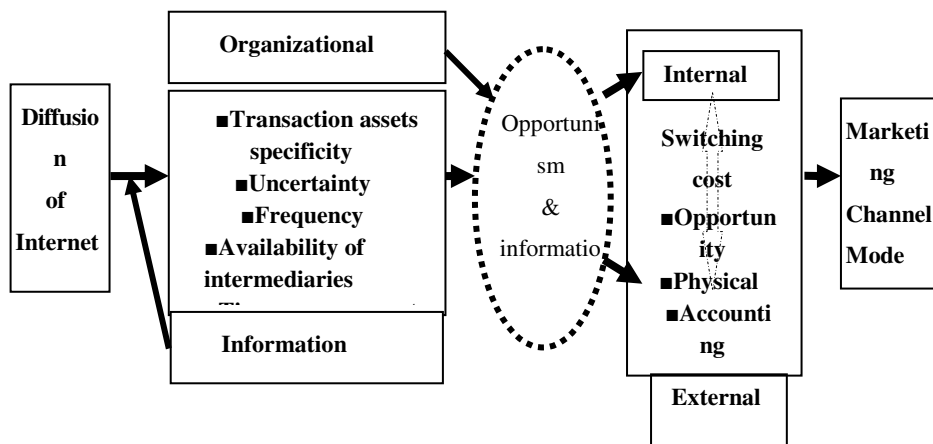


Fig. 1. How the Internet diffusion effect on firm’s choice of marketing channel mode

3.2 The Possible Factors of the Internet Diffusion Effecting on Transaction Cost

Information technology. Information dependence .Information dependence can be defined as the extent to which the value is added or exchanged through information rather than materials by departments (Afuah, 2003). Thus, the higher the information dependence is, the more likely a firm can add value over the internet . Meanwhile, Dewett and Jones (2001) propose that information technology has become a primary means of managing and reducing the uncertainties surrounding production and administrative processes.

P1a: the higher the information dependence between the units in marketing channel is, the more likely the diffusion of the Internet can strengthen the transaction relationship between them.

P1b: the higher the information dependence between the units in marketing channel is, the more likely the diffusion of Internet can reduce the internal uncertainty.

Information explicitness Knowledge is explicit if it is written, verbalized, or coded in drawings, computer programs, or other products (Polanyi, 1962).

P2a: the higher the information explicitness of the units in marketing channel is, the more likely the diffusion of the Internet can strengthen the transaction relationship between them.

P2b: the higher the information explicitness of the units in marketing channel is, the more likely the diffusion of the Internet can reduce the internal uncertainty.

Organizational technology. *Interdependence* is the extent to which value-adding departments depend on each other for information, materials, or resources to accomplish their tasks (Steensma and Corley, 2000; Tushman, 1979). Afuah (2003) suggests that high interdependence between units also means that one unit is more likely to have information that the other does not have.

P3a: the higher the interdependence between the units in marketing channel is, the more likely the diffusion of the Internet can reduce opportunism and the transaction cost between them.

Task variability is the number of exceptions that value-adding units encounter as they perform tasks in their value-adding activities (Daft and Macintosh, 1981) As task variability increases, the need for coordination and communication increases (Tushman, 1979).

P3b: the higher the task variability in marketing channel is, the more likely the diffusion of the Internet can reduce opportunism and transaction cost .

3.3 The Effect of Transaction Cost Determinants on Marketing Channel Choice

Transaction Asset specificity. Only a few physical assets are involved in marketing function, thus, the transaction specificity assets should involve three major parts: human assets, enterprise's reputation and ability of charging information.

Human assets refer to the transaction relationship between enterprises and customers. The specificity of human assets relies on the single transaction relationship between the enterprises and the customers, excluding the private relationship between customers and sales persons and the product's brand reputation.

P4a: the more specific the human assets are, the more likely the manufacturer prefers to apply the Internet direct marketing mode.

Enterprise's reputation is a trustworthiness which can let customer maintain the transaction relationship firmly without any force. Barney and Hansen (1994) note that some organizations are intrinsically trustworthy because it is critical to the organization's self-definition never to exploit a partner's vulnerabilities. When the intermediaries' reputation is high, it is easy for them to grasp the customer resource .

P4b: the better the enterprise's reputation is, the more likely the manufacturer prefers to apply the Internet direct marketing mode.

Ability of charging information We define information as all sources of information that can bring asymmetry and advantage. The more the enterprises grasp this kind of information, the more they grasp the market.

P4c: the stronger the enterprise's ability of charging information is, the more likely the manufacturer prefers to apply the Internet direct marketing mode.

Uncertainty. Internal uncertainty means the difficulty with which individual productivity can be metered (Alchian and Demsetz, 1972). We refer to this internal uncertainty as the difficulty of evaluating performance. When performance is difficult to evaluate, imperfect input measure is still preferable to defective output measure.

P5: the higher the internal uncertainty is, the more likely the manufacturer prefers to apply the Internet direct marketing mode.

External uncertainty is defined as environmental uncertainty. Volume unpredictability is the inability to forecast accurately the volumerequirements from the final consumers.

Availability of intermediaries .Here we divide intermediaries into two groups, one is nonspecific intermediary, and the other is specific intermediary.

Nonspecific intermediaries include wholesalers and retailers. The higher the availability of nonspecific intermediaries is, the more likely the manufacturer is to use direct selling mode.

P6a: the higher the availability of nonspecific intermediaries is, the less likely the manufacturer prefers to apply the Internet direct marketing mode.

Specific intermediaries generally are considered as a specific function entity including logistics and transportation enterprises, information technology enterprises and etc. The more the specific intermediaries is and the more their ability of suiting for the manufacturer, the less the cost is.

P6b: the higher the availability of specific intermediaries is, the more likely the manufacturer prefers to apply the Internet direct marketing mode.

Time span to feedback.Time span to feedback refers to the time span from ordering the product to transferring the ownership and the entity of it. When the requirement of time span to feedback is not very high, the manufacturer will not employ the intermediaries.

P7: the longer the time span to feedback is, the more likely the manufacturer prefers to apply the Internet direct marketing mode.

Switching Cost. Coughlan (2001) defined switching cost in *Marketing Channel* as a temporary loss for giving up the established channel to a new one. He suggested that estimating switching cost should include the opportunity cost of channel switching; the psychical cost of employees and the accounting cost of rebuild a channel.

The opportunity cost of channel switching arises from the former volume requirement disappearing. The difference between the former market requirement and the latter one relies on the customers' dependence on retailers.

P8a: the higher the opportunity cost, the less likely the manufacturer prefer to switching marketing channel.

The Psychical Cost of employees is due to the turnover of personnel.

P8b: the higher the psychical cost is, the less likely the manufacturer prefers to switching marketing channel.

The Accounting Cost of recruiting, training, and infrastructure building is explicit, and it is easy to compute and consider.

P8c: the higher the accounting cost is, the less likely the manufacturer prefer to switching marketing channel.

4 Further Research

Our research gives a direction of a further empirical examination. On one hand, excluding switching cost, we analyze the relationship between the transaction cost determinants and the application of internet direct marketing channel; On the other hand, including switching cost, we must consider the impact of switching cost on

switching marketing channel. In addition, the survey should be in the firms, which are different in information technology and organizational technology.

Acknowledgement. This research was supported by the Hubei Provincial Department of Education under Grant 060164 and the National Natural Science Foundation of China under Grant 70972091. Xiaoliang Feng is the correspondent.

References

1. Afual, A.: Redefining Firm Boundaries in The Face of The Internet: Are Firms Really Shrinking? *Academy of Management Review* 22(1) (2003)
2. Alchian, A.A., Demsetz, H.: Production, Information Costs, and Economic Organization. *American Economic Review* 62(5), 777–795 (1972)
3. Anderson, E.: The Salesperson as Outside Agent or Employee: A Transaction Cost Analysis. *Marketing Science* 4(3) (Summer 1985)
4. Anderson, E.: *Transaction Cost of Analysis and marketing-Transaction Cost Economic and Beyond*. Kluwer Publishing, London (1996)
5. Anderson, E., Weitz, B.: The Use of Pledges to Build and Sustain Commitment in Distribution Channels. *Journal of Marketing Research* 24 (February 1992)
6. Barney, J., Hansen, M.: Trustworthiness as a source of competitive advantage. *Strategic Management Journal* 15, 175–190 (1994)
7. Coughlan, A.T.: *Marketing Channel*. Prentice-Hall Press (2001)
8. Daft, R.L., Macintosh, N.B.: A Tentative Exploration into the Amount and Equivocality of Information Processing in Organizational Work Units. *Administrative Science Quarterly* 26(2), 207–224 (1981)
9. Dewett, T., Jones, G.R.: The role of information technology in the organization: a review, model, and assessment. *Journal of Management* 27(3), 313–346 (2001)
10. Heide, J.B., John, G.: The Role of Dependence Balancing in Safeguarding Transaction-Specific Assets in Conventional Channels. *Journal of Marketing* 52(1) (1988)
11. Malone, T.W., Yates, J., Benjamin, T.: *The Logic of Electronic Markets*. Harvard Business Review (May-June 1989)
12. Neves, M.F.: A Model for the Distribution Channels Planning Process. *Journal of Business and Industrial Marketing* 16(6-7) (2001)
13. Polanyi, M.: *Personal knowledge*. University of Chicago Press, Chicago (1962)
14. Stapleton, D.: The Location-Centric Shift from Marketplace to Market space: Transaction Cost- Inspired Propositions of Virtual Integration via an E-Commerce Model. *Advances in Competitiveness* 9(1) (2001)
15. Steensma, H.K., Corley, K.G.: On the Performance of Technology-Sourcing Partnerships: The Interaction between Partner Interdependence and Technology Attributes. *Academy of Management Journal* 43(6), 1045–1067 (2000)
16. Tushman, M.L.: Impacts of Perceived Environmental Variability on Patterns of Work Related Communication. *Academy of Management Journal* 22(3), 482–500 (1979)
17. Williamson, O.E.: Transaction-Cost Economics: The Governance of Contractual Relations. *The Journal of Law and Economics* 22 (October 1979)

Study on Diluted Symmetric Discrete-Time Hopfield Network

Li Tu and Yan Wang

Department of Computer Science Hunan City University, Yiyang, Hunan, 413000, China
tulip1907@163.com, shaoguanzai@sina.com

Abstract. An efficient system designing method for endowing the diluted symmetric discrete-time Hopfield networks with retrieval properties is proposed based on the matrix decomposition and connection elimination method. Numerical simulations show that the blurred patterns can correctly retrieved, and about 80% wiring cost can be reduced. This method can be applied in the design of large-scale neural networks, this greatly reduces the cost of neural circuit design.

Keywords: diluted symmetric discrete-time Hopfield network, matrix decomposition, connection elimination method.

1 Introduction

Artificial neural network is used to mimic the human brain structure and function. Some studies show that the brain's synaptic connections among neurons is sparsity. Usually the more the number of synaptic connections, the more resources consumed, the consumption of diluted network is significantly less than the fully connected network, so a lot of resources can be saved.

2 Network Model

The following is a n-neuron Hopfield network $x_i(t) = \pm 1$ is the i-th neuron state. Set $X(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$, it is the network state at time t, then the network state in time t+1 is determined by the following formula:

$$X(t+1) = \text{SGN}(WX(t)) \quad (1)$$

in this formula $\text{SGN}(WX(t)) = [\text{sgn}(f_1(t)), \dots, \text{sgn}(f_n(t))]^T$, and $f_i(t) = \sum_{j=1}^n w_{ij} x_j(t)$, the sign

function $\text{SGN}(\cdot)$ is defined in : $\text{sgn}(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$, $W = [w_{ij}] \in R^{n \times n}$, and $w_{ij} = w_{ji}$.

3 System Design

Energy function of network(1) is

$$EX = -\frac{1}{2} X^T W X \tag{2}$$

$\lambda_1(W)$ and $\lambda_n(W)$ are the maximum and minimum eigenvalue of Matrix W, then

$$-\frac{n\lambda_1(w)}{2} = -\frac{\lambda_1(w)}{2} X^T X \leq EX \leq -\frac{\lambda_n(w)}{2} X^T X = -\frac{n\lambda_n(w)}{2} \tag{3}$$

$X^{(j)} = [x_1^{(j)}, x_2^{(j)}, \dots, x_n^{(j)}]^T, (j = 1, 2, \dots, m)$ is the vector mode need to be

recognized by the Hopfield neural network. We need to design appropriate parameter W to meet the following conditions.

$\Delta E \leq 0$, it is $W \geq 0$, (i)

$X^{(j)}$ is the balance point, (ii)

$X^{(j)}$ is the local minimum point of formula $E(\cdot)$, $j = 1, 2, \dots, m$, (iii)

When the state of the network synchronization was updated in parallel, the energy network can be expressed as:

$$\begin{aligned} \Delta E &= E(X(t+1)) - E(t) = -\frac{1}{2} [X^T(t+1)WX(t+1) - X^T(t)WX(t)] \\ &= -\frac{1}{2} [(\Delta X)^T WX(t) + X^T(t)W\Delta X + (\Delta X)^T W\Delta X] = -(\Delta X)^T WX(t) - \frac{1}{2} (\Delta X)^T W\Delta X \end{aligned} \tag{4}$$

and $\Delta X = X(t+1) - X(t) = [\Delta x_1, \Delta x_2, \dots, \Delta x_n]^T, \Delta x_i = x_i(t+1) - x_i(t), i = 1, 2, \dots, n$, in this formula,

$$\Delta x_i = \begin{cases} 0, & x_i(t) = \text{sgn}(f_i(t)) \\ 2, & x_i(t) = 1, \text{sgn}(f_i(t)) = -1 \\ -2, & x_i(t) = -1, \text{sgn}(f_i(t)) = 1 \end{cases} \tag{5}$$

the above equation shows $-(\Delta X)^T WX(t) \leq 0$, if $W > 0$, then $-\frac{1}{2} (\Delta X)^T W\Delta X \leq 0$.

When $W \geq 0$, with the network status updated, the network energy function $E(\cdot)$ monotonically decreasing, and $\Delta E = 0$, $X_{(t)}$ is the balance point.

We proposed an algorithm based on a fully connected network.

Step 1) Set $\bar{Y} = [X^{(1)}, X^{(2)}, \dots, X^{(m)}]$;

Step 2) Carried out \bar{Y} by singular value decomposition, $\bar{Y} = \hat{U}\Lambda\hat{V}^T$, \hat{U} and \hat{V} are the orthogonal matrix, Λ is a diagonal matrix, its elements are the singular values of matrix \bar{Y} ;

Step 3) Calculate $W = \hat{U}L\hat{U}^T$, $l = \text{diag}[l_1, l_2, \dots, l_n] \in R^{n \times n}$, and $l_i = \begin{cases} \tilde{\alpha}_1, & i \leq m \\ \tilde{\alpha}_2, & i > m \end{cases}$, $i = 1, 2, \dots, n$, and $0 < \tilde{\alpha}_2 < \tilde{\alpha}_1$.

The above design method shows :

$$W\bar{Y} = \hat{U}L\hat{U}^T\hat{U}\Lambda\hat{V}^T = \hat{U}\Lambda\hat{V}^T = \tilde{\alpha}_1\hat{U}\Lambda\hat{V}^T = \tilde{\alpha}_1\bar{Y} \tag{6}$$

Then $X^{(j)} = \text{SGN}(WX^{(j)})$, $j = 1, 2, \dots, m$, so $X^{(j)}$ can be designed as the balance point of network (1), this satisfies the condition (ii), in addition, $W = \hat{U}L\hat{U}^T$ can be seen as orthogonal decomposition, then $\lambda_1(W) = \tilde{\alpha}_1$, $\lambda_n(W) = \tilde{\alpha}_2 \geq 0$, this satisfies the condition (i), and

$$E(X^{(j)}) = -\frac{\tilde{\alpha}_1}{2}(X^{(j)})^T X^{(j)} = -\frac{n\lambda_1(W)}{2}, j = 1, 2, \dots, m \tag{7}$$

We can prove that $X^{(j)}(j = 1, 2, \dots, m)$ is the local minimum point of formula $E(\cdot)$. There are some unimportant connections in the network. If $|\sigma|$ is small

enough, and $f_i(t) \neq 0$, then $\text{sgn}(f_i(t)) = \text{sgn}(f_i(t) + \sigma)$, $i = 1, 2, \dots, n$. We get a diluted

network by removing some unimportant weights. The strength of synaptic connections

are $\bar{\varphi}(w_{ij}) = \frac{|w_{ij}|}{\max_{1 \leq j \leq n} |w_{ij}|}$, if $\bar{\varphi}(w_{ij}) \leq c$, the synaptic connection is unimportant, and

$c \in [0, 1]$ is a removing parameter.

4 System Simulation

Build a fully connected network of 300 neurons, the network needs to remember 20 linearly independent binary vectors, The network parameters are: $n = 300$, $m = 20$, $\tilde{\alpha}_1 = 2$ and $\tilde{\alpha}_2 = 0$. The sparsity of the symmetrical network is :

$$\Gamma = \frac{\text{The number of zero weights dendritic connections}}{n \times n} \tag{8}$$

We use the following function to describe the performance of the network:

$$\Psi(X) = \sum_{i=1}^n x_i x_i^* / n \tag{9}$$

$X = [x_1, x_2, \dots, x_n]^T$ is the final output of the network, the recovery mode, $X^* = [x_1^*, x_2^*, \dots, x_n^*]^T$ is the desired output, to the model to be fully restored, $\Psi(X) = 1$. To each storage mode, We take the opposite value in each dimension of the vector with 15% probability, and we got 20 damaged models ($X_j(0)$, $j = 1, 2, \dots, 20$), we calculated $\Psi(X_j(0)) = 0.7$, $j = 1, 2, \dots, 20$. The damaged patterns were provided to the network as the initial states. The network state continuously synchronized parallel updated, when the network eventually converged to a stable state, we got the average network performance $\bar{\Psi}(X)$, and $\bar{\Psi}(X) = \frac{1}{20} \sum_{j=1}^{20} \Psi(X_j)$.

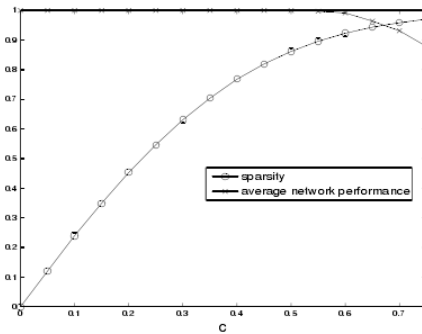


Fig. 1. Statistical properties of the Hopfield neural network

Figure 1 shows the changing of network sparsity and performance when the variable c changes from 0 to 0.75 (Step value is 0.05). When $c \leq 0.5$ removing unimportant synaptic connections does not reduce the performance. Especially when $c = 0.5$, about 86.2% of synaptic connections are removed, but the diluted network can still make the pattern recognition accurate.

Figure 1. shows, when $0.5 < c \leq 0.75$, although more synaptic connections were removed, the network performance is still greater than 0.88.

When $c=0.70$ and $c=0.73$, statistical properties of the network is shown in Table 1. In Table 1, \bar{L}^n is the shortest average path, D_c is the average degree, A_c is the average clustering coefficient.

Table 1. Statistical properties of the Hopfield neural network

Delete parameters	\bar{L}^n	D_c	A_c	Γ	$\bar{\Psi}(X)$
C=0.70	2.70	12.6	0.310	0.958	0.96
C=0.73	2.90	10.1	0.340	0.967	0.90

Table 1 shows, when $c=0.70$ and $c=0.73$, the length of network average paths was 2.7 and 2.9, most of the neurons are not directly connected, However, a neuron can reach another neuron only by moving a few steps, and the network average clustering coefficient was 0.31 and 0.34.

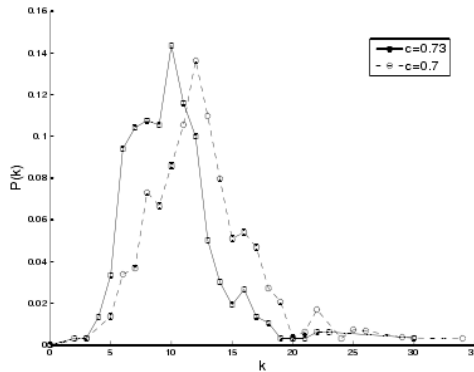


Fig. 2. Distribution of Degree when $c = 0.70$ and $c = 0.73$

When $c=0.70$ and $c=0.73$, the degree distributions $P(k)$ of the Hopfield neural network were showed in Figure 2. $P(k)$ is the proportion of neurons.

5 Conclusion

The simulation results show that, within a certain range the performance of the network does not reduce by deleting some unimportant weights, and compared with the fully connected network, diluted discrete-time Hopfield network has both smaller consumption and high quality of pattern recognition. It is well known that, Hopfield network can be achieved with electronic circuits, the results showed that about 80% of synaptic connections can be deleted, and it has important implications in the application of neural electronic circuits.

References

1. Huang, Y., Yang, X.S.: Hyperchaos and bifurcation in a new class of four-dimensional Hopfield neural networks. *Neurocomputing* 69, 1787–1795 (2006)
2. Lu, H.: Global exponential stability analysis of Cohen-Grossberg neural networks. *IEEE Trans. Circuits Syst.* 52(8), 476–479 (2005)
3. Huang, W.Z., Huang, Y.: Chaos of a new class of Hopfield neural networks. *Appl. Math. C.* 206, 1–11 (2008)

Research on the BM Algorithm Based on the Intrusion Detection System

Xuewu Liu¹, Lingyi Peng², and Chaoliang Li³

¹Hunan University of Commerce Beijin College

²Hunan First Normal University

³Hunan University of Commerce,
Changsha, China

{12870595, 494680234, 522396825}@qq.com

Abstract. Intrusion detection is a mean of proactive network security defense, which can enhance network security by analyzing the data stream on the network to identify potential threat of invasion. In the intrusion detection system, the detection engine is the core part. Although the BM algorithm based on the traditional pattern matching has its advantages, it does not take into account the matched suffix and the adjacent relationships between the current characters leading to the match fails, which results in its low efficiency. This paper will make some improvements to the BM algorithm, so that these can improve the detection accuracy and system performance by reducing the amount of calculation in detection process.

Keywords: Network security, Intrusion detection, The BM algorithm.

1 Preface

With the rapid development of network technology, Network openness, sharing and interconnection degree expand ceaselessly, The network has also brought increasingly serious security problem. In view of the existing security hidden troubles in the network, People have developed many safety precautions, Put forward many solutions to the security problems of information network thought. At present the commonly used network information security technology: Access control technology; Data encryption technology; Firewall technology and so on. The traditional network security technology to protect network security plays a very important role, But they also have many defects, Is still not very well to ensure network security. Intrusion detection system is a newly-developed dynamic network security technology, It is important because it provides the effective network security intrusion detection and take the corresponding protective measure. Intrusion detection system can real-time capture and analysis of data, And through combining and network monitoring system, According to the specific safety data in the database, Through the analysis, can be rapidly found dangerous attack characteristics, And to sound the alarm, At the same time to provide some protection measures. The characteristics of the intrusion detection system is a real-time, dynamic detection and active defense, It can overcome the shortcomings of the traditional network security measures are not enough.

2 The Summary of Intrusion Detection System

Intrusion detection system is a complete intrusion detection function of the software, hardware combinations. Intrusion detection system includes three functional components: Event stream information source; Detect intrusion signs analysis engine; Response response components based on the analysis of the result of engine. Intrusion detection system architecture as shown in figure 1.

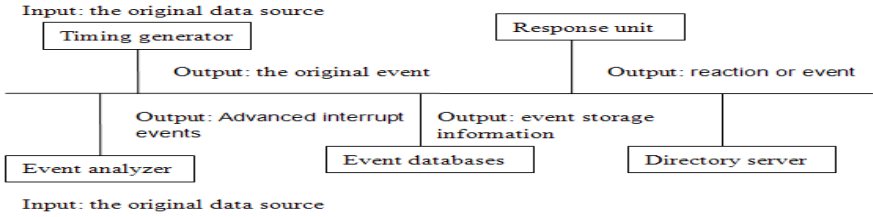


Fig. 1. Intrusion detection system architecture diagram

- (1) The event generator: it is mainly responsible for the collection of primary data, Tracking data, log files etc. Then use the collected raw data into event, to provide this event to the other part of the system.
- (2) Event analyzer: Event analyzer is the analysis of event data and components to transmit its data. For example, analyzes the input event, detect intrusion signs, or to describe the intrusion response response data, can send event analyzer to analyze.
- (3) Event databases: It is responsible for the storage of raw data or has been processed data. It receives the data and save from the event generator or event analyzer, It can be a complex database or a simple text.
- (4) Response unit: It is aimed at the analysis of the results of the analysis component, According to the response strategies of action, Commands are issued in response to the attacks.
- (5) Directory server: Directory server for each component positioning other components, And control other components transfer data certification and the use of other components, In order to prevent the intrusion detection system itself under attack.

3 Intrusion Detection System Architecture and Deployment

3.1 Intrusion Detection System Architecture Design

The architecture of intrusion detection system includes packet capture module、 The data analysis module、 The rule analysis module、 The response module、 Log subsystem module. The data analysis module includes a protocol analysis、

Pretreatment and detection engine three modules. The data capture module access device is realized by means of the data link layer, It is the first part of the intrusion detection system, Its function is from the Ethernet packet capture. Data analysis module is the core of the intrusion detection system, All network data packets in here is analyzed、 pretreatment、 detection, Such as abnormal behavior, Warning to the console, Even respond. The rule analysis module is used to record the custom rules to describe the grammar described the attack characteristics and the corresponding response rule, control event detection engine. The response module includes take the corresponding measure to determine the intrusion behavior, When the system detects the invasion at the time, only through the response to treatment. Log system storage of various security log information, for the invasion behavior of audit and track work in hindsight. Its structure is as shown in figure 2:

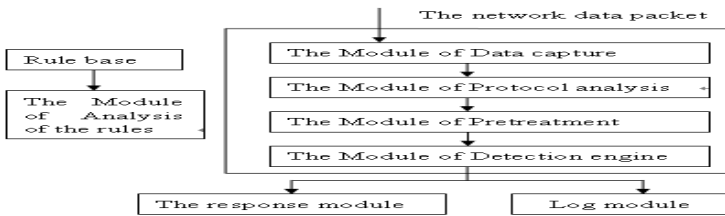


Fig. 2. The structure of intrusion detection system based on Network

3.2 Intrusion Detection System Deployment

For intrusion detection system, Its type is different, different application environment, deployment solution will also be different. We deploy intrusion detection system based on firewall network.the effect of firewall system is defense against the attack from an external network, It can be more effective safety management and intrusion detection system to cooperate with each other. Usually the intrusion detection system is deployed in a firewall in this case, for the two time after time after filtering firewall defense, as shown in figure 3.

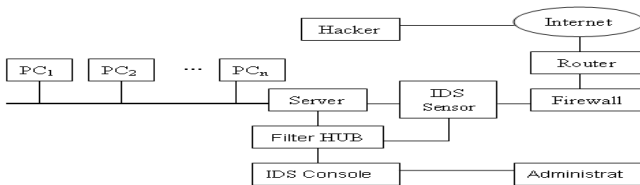


Fig. 3. Network intrusion detection system deployment diagram based on Firewall

4 Intrusion Detection System Based on Pattern Matching

Intrusion detection system based on pattern matching is proposed by Kumar in 1995, at present, pattern matching has become the most widely in the field of IDS intrusion detection means and mechanism. There are a variety of pattern matching algorithm, The typical algorithms including the BF algorithm、BM algorithm and KMP algorithm. BM algorithm is widely used because of its high efficiency.

4.1 The BM Pattern Matching Algorithm Analysis

BM algorithm (Boyer.Moore algorithm) by the Boyer in 1997 and Moore to design a string single pattern exact matching algorithms. The algorithm is generally considered the most efficient string matching algorithm in the application of, Many text editors's " search " and " replace " feature is the use of BM algorithm integrity or simplified version. The next we introduce the BM pattern matching algorithm.

1. The basic idea of BM algorithm

First we use string "P" intrusion feature string, use "T" said detection packet, The length of "P" is set to "m", the length of "T" is expressed as "n" ($n \geq m$), The first character and the last character of "P" are expressed as P_1, P_2, \dots, P_m , The first character and the last character of "T" are expressed as T_1, T_2, \dots, T_n . To make P and T from left alignment, even if the T_1 and P_1 align left, The rightmost characters from P to start match. Comparison of P_m and T_m are the same, If the matching is successful, moves one position to the left, Comparison of P_{m-1} and T_{m-1} are the same, And so on, until P all match or mismatch successful case. If the matching failure occurred in P_j and T_i , at this time to the right P can use a bad character, By using the function $R(x)$ obtained the right moving distance. When the X is in the P, $R(x) = P$ in the right side of the X position; When the X is not in P, $R(x) = m$, Where x is a character in T.

2. BM algorithm analysis

(1) BM algorithm for character comparison is from right to left, While the general string matching algorithm for character comparison is from left to right.

(2) General search string matching algorithm, When the lookup failure, backward movement of a character will target string T then repeated comparison, the BM algorithm can obtain the information of mobile multiple characters according to before the matching fails.

(3) BM algorithm to search the time complexity is $O(m+rn)$. The lookup stage the worst-case time complexity of $O(m * n)$, The best case time complexity is $O(n/m)$.

(4) Because BM algorithm does not consider the matching suffix and cause matching failure current character between the adjacent relation, lead the overall algorithm efficiency is not high.

4.2 Improved BM Algorithm

1. The first of Improved BM algorithm

(1) The basic idea

When match P and T, Consider the P of each character are included in T. If in the P at least one of the characters do not appear in the T, be sure not to match, don't make the following match. Otherwise required BM algorithm for matching. Secondly. If in the P each character appears in the T, Their order the same is it right?

(2) Algorithm description

The following is the algorithm pseudo code:

```
preprocess(char*T, int len)// the preprocessing stage
{
int map[256] ;
for(int i=0 ; i<len ; i++)
map[t[i]]=1 ;
}
search(char*P,char*T,int len_P,int len_T)// search stage
{
for(int i=0 ; i<len_p ; i++)
{
if(map[p[i]]!=1)
return NOT_EXIST ;
}
return BM_New(P,len_P,T,len-T) ;
}
```

As can be seen from the above pseudo code, the preprocessing stage time complexity and data packet size of the load in a linear relationship, phase time complexity and all pattern length and linear relationship. The improved algorithm can avoid the character check one by one, reduce unnecessary comparisons, can effectively improve the efficiency of operation.

2. The second of Improved BM algorithm

(1) The basic idea

The BM algorithm is based on the text string corresponding to the pattern string the final location of the character information to determine the offset,so the maximum offset is M. In order to improve the offset as the starting point in the improved algorithm, To determine the offset press text string corresponding to the pattern string

last character position of the next character information, Sliding median maximum increased to $m+1$ when they don't match. In the algorithm design, Based on BM algorithm, only using the bad character of suffix rules, calculation of skip values of an array. Consider the next character in calculation of skip array, The right amount of decision through the use of a character.

(2) Algorithm description

Int*improve_sk ip(char*ptrn,int len_P)/* the function is calculation of the skip array program*/

```
{
int i, *skip=(int*)malloc(256*sizeof(int)) ;
for(i=0 ; i<256 ; i++)skip[i]=len_P ;
for(i=0 ; i<len_P-1 ; i++)
skip[ptrn[i]]=len_P-i-1 ;
return skip ;
}
void N_BM(char*P, char*t, int ship[])/* The function is BM improved algorithm */
/
{
for(i=0 ; i<maxchar ; i++)
skip[i]=m+1 ;
for(i=0 ; i<m ; i++)
skip[p[i]]=m-i :
k=m-1 ;
while(k<n)
{
J=m-1 ;
while(j>=0)&&(p[j]==t[k+j-(m-1)])
{
J-- ;
}
If(j==0)
return t[k] ;
else
k=k+ship ;
}
```

The BM algorithm's time complexity of $O(m, n)$ in the stage of searching, but the improved algorithm of time complexity is $O(m(N, m))$ in the stage of searching, on the basis of the original pattern string mobile, Accelerated to jump back in the amplitude of the match after failure, thereby reducing the number of comparisons, Improve operational efficiency, Especially the efficiency is more obvious in the string is more complex cases.

5 Summary

Intrusion detection is the access control, firewall and other traditional security protection technology after a new generation of protection technology, It is the reasonable supplement to traditional security protection technology. But current intrusion detection technology are far from mature. The traditional packet capture mechanism and detection engine, It is difficult to solve in high-speed network environment of tundish and false alarm problem. This article has conducted the thorough research on the BM pattern matching algorithm, put forward higher efficiency of improved BM algorithm. Intrusion detection system can effectively improve the matching efficiency not high question in a high-speed network.

Acknowledgements. Hunan science research foundation on high education (11C0746).

References

1. Fu, X., Yu, W., Cheng, D., et al.: On Recognizing Virtual Honeypots and Countermeasures. In: The 2nd IEEE International Symposium on Dependable, Autonomic and Secure Computing, vol. 9, pp. 220–230 (2006)
2. Leita, C., Mermoud, K., Daeier, M.: ScriptGen:all automated script generation tool for honeypot. In: 21st Annual Computer Security Applications Conference, vol. 9, pp. 125–135 (2008)
3. Zhang, F., Zhou, S., Qin, Z., et al.: Honeypot:a supplemented active defense system for network security. In: Proceedings of the Fourth International Conference, PDCAT 2009, vol. 8, pp. 231–235 (2003)
4. Kreibich, C., Crowcroft, J.: Honeycomb-Creating intrusion Detection Signatures Using Honeypots(EB/PDF) (2011), <http://www.sigcomm.org/I-IotNets-II/papershaoneycomb.pdf>.
5. Domscif, M., Holz, T., Mathes, J., Weisemoller, I.: Measuring Security Threats with Honeypot Technology, pp. 21–29 (2009)
6. Yah, L.K.: Virtual honeynet revisited. In: Proceedings from the Sixth Annual IEEE SMC Information Assurance Workshop, vol. 9, pp. 230–240 (2010)

Modeling and Analysis of DNA Mutation Effects on Protein Structure with Finite State Machine

Rui Gao¹, Wen-Song Hu¹, and Cheng-Qiu Zhang²

¹ School of Control Science and Engineering, Shandong University, P.R. China, 250061

² Shandong University Hospital, P.R. China, 250061

{Gaorui, zcqhuxx}@sdu.edu.cn, wensonghusdu@yahoo.com.cn

Abstract. This paper extends our early study on discrete events system formulations of DNA hybridization, and focuses discussions on gene mutation in Molecular Biology. Finite state machine (FSM) theory is extensively applied to express the key concepts and analyze the processes related to the biological phenomena mentioned above. The goal is to mathematically represent and interpret metabolism and the effects to structures of protein macro molecule caused by gene mutation. We hope the proposed model will provide a foothold for introducing the systems science and the control theory tools in Molecular Biology.

Keywords: Finite state machine, mathematical formulation, gene mutation, Molecular Biology.

1 Introduction

A mathematical model is helpful in understanding the theoretical aspects of genetic information transmission, and also useful to solve DNA computation problems with mathematical tools. With the presented models, character-based DNA computation is converted into a numerical computation problem. Related propositions about DNA hybridization have also been presented [1][2]. In 2009, the author and his partners proposed a unified mathematical model of the Central Dogma of Molecular Biology [3], which built the state-space model for the experimental biochemical process. Based on the modeling work, the author modeled protein molecular second structure, torsion angle in amino acids along polypeptide with matrix group, and gave the simple process model of DNA duplication with the FSM.

Another important model in Molecular Biology is hidden markov models (HMM). HMM is a statistical model that considers all possible combinations of matches, mismatches, and gaps to generate an alignment of a set of sequences. These models are primarily used for protein sequences to represent protein families or sequence domains, but they are also used to describe patterns in DNA sequences, such as RNA slice junctions. In 1987, Lander and Green used HMM in the construction of genetic linkage maps [4]. In 1989, Churchill employed HMM in sequence analysis to produce an HMM that represented a sequence profile (a profile HMM) to analyze sequence composition and patterns [5]. In the 1990s, HMMs have been widely used to model sequences and proteins. White *et al.* used HMM to model super families of proteins [6]. Asai *et al.* applied HMMs to predict the secondary structure of proteins, and obtain higher prediction rates than previously achieved [7]. HMMs are also used in

producing multiple sequence alignments [8][9][10][11]. These HMMs generate sequences with various combinations of matches, mismatches, insertions, and deletions, and give these sequences probabilities depending on the values of the various parameters in the models. In 2010, the author applied FSM to analyze the process of DNA hybridization, and then got the criteria about the perfectibility of DNA hybridization [12]. Based on the FSM discussed above, we will study the process of gene mutation of DNA strands in this paper. Then we will further discuss the effects on protein molecular structures caused by DNA mutation.

This paper is organized as follows. In Section 2, a new FSM model is proposed to judge if the mutation happens and analyze the specific effects on protein molecular structures. The conclusions are presented in Section 3.

2 FSM Model Analysis of Effect on Protein Molecule Caused by Gene Mutation

Based on the Central Dogma [13], the genetic information stored in the bases of DNA strand is transferred to protein molecules through mRNA. The DNA strand same as the mRNA strand (except for *T* and *U*) is called the coding strand. The other DNA strand that directs mRNA strand duplication by the base-pairing principle is called the template strand.

Every three nucleotides on mRNA strand are translated into an amino acid along polypeptide. The three nucleotides are called the genetic code, or a codon [1]. The translation begins from the start codon *AUG*, successively reads the codons in the direction of $5' \rightarrow 3'$, and then stops at the stop codon. So the polypeptide with specific nucleotide sequence is generated. The amino acids in the new polypeptide and its order are decided by the bases of the original DNA strand. As a result, the produced proteins are controlled by the genes.

There are 20 classes of amino acids and 64 genetic codes in general [2]. So one kind of amino acid may correspond to one or several genetic codes, which is called the synonymous codon.

For analyzing the effect on the protein molecular structures caused by DNA mutation thoroughly, we present the following FSM model (see Fig. 1).

There are $m+2n+12$ states in the model, where m, n are decided by actual requirements. The states are divided into two classes. One class of states including $S_{4,k}, S_{11,l}, S_{12,l} (k=1,2,\dots,m; l=1,2,\dots,n)$ are the analyzing states, and the others are transmission states. If there is no DNA strand, mRNA strand or base in a certain state, the state is defined as *NULL*. We require the direction of reading strand as follows. If the read DNA strand is the coding strand, the direction of reading is $5' \rightarrow 3'$. Otherwise, if the read DNA strand is the template strand, the direction of reading is $3' \rightarrow 5'$.

Initialization Stage. In this stage, the DNA strand is read in the FSM model. All of the counters are reset. Counter 2 and Counter 4 are invalid in this stage. The transitions $T_{3,k}, T_{4,k}, T_{10,l}, T_{11,l} (k=1,2,\dots,m; l=1,2,\dots,n)$ will be enabled just once in order. The transition T_{12} is invalid. The transition $T_{6,1}$ is enabled under the condition that all bases are transferred into S_6 and stored in sequence. There are two steps in this stage.

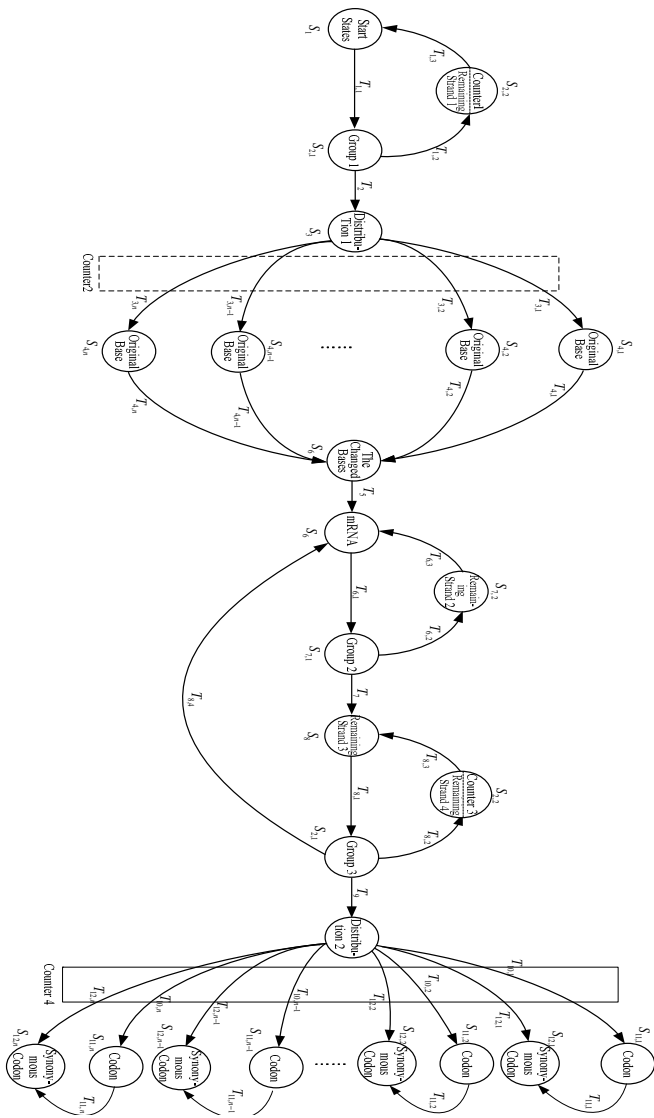


Fig. 1. FSM model of DNA mutation effects on protein structure

1. When a DNA strand is read in S_1 , the transition $T_{1,1}$ is enabled. Then the state $S_{2,1}$ chooses the first base of the DNA strand, and keeps the remained strand in it. If the state $S_{2,1}$ is not *NULL*, the transitions $T_{1,2}$ and T_2 are enabled. The remained DNA strand is transferred into the state $S_{2,2}$, and the first base is transferred into S_3 . In $S_{2,2}$, there is Counter 1. When the transition $T_{1,2}$ is enabled, Counter 1 is added by 1. The remained DNA strand is transferred into the state S_1 through the enabled

transition $T_{1,3}$. The base in S_3 enables the according transition $T_{3,1}$. The first base is transferred into the state $S_{4,1}$, and then enables the transition $T_{4,1}$. Here the transition $T_{4,1}$ acts as a convertor, and the transformation rules are stated as the follows. If the DNA strand is the coding strand, then T is converted to U and the other three bases are maintained. On the other hand, if the DNA strand is the template strand, then the transformation rules are that A is converted to U , T is converted to A , G is converted to C , and C is converted to G . The state $S_{4,1}$ keeps the first base and transfers the converted bases into the state S_5 . The base in S_5 enables the transition T_5 , and is transferred into the sequence in S_6 . When the base in $S_{4,1}$ is converted into the state S_5 , the remained strand in S_1 enables the transition $T_{1,1}$ again, and is transferred into the state $S_{2,1}$. In the case that the state $S_{2,1}$ is not *NULL*, $S_{2,1}$ chooses the first base of the remained strand, transfers it to the state S_3 through the transition T_2 , and transfers the remained strand to the state $S_{2,2}$ through the transition $T_{1,2}$ at the same time. Counter 1 is added by 1. The remained strand in $S_{2,2}$ enables the transition $T_{1,3}$, and is transferred into the state S_1 . The base in S_3 enables the transition $T_{3,2}$, and is transferred into the state $S_{4,2}$. The base in $S_{4,2}$ is converted and transferred into the state S_5 through the enabled transition $T_{4,2}$. At the same time, the state $S_{4,2}$ keeps the original base before conversion. The base in S_5 enables the transition T_5 , and is transferred into the sequence in S_6 . The first step of the initialization stage will stop till the last base of the DNA strand is read, converted and stored in S_6 , where we note that m is not less than the number of the read DNA strand.

2. After the first stage, there is the mRNA strand corresponding to the read DNA strand in S_6 . The mRNA strand in S_6 enables the transition $T_{6,1}$, and then transferred to $S_{7,1}$. The state $S_{7,1}$ chooses the first three bases of the strand, keeps the remained strand, and judges if the first three bases are the start codon. If not, the first base of the mRNA strand is erased, and enables the transition $T_{6,2}$. The remained strand is transferred into the state $S_{7,2}$, and then transferred into the state S_6 through the enabled transition $T_{6,3}$. The process repeats until the first three bases of the remained strand are the start codon. Consequently, the start codon is deleted and the transition T_7 is enabled. The remained mRNA strand is transferred into the state S_8 . The sequence in S_8 enables the transition $T_{8,1}$, and is transferred into $S_{9,1}$. The state $S_{9,1}$ chooses the first three bases in the sequence, keeps the remained sequence, and judges whether the three bases are the stop codon. If not, the transitions $T_{8,2}$ and T_9 are enabled, the first three bases are transferred into the state S_{10} as a codon, and the remained strand is transferred into the state $S_{9,2}$. There is a counter, Counter 3 in $S_{9,2}$. When $T_{8,2}$ is enabled, Counter 3 is added by 1. The remained strand in $S_{9,2}$ enables the transition $T_{8,3}$, and then is transferred into the state $T_{8,3}$. The three bases in S_{10} enable the transition $T_{10,1}$, and are transferred into the state $S_{11,1}$. The codon in $S_{11,1}$ enables the transition $T_{11,1}$, and is transferred into the state $S_{12,1}$, which stores all the synonymous codons with the codon in $S_{11,1}$. For example, we suppose that the codon in $S_{11,1}$ is *AUU*, corresponding to *Leucine* (refer to Fig. 4) we find that there are three codons that describe *Leucine*, which are *AUU*, *AUC* and *AUA*, respectively. Then the codon set in $S_{12,1}$ is $\{AUU, AUC, AUA\}$. At the same time when the transition $T_{11,1}$ is enabled, the remained strand returning to S_8 enables the transition $T_{8,1}$. Repeat the steps mentioned above, till the first bases of the remained strand in $S_{9,1}$ are the stop codon. Then we erase the stop codon and enable the

transition $T_{8,4}$. The remained strand is moved to the state S_6 through the transition $T_{8,4}$. The operation will stop till the number of the remain strand is zero. When all the states in front of the state S_{12} , except for the state S_4 , are all *NULL*, the initialization stage ends. Thus, the effective fragment of gene along the mRNA strand used to code protein is extracted. With the value of Counter 1, the length of the read DNA strand, i.e., the number of the bases along the DNA strand, can be obtained. The length of the effective fragment of the mRNA that makes up the protein coding, i.e., the number of the codons along the mRNA strand, can also be procured with Counter 3.

Analysis Stage. In this stage, the new strand after metabolism is read in the FSM model. This stage judges whether there are mutations happen along the new generated DNA strand and whether the mutations change the structure of the protein molecules. The transitions $T_{3,k}$, $T_{4,k}$, $T_{12,l}$ ($k=1,2,\dots,m;l=1,2,\dots,n$) can be enabled once orderly, and the transitions $T_{10,l}$ and $T_{11,l}$ ($l=1,2,\dots,n$) are invalid. Moreover, the transition $T_{6,1}$ will be enabled when satisfied with the following two conditions. One is that all the bases along the new generated strand have been converted and have been stored in S_6 , namely, all the transitions in front of the transition $T_{6,1}$ are not enabled. The other one is that the value of Counter 2 is nonzero. To distinguish the counter state in the two stages, Counter 1 in the initialization stage has been denoted as Counter 1', and in the current stage all the counters are reset. The stage can be divided into two steps as follows.

1. When the new strand is read in the FSM model, the running process from the state S_1 to state S_3 is similar to that of the initialization stage. The first base of the new strand enters the state S_3 and then enables the transition $T_{3,1}$. Here the transition $T_{3,1}$ executes a comparison operation. If the base in S_3 is the same as the base in $S_{4,1}$, the base in $S_{4,1}$ will be maintained and enable the transition $T_{4,1}$, which acts as an convertor. Same as the initialization stage, the converted bases are stored in S_5 . The base in S_3 enables the transition T_5 , and then is transferred into the sequence in S_6 . If the base in S_3 is not same as the base in $S_{4,1}$, Counter 2 is added by 1, and the base is transferred from the state S_3 to $S_{4,1}$. The base in $S_{4,1}$ enables the transition $T_{4,1}$, and enters the state S_5 after conversion. The base in S_5 enables the transition T_5 , and then transferred into the sequence in S_6 . Next, the second base of the new strand enters the state S_3 and enables the transition $T_{3,2}$. The comparison between the base in S_3 and that in $S_{4,2}$ is executed. The corresponding actions are similar to the above mentioned, till the last base of the new strand converts, and enters the sequence in S_6 . The first step in the initialization stage is completed. Now, based on the value of Counter 2, we can judge whether the mutation happens along the DNA strand after the metabolism. If the value of Counter 2 is zero, there is no mutation happens, and all the transitions behind T_5 are invalid. The FSM will end at the state S_6 . If the value of Counter 2 is nonzero, there must be mutations happening along the new generated strand. The FSM will continue to run but the length of the DNA strand is not changed. The ratio of the value of Counter 2 and that of Counter 1 is the practical mutation ratio.

2. The duty of this step is to judge if the mutation changes the corresponding protein molecular structures when the mutation happens. If the value of Counter 2 is nonzero after the first step of the initialization stage, namely, there is mutation happens along the new generated strand, the transition $T_{6,1}$ is enabled and all the transitions in front

of the state S_6 are invalid. When the transition $T_{6,1}$ is enabled, the running process from the state S_6 to S_{10} is similar to the second step of the initialization stage. The first codon entering the state S_{10} enables the transition $T_{12,1}$, and $T_{12,1}$ executes a comparison operation and judges whether the codon in S_{10} belongs to the state $S_{12,1}$. If not, Counter 4 is added by 1. Otherwise, the value of Counter 4 remains the same. Then the second codon entering the state S_{10} enables the transition $T_{12,2}$, and $T_{12,2}$ executes a comparison operation and judges whether the codon in S_{10} belongs to the state $S_{12,2}$. The process will continue till the first three bases of the remained strand in $S_{9,1}$ are the stop codon. The state $S_{9,1}$ changes into *NULL*. When the valid codon in front of the stop codon completes the comparison operation, the step ends. Till now, based on the value of Counter 4 we can judge whether the mutated, new generated DNA strand affects the corresponding protein molecular structures. If the value of Counter 4 is zero, the new generated strand may not change the protein molecular structures though the mutation happened. Otherwise, if the value of Counter 4 is nonzero, the mutated strand changes the corresponding structures. However, there are two special cases. One case is that there is no start codon measured in $S_{7,1}$, and the other one is that there is no start codon measured in $S_{9,1}$. In either of the cases mentioned above, the value of Counter 4 is set to ∞ .

The above-mentioned three examples demonstrate the running process of the FSM models. When DNA mutation happens, though the new generated protein molecular structures are not changed through the FSM model, the protein synthesis may still be affected. In addition, the model proposed here is not applicable to polycistron mRNA transcribed from recessive genes.

Here we notice that the models mentioned above are useful in diagnosis of diseases caused by gene mutation. For example, hemoglobin of a healthy person is a quartet composed of two DNA strands and the other two DNA strands, which interacts among them. There are 141 and 161 amino acids arranged in order along the DNA helix, respectively. In the patient with sickle cell anemia, the sixth amino acid has changed. In the healthy, the sixth amino acid is *Glutamic acid*, but in the patient with sickle cell anemia it is *Valine* instead. That is, the codon *CTT* is changed into *CAT* along the decoding strand of the DNA helix. It causes disease. If the DNA strand that guides hemoglobin synthesis is read in the FSM model, the read strand can judge whether the mutation happens. If so, the effect on the structure of the protein molecule can be further analyzed and then the disease is diagnosed.

3 Conclusions

The FSM model descriptions of gene mutations have been presented in this paper. Based on the above-mentioned model, the processes of gene mutations have been analyzed, and whether the mutations happen and bring many changes on protein structures has been judged. Also the control methods that guide the gene mutation have been proposed theoretically. It hopes that the main results of this paper can provide a sound basis for analyzing the related biological process.

Acknowledgment. This research was supported by the National Natural Scientific Foundation under grant 61174218, the Independent Innovation Foundation of Shandong University under grand 2010TS016, the Natural Science Foundation of Shandong Province under grant ZR2010FM030, the Outstanding Young Scientists Incentive Funding of Shandong Province under grant BS2009DX041.

References

1. Zhang, M.J.: On constructing a molecular computer. In: Lipton, R., Baum, E. (eds.) *DNA Based Computers*, Providence, RI: Amer. Math. Soc. ser. Discrete Mathematics and Theoretical Computer Science (DIMACS), pp. 1–21. Amer. Math. Soc., Providence (1996)
2. Zhang, M., Cheng, M.X., Tarn, T.J.: A mathematical formulation of DNA computation. *IEEE Trans. on Nanobioscience* 5, 32–40 (2006)
3. Gao, R., Yu, J.Y., Zhang, M., Tarn, T.J., Li, J.S.: A Mathematical Formulation of the Central Dogma of Molecular Biology. In: Zhang, M., Xi, N. (eds.) *Nanomedicine: A System Engineering Approach*, pp. 81–116. Pan Stanford Publishing Pte. Ltd, Singapore (2009)
4. Lander, E.S., Green, P.: Construction of multilocus genetic linkage maps in humans. *Proceedings of the National Academy of Sciences of the United States of America* 84, 2363–2367 (1987)
5. Churchill, G.A.: Stochastic models for heterogeneous DNA sequences. *Bull. Math. Biol.* 51, 79–94 (1989)
6. White, J.V., Stultz, C.M., Smith, T.F.: Protein classification by stochastic modeling and optimal filtering of amino-acid sequences. *Math. Bios* 119, 35–75 (1994)
7. Asai, K., Hayamizu, S., Onizuka, K.: HMM with protein structure grammar. In: *Proceedings of the Hawaii International Conference on System Sciences*, pp. 783–791 (1993)
8. Eddy, S.R.: Hidden Markov models. *Curr. Opin. Struct. Biol.* 6, 361–365 (1996)
9. Eddy, S.R.: Multiple alignment using hidden Markov models. In: *ISMB*, vol. 3, pp. 114–120 (1995)
10. Baldi, P., Chauvin, Y., Hunkapillar, T., McClure, M.A.: Hidden Markov models of biological primary sequence information. *Proc. Natl. Acad. Sci.* 91, 1059–1063 (1994)
11. Krogh, Brown, M., Mian, I.S., Sjölander, K., Haussler, D.: Hidden Markov models in computational biology. Applications to protein modeling. *J. Mol. Biol.* 235, 1505–1531 (1994)
12. Gao, R., Yu, J.Y., Zhang, M.J., Tarn, T.J., Li, J.S.: Systems Theoretic Analysis of the Central Dogma of Molecular Biology: Some Recent Results. *IEEE Trans. on NanoBioscience* 9(1), 59–70 (2010)
13. Crick, F.H.C.: Central dogma of molecular biology. *Nature* 227, 561–563 (1970)
14. Crick, F.H.C.: On protein synthesis. In: *Symp. Soc. Exp. Biol.*, vol. XII, pp. 139–163 (1958)

Design and Implementation of Solar Thermal Utilization Information System Based on GIS

Yu Zhang, Shijun He^{*}, Shutai Zhang, Zhonghua Chen, Wenjun Zhou,
Yanyan Dai, Shiting Zhao, and Fan Bai

College of Information Technology
Shanghai Ocean University
Shanghai, China
sjhe@shou.edu.cn

Abstract. The establishment of solar thermal utilization information system is still in the exploratory phase and the data of solar engineering was not taken sufficient use. In this paper, geographic information system was integrated into management information system by Visual C++. The solar thermal utilization information system was implemented with the software architecture of data layer/ business logic layer/ presentation layer. The graphical user interface of the information system was based on MapObject component technology of GIS. Data storage and management were realized through the ADO accessing way of SQL Server database. The fusion of spatial data and attribute data was completed by ArcSDE. And aided decision models were calculated by MATLAB. This method to design the solar thermal utilization information system is effective to apply computer technology, database technology, GIS technology and scientific theories of solar thermal utilization into solar energy engineering, and it has certain practical significance.

Keywords: Solar, Information System, GIS, Data Management, Aided Decision.

1 Introduction

Solar thermal utilization systems include solar water heating system, solar greenhouse, solar refrigeration system, solar drying, solar cookers, etc. Due to the utilization of solar thermal started late and the technology is very immature, solar engineering data most only has simple forms to storage, such as reports and atlas. That was not convenient for data management and the information was not made full use of. The expression of data information was not intuitive, and the link of data and graph was not closely enough, so that managers could make no decision facing these data and reports.

^{*} Corresponding author: Shijun He (1965-), male, Zhengzhou, Henan, China, Professor. Research fields: Marine information processing, sensor and digital technology, Digital Ocean.

Application of geographic information system technology in information system can be effectively integrated different terrain, physiognomy, natural resources, social environment and other information into the information system. References [1-4] has given the cases of information systems base on different technologies, such as VC++, SQL Server, MapX, GIS component technology, Visual C# and ArcGIS Engine.

In general, the design of new information system requires making the full use of the existing management information system, and integrating of spatial information and the aided decision making information. Therefore, reasonable software architecture is very important. References [5-6] implemented information systems based on different three-layer software architectures.

The development of solar thermal utilization information system (STUIS) needs a variety of software to realize the calculation model, results display, data storage, data management and other functions. By using mixed programming technology, the communication of different software and interface technology were implemented in reference [7]. In this paper, the GIS technology is applied into the solar thermal utilization information system, and so are the three hybrid software architecture based on B/S and C/S and the interface technology.

2 Requirements and Objectives of STUIS

2.1 General Objective of STUIS

STUIS database is built by data resources which are spatial information, attribute information and documentation. GIS component technology is applied into the solar energy engineering and the spatial and non-spatial information related to solar thermal utilization, such as product introduction, energy consumption query and energy saving evaluation, is effective in acquire, storage, update, management, analysis and display. STUIS can provide modern technology tools for solar thermal utilization system and informationize the process of solar energy thermal utilization and make it convenient for users to obtain the information of the system, such as water temperature of the heat storage tank and solar irradiance. At the same time, the system manager can effectively control the solar energy collecting devices, and it will do well to energy saving and emission reduction. The final purpose is to realize informatization and cyberization of solar thermal utilization system, and to apply better services for users.

2.2 Requirements of System Functions

STUIS provides the basic map operations functions, such as amplification and narrow. In addition, it still can be achieved to query the information of the device running states and related service information, such as water temperature, supplying time and consumption records, through different ways.

Geographic information database, a variety of thematic maps and attribute database of solar thermal utilization region, is needed to realize scientific management of spatial data and non-spatial data. Attribute database is based on the information acquisition by solar energy devices. GIS functions are necessary, for example, thematic information extraction, thematic map drawing and statistical analysis. It should be available to provide aided decision making, equipment and meteorological element information query, energy saving scheme recommended and other information services. Moreover, through the internet, users can also get access to the system and obtain the information of energy consumption and energy saving.

3 System Design

Operating system: Windows XP; database management software: Microsoft SQL Server; secondary development software: ESRI MapObject module of GIS; visual programming environment of system integration: Visual C++, C#.

3.1 Database Design

ArcGIS Geodatabase data model is used to design and establish the geographic information database. Throughout ArcSDE, the solar thermal utilization information data is organized and managed by SQL Server, a large-scale relational database.

Spatial Database. Spatial database stores basic geographic information, such as latitude, longitude, and altitude. Through the developed GIS interface, information query and management of a point on the map can be achieved.

Attribute Database. Attribute database include device information table, system status information table and device information table. The relationships between the tables include foreign key constraint and integrity constraint. Device information table provides the basic attribute information of solar energy devices. System status information table gives the real-time running process data of the device parameters. Device information table provides service information query such as the user's consumption. The auxiliary energy consumption in a certain period of time can be calculated by the meter scale in Table 3. The collected thermal can be calculated by the meteorological parameters in Table 3. Then energy saving is carried out. And according to the energy-saving evaluation and analysis results, reasonable energy saving recommended scheme could be given.

3.2 Function Modules Design

As shown in Fig.1, STUIS includes three subsystems and mainly orients to information processing and information services.

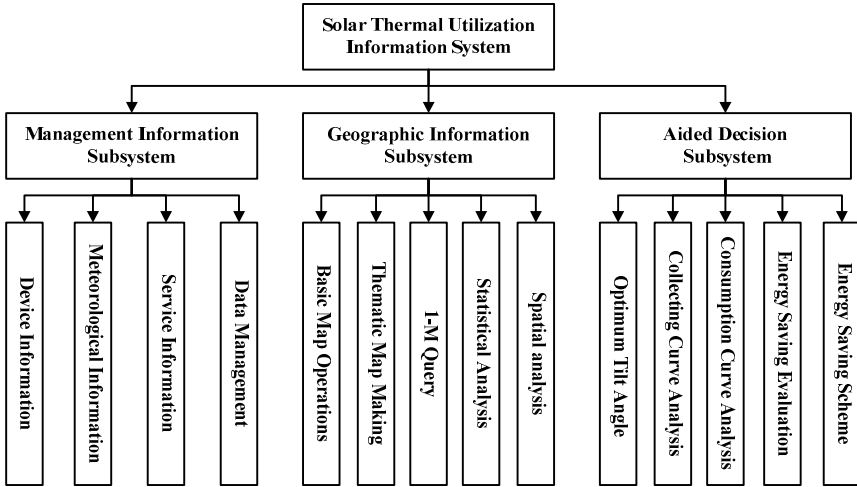


Fig. 1. Functional Framework of STUIS

Management Information Subsystem. The subject of this subsystem is the relational database management system. The operations are add, delete, update, search and other operations of attribute database. The functions include device information query, meteorological element information query, service information query and data management. Device information includes the device type, device name, thermal performance parameters and running state. Meteorological elements include solar irradiation, temperature, wind speed, solar hour angle etc. Service information includes of supplying time, consumption records etc. Data management includes data insert, delete, modify, backup, restore and print output.

Geographic Information Subsystem. The subject of this subsystem is vector processing. The operations are basic geographic information edit, query, analysis, management and other operations. The functions include basic map operations functions, thematic mapping, 1-M query, statistical analysis and spatial analysis. The numbers accorded to lemmas, propositions, and theorems, etc. should appear in consecutive order, starting with Lemma 1, and not, for example, with Lemma 11.

Aided Decision Subsystem. The subject of this subsystem is data analysis. The operation is to calculate some data in relation tables and accomplish the results analysis. The functions include calculation of the optimum tilt angle of solar collector, energy-saving effect evaluation, collecting curve analysis, energy consumption analysis and energy saving scheme recommended.

3.3 Software Interface Design

Data sharing of different software platform is through a variety of interfaces between them. [8] ADO (ActiveX Data Object) is a popular client database access technology in VC++, and it provides three basic interfaces: `_ConnectionPtr`, `_CommandPtr` and `_RecordsetPtr`, which are for users to call. Communication between VC++ and SQL Server can be achieved by ADO. Calling the MATLAB engine can realize the connections between VC++ and MATLAB. Thus, it will be helpful for the operations of calculation models. With GIS component technology, MapObjects can be directly embedded into the VC++ development environment. In addition, relevant service publication can be completed by Visual C# with the mode of B/S.

4 Implementation of Key Technology in STUIS

Implementation of STUIS requires multiple development software. In this paper, SQL Server is used to establish and maintain the database and implement the basic data storage and management. MATLAB is used for model structures and calculation. GIS component technology is applied in VC++ environment to develop a friendly graphical user interface. The calculation, display and data access of solar thermal utilization information can be efficient through the GUI, and resource sharing and data fusion between multiple software are achieved through the platform.

4.1 Three-Layer Hybrid Software Architecture

The three-layer hybrid architecture, which is combined C/S and B/S, consists of data layer, business logic layer and presentation layer. (As shown in Fig.2) Specific application requirements and user interactions are in presentation layer. Business logic layer is responsible for the presentation layer and transform user needs into basic functions of a specific server. Data layer is used to store data and provide basic data services.

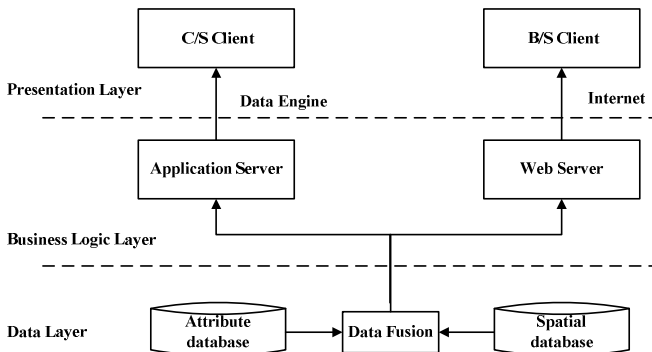


Fig. 2. Three-layer Hybrid Software Architecture

By C/S, administrator can carry out data management and upload to server according to the actual situation. General users can directly obtain the information services provided by B/S server through network browser, and give personal feedback advices. Generally speaking, C/S is mainly aimed at the data management and update for administrator, and B/S is mainly used to release information services for general users.

4.2 Data Fusion Technology

In the development process, the user often requires making full use of attribute data in existing MIS. So the coupling of GIS and MIS is very important. With component technology, GIS is integrated into MIS, and resources share become effective.

ArcSDE, a spatial database engine which is provided by ESRI, is a GIS channel between ArcGIS and relational database. ArcSDE for SQL Server provides a database management mode, which takes ArcSDE as an engine and SQL Server as the central database. In this mode the spatial data and attribute data are in the same database environment. Table connection is complete through a foreign key, and the data fusion of spatial data and attribute can be realized.

4.3 Calculation Engine Technology

The MATLAB engine consists with a group of functions. Through these functions, users can complete the calculation and drawing tasks by MATLAB. That is to say, MATLAB is taken as a calculation engine. SIMULINK, a kit of MATLAB, can be seamlessly combined with MATLAB. The interface between VC++ and MATLAB can be used to operate the SIMULINK model. In this way, SQL Server is used to complete the data storage and management, and the work that is numerical calculation, virtual simulation and decision support is left to MATLAB. Through the interaction of MATLAB and SQL Server, powerful numerical analysis, simulation, modeling and other functions of MATLAB are taken to achieve complementary advantages.

4.4 Component GIS Technology

ArcGIS not only has strong and stable GIS basic functions, but also provides rich COM objects, which can be used by VC++. MapObjects are a group of GIS function components supplied for the developer of drawing maps. It is composed of a plurality of controls and a series of programmable objects, which can be inserted directly into VC++. With the help of VC++, all the process and data of the system will be under the designer's control, and GIS functions can be added according to the requirements of the system. At the same time, it will present fast running speed and good reliability.

5 Conclusions

The design and implementation of STUIS needs kinds of development software, and the interfaces design is the key factor of information sharing between different software platforms. Solar engineering data includes spatial data and attribute data.

With the help of component GIS technology, geographic information system is integrated into the management information system, and the data fusion of spatial data and attribute can be realized. By developing three-layer hybrid software architecture (presentation layer/business logic layer/data layer), administrator users manage data through the C/S mode, and general users browse services through the B/S mode. Database management software is used to complete the data storage and management, and the work of decision-making model calculations is left to calculation engine. So the complementary advantages of different software are achieved. Design and implementation of solar thermal utilization information system has practical significance for digital and information-based development of solar energy engineering.

Acknowledgments. Financial support for this work provided by the municipality subject of science and technology commission of shanghai, contract NO: 10510502800, is gratefully acknowledged.

References

1. Xiu, W., Chi, T.: City Geographic Information System (GIS). Beijing Hope Electronic Press, Beijing (2001)
2. Lu, Y., Wang, X., Wu, H.: The Reserch and Realization of Honghu Marsh Protection and Ecological Restoration Inspect Information System Based on GIS Component Technology. *Journal of Central China Normal University (Natural Sciences)* 39(1), 141–144 (2005)
3. Cai, B.: Study and Implementation of Tourism in Formative System Based on GIS City. *Shanxi Architecture* 34(27), 348–349 (2008)
4. Zou, J., Sun, G.: Research and Realization of Town Cadastral Management Information System Based on the ArcGIS Engine and C#. *Urban Geotechnical Investigation & Surveying* 4, 29–31 (2009)
5. Gao, Q., Guo, S.: Research and Implementation of School Education Information System Based on the Campus Network. *Science & Technology Information* 31, 266–267 (2007)
6. Liu, T., Zhang, X., Huang, Q.: Research and Realization of Water Resource Information System Based on Hybrid Software Architecture. *Acta Scientiarum Naturalium Universitatis Sunyatseni* 46(1), 133–135 (2007)
7. He, S., Zhang, Y., Yu, Y.: Application of ArcGIS/MATLAB/ SQL Server/ VC++ Mixed Programming in Solar Energy Engineering. *Advanced Materials Research* 171-172, 323–328 (2011)
8. He, S., Zhang, Y., Zhou, W.: A Method to Calculate the Optimum Tilt Angle of Solar Collector. *Acta Energiae Solaris Sinica* (in Press)

Research of Lip Contour Extraction in Face Recognition

Kunlun Li¹, Ruining Xin¹, Miao Wang¹, and Hui Bai²

¹ College of Electronic and Information Engineering, Hebei University, Baoding 071002, China

² Software School, Fudan University, Shanghai, 201203, China

likunlun@hbu.edu.cn, xinruining1986@163.com, hbhdwm800@sina.com

Abstract. The lip contour extraction is widely used in facial expression classification, lip recognition, fatigue driving test and so on. Its research and development not only has great theoretical significance, but also has wide application prospect and practical significance. In this paper, we study the lip contour extraction problem in face recognition. According the practical difficulties in traditional lips contour extraction method, we propose two improved level set methods based on Fisher transformation and semi-supervised learning algorithm for lip contour extraction. Experiment results prove the effectiveness of this method.

Keywords: Lip contour extrac, Fisher transformation, Semi-supervised learning, Level set method.

1 Introduction

Lip contour detection is a challenging and important problem in computer vision due to the variation human's expressions and environmental conditions. It can be classified in three categories. The first category of techniques is threshold based method [1-2], which enjoys a central position in applications of image segmentation because of its intuitive properties. The second category is edge and line oriented approaches [3-5], in which lines or boundaries defined by contrast in luminance, color or texture are detected. The last category is hybrid approaches [6], which aimed at consistency between regions and region boundaries. There is a very large of techniques for each specific class of methods.

In this paper, we compare the three categories lip contour detection method and propose a new method combines the semi-supervised learning method, and then we present some experiment results.

2 Contour Detection

The contour detection method mainly classified in three categories: threshold based method approaches, edge and line oriented approaches and LDA algorithm approaches. In this paper we compare the three categories method, and we present a hybrid approach combines with level set method and semi-supervised LDA algorithm for the lip contour decoction.

2.1 Threshold Based Method

Image threshold segmentation is a widely applied segmentation technology. In this method, the image is looked as the combination of the target area and background region.

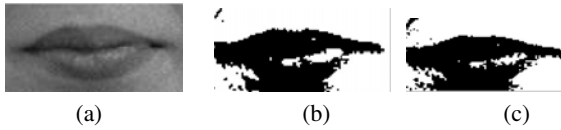


Fig. 1. (a) Gray image (b) iteration threshold binary image (c) Otsu threshold method

Through the experiments we can see that the results of the threshold method is affected for the difference between grayscale information of lip color and that of skin color is small.

2.2 Edge and Line Oriented Approach

1). Local Contour Detector

a) Differential methods

The earliest linear filtering approach such as the Sobel, Prewitt, Robert and Canny [7]detectors, the most important limitation of these methods is that they do not distinguish between texture edges on one hand and region boundaries and object contour on the other hand.

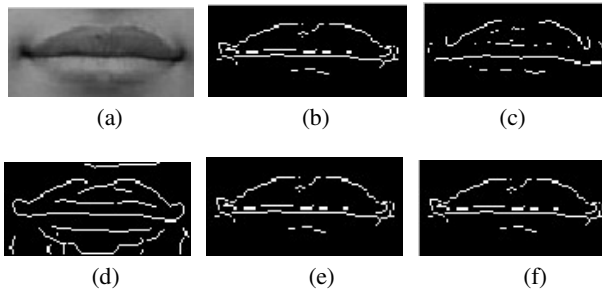


Fig. 2. (a)Gray image (b)Sobel (c)Log (d)Canny (e)Prewitt (f)Robert

b) Morphological edge detectors

Mathematical morphology theory is introduced by Matheron for analyzing geometric structure of metallic and geologic samples. It was first applied to image analysis by Serra [8].

There are two basic morphological operators: erosion and dilation. Erosion is a “shrinking” operator in the values of $F \ominus B$ are always less than or equal to the values of

F. Dilation is an “expansion” operator in the values of $F \oplus B$ are always larger than or equal to the values of F.

By the erosion and dilation operators defined above, we can detect the edge of image F, the erosion residue edge detector is:

$$Ee(F)=F-(F \ominus B) \tag{1}$$

The dilation residue edge detector is:

$$Ed(F)=(F \oplus B)-F \tag{2}$$



Fig. 3. (a) Gray image (b) dilation (c) erosion (e) morphological edge detection

The tradition morphological algorithm only considers of intensity information of the edge and ignores the direction information of the edge.

C)LDA approaches

Fisher linear discriminate analysis (LDA) is a traditional statistical technique. Here we use the Fisher linear discriminate analysis to enhance the gradient information of the lip contour and then we can obtain the lip contour clearly [9].

Here, the G and B components of RGB color space are set to a vector x which is used to distinguish the lip color and skin color. In the training process, we manually extract patches of 50 people's lips and skin regions as the training samples. By utilizing Fisher transformation to the vector $x = (G, B)^T$, we obtain a function that can be used to discriminate between the two classes. This function is calculated by using the within-class scatter matrix and defined as:

$$Fisher(x) = W \cdot x^T \tag{3}$$

The projection vector is calculated by:

$$W = S_w^{-1}(m_1 - m_2) \tag{4}$$

The within-class scatter matrix Sw, is:

$$S_w = S_1 + S_2 \tag{5}$$

$$S_1 = \sum (x - m_1)(x - m_1)^T \tag{6}$$

$$S_2 = \sum (x - m_2)(x - m_2)^T \tag{7}$$

The sample mean vector of each class, m_k is defined as:

$$m_k = \frac{1}{n_k} \sum_{x \in x_k} x_k \quad (k = 1, 2) \tag{8}$$

Where x_k is the set of the vectors in k^{th} class and n_k is the number of the vectors in the class.

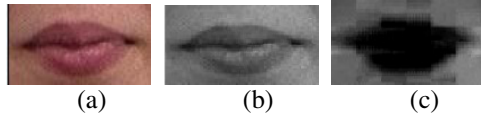


Fig. 4. (a) Original image (b) gray image (c) the result of FDA algorithm

In the context of pattern classification, LDA seeks for the best projection subspace such that the ratio of the between-class scatter to the within-class scatter is maximized. But the labeled data are expensive to obtain, as they require the efforts of human annotators. Contrarily, in many cases, it is far easier to obtain large numbers of unlabeled data. The problem of effectively combining unlabeled data with labeled data is therefore of central importance in machine learning. In this paper, we apply the Semi-supervised LDA algorithm to improve the traditional LDA [10] method. The following is the algorithm step of this method:

- (1) Find a orthonormal basis of the subspace $R (X_k)$:
- (2) Map data points into subspace $R (X_l)$:

$$x_i \rightarrow z_i = P^T x_i, i = 1, \dots , l, l + 1, \dots , n \tag{9}$$

- (3) Construct the scatter matrixes: Construct the between-class scatter matrix S_b and total scatter matrix S_t as

$$S_b = Z_l B Z_l^T = P X_l B X_l^T P^T \tag{10}$$

$$S_t = Z_l C Z_l^T = P X_l C X_l^T P^T \tag{11}$$

- (4) Eigen-problem: Compute the eigenvectors with respect to the nonzero eigenvalues for generalized eigenvector problem:

$$S_b b_i = \lambda_i S_t b_i, i = 1, 2, \dots c - 1 \tag{12}$$

We obtain the transformation matrix $B = [b_1, \dots , b_{c-1}]$.

- (5) Construct the adjacency graph: Construct the p-nearest neighbor graph matrix W as in (13) to model the relationship between nearby data points and calculate the graph Laplacian matrix $L = D - W$.

(6) Eigen-problem: Compute the eigenvectors with respect to the nonzero eigenvalues for the generalized eigenvector problem:

$$S_b a_i = \lambda(aS_i + (1 - \alpha)PXLX^T P^T) a_i, i = 1, \dots, c - 1 \tag{13}$$

(7) Embedding: The data point can be embedded into $c-1$ dimensional subspace by:

$$x \rightarrow y = A^T z = A^T P^T x = (PA)^T x \tag{14}$$

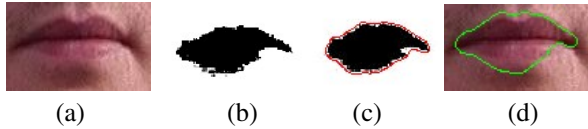


Fig. 5. (a) Original image (b) the result of Semi-supervised LDA algorithm (c) the result of applying median filtering method to remove noise (d) contour detection result

2). Global Algorithm

Global method including the level set method, snake model methods, this section we mainly introduced the level set method. Level set method is an effective tool for describing the curve with curvature related speed evolution. It has been widely used in image segment and computer visual. It is based on the change of the image gradient, the experiment results is in Figure 6.

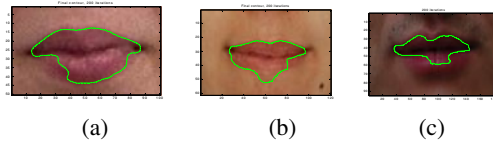


Fig. 6. The results of the traditional level set method

The level set method described in this Section is based on the change of the image gradient. As Figure 7 shown the contour detection may not be accurate if the object contour is not obvious or the gradient information is too weak.

2.3 Hybrid Approaches

Through the analysis and experiments above we can see, these three methods have their advantages shortcomings respectively. The semi-supervised LDA can Learning from labeled and unlabeled data and enhances the gradient information of the lip contour. It can improve the level set method and solve the limitation of the traditional level set method. So in this paper, we present a new hybrid approach combined with semi-supervised LDA method and level set method.



Fig. 7. (a) Original image (b) Edge preserving smoothing result (c) semi-supervised FDA (d) contour detection result

3 Experiment

For testing the performances of our approach, we built a color lip image database. The database contains the data of 100 images. The data set included the lighting, facial expressions and attitude changes.



Fig. 8. Some of the accuracy results of the lip contour detection

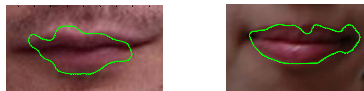


Fig. 9. Depicts such poor detection results

Figure 8 shows some of the better results. From failure cases of our method we observed that there are leakages between upper lip, nose and around the edge of the lip. Figure 9 depicts such poor detection results.



Fig. 10. Detection results of the traditional ASM method

We compared the improved approach with the traditional Active Shape Models. Figure 10 shows the detection results by using traditional ASM. The traditional ASM needs a large set of training set to include all types of lips, the accuracy and speed depend on the initialization of the model parameters. It will increase the contour matching scope and impact detection speed if the initialization is not accurate.

4 Conclusion

The lip contour extraction technology in face recognition is a hot research topic in computer vision. Its research and development not only has great theoretical significance, but also has wide application prospect and practical significance. In this paper, we contrast contour preprocessing method of average filtering, bilateral filtering and edge preserving smoothing technique to eliminate lip image noise and compare the threshold based method, edge and line oriented approaches for lip contour detection. Through the experiment results and analysis of the advantage and disadvantage of these approaches we propose a hybrid approach combines with level set method and semi-supervised Fisher transformation method for lip contour detection. This method applies the semi-supervised idea and combines with the advantages of the two methods which solve the problem of time costing and human annotators efforts. The experiments results proved that it improves the accuracy of detecting the lip contours.

Acknowledgment. This work is supported by the National Natural Science Foundation of China under Grant No. 61073121 and the Natural Science Foundation of Hebei Province under Grant No. F2009000215.

References

1. Fu, K.S., Mui, J.K.: A survey in image segmentation. *Pattern Recognitions* 13, 3–6 (2006)
2. Murthy, C.A., Pal, S.K.: Histogram thresholding by minimizing gray-level fuzziness. *Information Sciences* 60, 107–135 (2002)
3. Jackway, P.T.: Morphological scale-spaces. *Adv. Imag. Elec. Phys.* 119, 123–189 (2001)
4. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: a survey. *Med. Image Anal.* 1(2), 91–108 (1996)
5. Cufi, X., Munoz, X., Freixenet, J., Marti, J.: A review of image segmentation techniques integrating region and boundary information. *Adv. Imaging Electron. Phys.* 120, 1–39 (2002)
6. Ziou, D., Tabbone, S.: Edge detection techniques-an overview. *Int. J. Pattern Recognit., Image Anal.* 88, 297–345 (1994)
7. Ganesan, L., Bhattacharyya, P.: Edge detection in untextured and textured images-a common computational framework. *IEEE SMC* 27(5), 823–834 (1997)
8. Chu, C., Aggarwal, J.K.: The integration of image segmentation maps using region and edge information. *IEEE Trans. Pattern Anal. Machine Intell.* 15, 1241–1252 (1993)
9. Wang, R., Gao, W., Ma, J.: An Approach to Robust and Fast Locating of Lip Motion. *Chinese Journal of Computers*, 866–871 (2001)
10. Yang, W.-Y., Liang, W., Xin, L., Zhang, S.-W.: Subspace Semi-supervised Fisher Discriminant Analysis. *ACTA Automatica Sinica* 35(12), 1513–1519 (2009)

Public-Keys-Based Entity Authentication Protocol with Digital Signature for Cognitive Radio Networks

Chao Wang and JiWen Guo

School of Computer & Communication Engineering, University of Science and Technology
Beijing, Beijing, 100083, P.R. China

Abstract. Considering the characters of Cognitive Radio Networks (CRNs), a new authentication protocol is proposed in this paper. Its security is based on the algorithmically unsolvable result of the large integer factorization problem. The public key encryption and signature are adopted to guarantee security of the transmission message. The entity authentication protocol based on public-key is analyzed in detail in this paper and shows its ability to prevent the network from Non-repudiation, Replay Attack, Interleaving Attack and Man-in-the-middle Attack.

Keywords: Cognitive radio networks, entity authentication, public-keys.

1 Introduction

Cognitive radio (CR) technology is to solve deficiency of spectrum source and its key idea [1] is to improve utilization rate of spectrum by mean of dynamic spectrum management. With development of wireless networks, people pay more and more close attention to cognitive radio networks (CRNs). However, the secure problem is studied less than spectrum sense and spectrum allocation. In [2], some basic secure concepts are introduced for security of CRNs and many kinds of attacks are showed, which may come from external environment or internal environment.

Burbank presented the current security problems of IEEE 802.22 standard and give some methods to solve these problems [3]. What's more, some literatures introduced the security on CRNs [4, 5].

It is necessary to consider of primary users (PUs) and secondary users (SUs) together in CRNs. It is important for all users to authenticate identity of users, who is going to occupy spectrum bands, especially the SUs. Thus authentication protocols play an important role in the security of CRNs.

In CRNs, it is important for studying security to guard against all kind of attacks, such as fabricating information, eavesdropping, hijacking and selfish behavior. Then it is necessary to research the security of CRNs to guarantee integrality, reliability, confidentiality and non-repudiation of data [2]. Because the SUs employ spectrum bands by mean of opportunity access it is useful to research identification [6], which is performed by BSs.

We propose a novel authentication protocol in this paper in order to shorten the authentication delay and guarantee its security. Because this scheme is designed by mean of public-key encryption and digital signature and its transmission data is secure. By analysis, it can prevent the network from man-in-the-middle attack.

2 Network Model

The network model for CRNs, mainly consists of four parts: Access user, Access Point (AP), Base Station (BS), Database. The communication among users can be interlinked by wireless link or wire link. Each BS has two different roles: Users Registration and Verification, which directly interlinks database to store public key, unique identifier and information for SUs. A part of BSs has a duty to link the primary networks such as Internet networks or the others CRNs. The architecture of network model is showed in Figure 1.

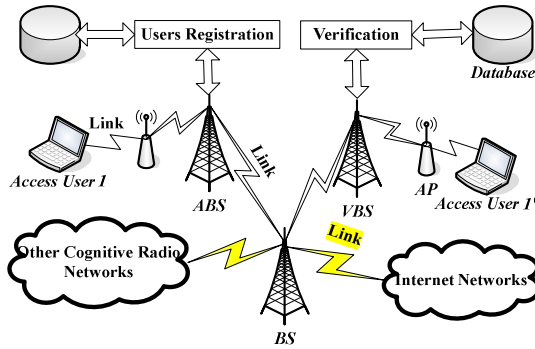


Fig. 1. Network model for CRNs

In the network model, before SUs execute an authentication protocol they need register relevant information in some BS firstly. Because SUs move to different BS ,they may need switch spectrum continually. In other words, one BS registering a SU_1 is known as Authorized Base Station (ABS) for it. The other BS, which allows SU_1 to use idle spectrum bands, is known as Verifiable Base Station (VBS) for it. However, the ABS may be coincide with the VBS for SU_1 .

In Figure 2, when SU_1 moves from VBS_1 to VBS_2 along direction of arrow, it may arrive to VBS_2 and also need use spectrum bands to transmit data. Then before SU_1 utilizes spectrum bands ,it must be confirmed the identity, namely which is executed re-authentication. Thus authentication delay becomes important indicator which effects efficiency of authentication protocol.

Definition 1. Authentication delay represents a period of the whole authentication. The initial time point is that some SU sponsors request. The finished time point is that it finishes transmitting data or is refused to enter CRNs. It consists of two parts: the first part is a time, which executes an authentication flow; the second part is a time of spending on re-authentication.

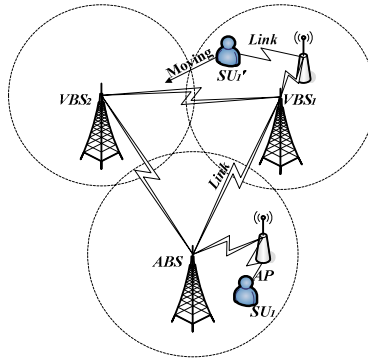


Fig. 2. Spectrum switching model for CRNs

3 Public-Key-Based Entity Authentication Protocol

3.1 Basis Idea

The RSA algorithm [7] generates asymmetrical key to fulfill the process of authentication. The security of RSA algorithm is based on the algorithmically unsolvable result of large integer factorization problem. The asymmetrical keys of our authentication protocol are generated by RSA algorithm. Thus the security of PEAP also depends on the algorithmically unsolvable result of large integer factorization problem. In this sub-section, we will introduce some basis notation and terminology as follows.

Let the number of SU be m in CRNs and SU_i be some secondary user, where $i = 1, 2, \dots, m$. Identifier ID_i denotes the unique identity of SU_i . *Request* and *Ack* denote request packet and respond packet, respectively. $Sign_A(.)$ denotes A 's signature mechanism. $Cert_{PK}$ denotes the public-key certificate containing A 's signature public key. $P_{PK}(.)$ denotes encrypting computation by public key PK . $Hash(.||)$ denotes a hash function.

Suppose that the number of BS be n in CRNs. The public keys are allocated to all BSs before the authentication procedure is executed. Each BS generates a pair of public-key by mean of RSA algorithm. The (e_j, l_j) denotes public key of BS_j and the d_j denotes its private key, where $j = 1, 2, \dots, n$. The BSs store the private key of own and broadcast the public key of own. Suppose all BSs don't suffer from any attack in which public keys are broadcasted. Thus each BS possesses its private and acquires public keys of the others BSs. Now that $l_i = l_k$ ($1 \leq i \neq k \leq n$) the public key (e_j, l_j) is marked by e_j .

3.2 Authentication Procedure

In this subsection, we present the procedure of PEAP and introduce its specific steps as follows.

Step1. SU_i generates a random number r_i by pseudo random generator. SU_i sends the request packet: $I = (Request, ID_i, r_i)$ to BS_j , which serves as ABS of SU_i . The ABS generates a random number r_{iD} by pseudo random generator. A pair of public-key is generated by the RSA algorithm, where public key e_i be stored in the ABS and private key d_i be allocated to SU_i by the secure channel. The ABS sends the respond packet: $II = P_{e_i}(ID_i, BS_j, r_i, r_{iD})$ to SU_i . Then the data packet II is decrypted to acquire the random number r_{iD} by private key d_i and verify whether the generated random number is identical to the received random number or not. If the result is right, SU_i sends the data packet: $III = (Cert_{si}, r_i, BS_j, Sign_{si}(r_i, r_{iD}, BS_j))$ to the ABS. The ABS verifies whether the digital signature is right or not. If the digital signature is valid, this user's information has been registered correctly.

Step2. SU_i sends the request packet: $IV = (Request, ID_i, BS_j, r_i)$ to BS_k , which serves as its VBS. Let e_k be the public key of the VBS and d_k be the private key of it. If the VBS exists utilized spectrum, it generates a temporary random number r_{ik} by pseudo random generator. The VBS sends the data packet: $V = P_{e_j}(ID_i, BS_k, BS_j, r_i, r_{ik})$ to its ABS.

Step3. After the ABS decrypts the data packet V and acquires random number, user information. Then the ABS generates a temporary random number r'_{ij} by pseudo random generator and returns the data packet: $VI = (Cert_A, r'_{ij}, Sign_A(r_{ik}, r'_{ij}, BS_k))$.

Step4. Verify whether it is right or not after the data packet VI is received. If this digital signature is valid the VBS returns the data packet: $VII = (Cert_V, r_{ik}, Sign_V(r'_{ij}, r_{ik}, BS_j))$.

Step5. Verify whether it is right or not after the data packet VII is received. If this digital signature is valid the ABS generates a temporary random number r_{ij} by pseudo random generator and sends data packet: $VIII = (Cert_A, r_i, e_i, P_{e_k}(r_{ij}), Sign_A(r_i, r_{ik}, r_{ij}, e_i, ID_i, BS_k))$ to the VBS.

Step6. Decrypt the data packet $VIII$ and verify whether it is right or not. If this data packet is valid the VBS returns allowable respond packet: $IX = P_{e_i}(ID_i, BS_k, r_i, r_{ij})$

to SU_i . After SU_i decrypts the respond packet IX the data packet: $X = (Cert_{S_i}, r_i, Sign_{S_i}(r_i, r_{ij}, ID_i, BS_k))$ is sent to the VBS.

Step7. When the VBS receives the data packet X verify whether it is right or not. If the data packet X is right SU_i will be permitted to enter the CRNs. Otherwise, the authentication procedure will be stop.

Step8. When spectrum switching takes place in the same BS re-authentication will not be executed for SU_i . If the VBS exists utilized spectrum the VBS sends allowable request packet: $XI = (ID_i, BS_k, Hash(r_i || r_{ij}))$ to SU_i . After SU_i receives request packet XI verify whether it come from BS_k . If this request packet is right SU_i will send the data packet: $XII = (Cert_{S_i}, r_i, Sign_{S_i}(r_i, r_{ij}, ID_i, BS_k))$ to the VBS.

Step9. When the VBS receives the data packet ,XII verify whether it is right or not. If the data packet is right it will be permitted to enter the CRNs again. If it travels to others BSs the authentication procedure will return to the upper Step2. Meanwhile, the present VBS, previous VBS play the role of a new VBS, a new ABS, respectively. Otherwise, it doesn't execute re-authentication.

In the authentication procure, when spectrum switching for any SU only takes place in the same BS, it will not be executed re-authentication. However, if each SU enter different AP belonging to the same BS it must be executed re-authentication. Because the probability of re-authentication is reduced hence authentication delay is shortened and certified efficiency is improved. The authentication procedure is showed as follows in Figure 3.

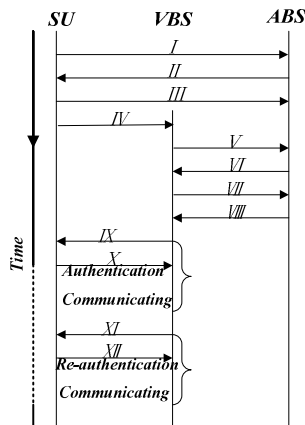


Fig. 3. Flow of authentication procedure

4 Conclusions

In this paper, a novel authentication protocol based on public-key was proposed for CRNs. The security of the proposed protocol is guaranteed by RSA algorithm. Our protocol excessively relies on the BSs which are reliable and secure. Thus these BSs become the targets which malicious nodes intrude. What's more, the authentication protocol is suitable for identity authentication of PUs. As the SUs and PUs make use of spectrum bands by mean of coordination they are in the nature of the analogous access method. Thus PEAP can effectively prevent from primary user emulation attacks.

Acknowledgements. This work is supported in part by the National Science Foundation of P. R. China under Grant (no. 60902042).

References

1. Akyildiz, I.F., Leea, W.Y., Chowdhury, K.R.: CRAHNS: Cognitive radio ad hoc networks. *Ad Hoc Networks* 7, 810–836 (2009)
2. Mahmoud, Q.: *Cognitive networks: towards self-aware networks*. Wiley & Sons Ltd., Chichester (2007)
3. Burbank, J.L.: Security in cognitive radio networks: the required evolution in approaches to wireless network security. In: *Proceedings the 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, pp. 1–7 (2008), doi:10.1109/CROWNCOM.2008.4562536
4. Bernthal, B., Jesuale, N.: Smart radios and collaborative public safety communications. In: *Proceedings the 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 1–20 (2008), doi:10.1109/DYSPAN.2008.70
5. Chen, R., Park, J.M., Reed, J.H.: Defense against primary user emulation attacks in cognitive radio networks. *IEEE Journal on Selected Areas in Communications* 26, 25–37 (2008)
6. Miller, R., Xu, W.Y., Kamat, P., Trappe, W.: Service discovery and device identification in cognitive radio networks. In: *Proceedings the 2nd IEEE Workshop on Networking Technologies for Software Define Radio Networks*, pp. 670–677 (2007), doi:10.1109/SDRN.2007.4348972
7. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM* 21, 120–126 (1978)

An IPA Analysis of Effective Teaching Activities in China

Haiyan Kong

Business School, Shandong University at Weihai
No.180 West Culture Road, Weihai, Shandong, China
Konghaiyan@sdu.edu.cn

Abstract. This study aims to explore students' perceptions of teaching activities both inside and outside the classroom. Data were collected in mainland China and a total of 212 valid data were obtained. To explore the detail information of students' perceptions of teaching techniques, importance performance analysis (IPA) method was used to analyze the data. It has been found that internship, practiced-oriented education, and field trip were regarded as the most important experiential learning. Students were satisfied with the learning environment in library, teachers' motivation and group works. A great gap was found between the expectation and performance of internship. The findings of this study may provide useful guidance to both academic and practitioners in terms of enhancing students' learning and skills.

Keywords: Teaching activities, learning satisfaction, IPA analysis.

1 Introduction

Nowadays, the competition in the labor market has been increasingly keen. Hunting for a good job has become a great concern for most undergraduate students. On the one hand, millions of college students graduated with an increase rate of 40% over the previous year [1]. On the other, most employers complained about the shortage of qualified graduates from schools in China [2]. This highly reflects the gap between school teaching program and industry demands. It is therefore urgent for universities or colleges to design special teaching programs, practice effective teaching activities, and to provide the society with qualified and high-quality labors.

Effective teaching methods play important roles in enhancing students' learning outcomes and satisfaction. Qualified teaching may help to widen students' horizons, enhance their career expectation, and strengthen their learning and application of classroom knowledge [3]. However, most teachers focused on traditional teaching or research instead of applying theories in practices. Without offering tailor-made programs and sufficient training, there was a gap between school curriculum and industry needs [4]. As a new generation, most young people want to explore more of the world, and they would like to accept new learning methods by experiencing learning. Effective teaching methods, such as experiential learning, may motivate the students to involve in the learning process, and then offer a significant link between the discipline and industry.

Experiential learning refers to multiple programs and systems which provide students in educational institutions with work-based applied learning opportunities. Although active teaching activities were critical for both students and the industry, most class methods could not meet the demands of students. Many students are passive learners and rarely participated in the classroom learning. Therefore, it is important to explore students' perceptions towards different kinds of teaching methods, and to examine which class activities are popular among the students.

2 Literature Review

Effective teaching methods contribute to knowledge transfer and students' learning satisfaction. Previous studies have explored the concepts and methods engaged in teaching activities [5]. It indicated that experiential education engages the learning activities in which the student is engaged in the field being studied. As a result, effective teaching methods, such as experiential education, provide a significant opportunity for students to link their knowledge learned in the class with the development of their employability skills. This is because experiential learning may increase students' learning motivation and satisfaction [6], and further enhance their understanding and interaction with the real world.

Many researchers have explored the concept and importance of experiential learning activities. It has been proved that effective teaching methods may develop students' learning outcomes. In experiential learning, students changed their roles from passive learners to active participants [7]. There are many experiential learning techniques practiced to enhance students' learning efficiency, such as internship, case study, field trip, role play, group discussion, service learning, and simulation games [8, 9, 10]. Gentry (1990) [11] further explained a range of techniques which may be possibly used in experiential learning, which include assessment centers, forums, group discussion, panel meetings, live cases, writing experiences, student-written textbooks, computer-assisted instruction, communication workshops, Delphi forecasting, on-the-job training, time management sessions, learning cooperatives, internships, job search preparation, field trips, game show formats, etc.

Although much has been written about customer satisfaction, there has been relatively little empirical work exploring students' perceptions of learning activities. This study, therefore, is to assess how teaching methods meet students' expectation by applying importance-Performance analysis (IPA), and to further provide valuable implications for academic and researchers.

3 Research Methodology

A questionnaire approach was used to collect students' information. The questionnaire was designed to collect data about perceptions towards different kinds of teaching activities. It was designed for sector-specific interview within an IPA conceptual framework. Participants were asked to rate each of the teaching methods attributes twice in a paired way: (a) to rate, according to the importance they attached to the attribute, the importance level of each teaching method on a Likert-type scale

of “1” (not at all important) to “5” (very important), and then (b) to rate, according to their satisfaction level of each of the same attributes on a scale of “1” (very dissatisfied) to “5” (very satisfied). At the end of the questionnaire, it was designed to collect respondents’ demographic information including gender, age, educational level, major, and grade.

The questionnaires were distributed to students majoring in human resources management, tourism management, marketing, finance management, accounting, and logistic management. The students were asked to complete self-administered questionnaires which were collected on the spots. A total of 212 valid questionnaires were obtained, providing necessary data for the current research. SPSS17.0 (the Statistical Package for the Social Sciences) was employed to analyze the data. Specifically, to reach the research objective that has been stated above, the IPA analyses were performed for the experiential learning activities respectively.

4 Findings and Discussion

Of the 212 students, 58% were male and 58% were female, age ranging from twenty to twenty five. The students surveyed were freshman (9%), sophomore (18.9%), junior students (65.1%), senior students (4.7%), and post graduates (10.3%). The respondents’ perceptions of teaching activities were analyzed. First, the importance of each experiential technique was evaluated, then the satisfaction level, and finally the gap between importance and perceptions was examined.

4.1 Importance Analysis of the Teaching Activities

The findings indicated that the most important experiential learning was internship with mean of 4.35, followed by practice-oriented education (4.26), and field trip (4.25). In addition, other teaching activities, such as library (4.16), mutual interaction (4.12), case study (4.10), lecturer (4.09), teacher’s motivation (4.06), group works (4.00), and project-based learning (4.00) were also had relatively higher scores according to the importance. On the contrary, certificate-oriented learning (3.37) and skills competition (3.51) were regarded as less important learning activities.

According to the mean of the importance, it is obvious that industry-based experiential learning activities were highly regarded as important by the students. This is because these kinds of experiential learning provide students the opportunities for deeper levels of learning and application of classroom learning. The results of this study were consistent with previous findings which indicated that learning is most effective when it is grounded in experience. For example, internship may provide the students opportunities to apply what they learn in the classroom in an actual, real-world work experience. By participating in the experiential learning programs, students gain the necessary experience and may help them to have a successful start of their career.

4.2 Performance Analysis of the Teaching Activities

The satisfaction level of teaching activities was also analyzed. It indicated that students were most satisfied with the learning environment in library (3.76), followed

by teacher's motivation (3.54), group works (3.53), and mutual interaction (3.53) between the teacher and students in the class. The university libraries may provide students satisfactory learning environment, facilities and information. While lecturing, teachers tended to communicate with students on the class. In addition, group works was listed as satisfactory learning activity, and students may learn from each other by working together or by brainstorm discussion. Class presentation offered students the stage and opportunity to present their ideas and elaborate their findings. Thus, it was also popular among the students. The data also showed that the satisfaction level of virtual learning (3.47), case study (3.46), lecture (3.39), challenging cup competition (3.38), and practice-orientated education (3.37) were relatively high.

However, they were not satisfied with the learning activity of overseas attachment program (2.91), as most students had few opportunities to study abroad due to limitation of quota or fund. Overseas attachment program offered students a special learning process by experiencing abroad. By attending this program, students may widen their horizons by understanding the teaching and subjects offered outside China. Some students took part in the attachment program by studying abroad for a period of time (one year or one semester), or by attending summer schools abroad. It was widely accepted as a good way of learning, but there required some formalities and fund. Thus some students may not afford to do so. Apart from that, most students hope to experience the real world by attending field trip, but this kind of experiential learning was practiced relatively less.

4.3 Gap between Importance and Satisfaction

In terms of the gap between students' perceptions of importance and satisfaction, the biggest difference is internship. All students strongly agreed that internships may develop job skills, and this item was the highest rated item in terms of importance. However, the satisfaction level of this item was low. That is, internship was highly regarded as the most important learning method, but the application of it was not satisfied. It has been stated that the internship is an opportunity to have an intensive, work-based exposure to a broad range of operations within a company. Internship offered student training under both academic and practitioner supervision. But in reality, the practice of internship can not meet the demands of students, due to short term and less supervision. Thus, this study may provide insights for teachers to arrange necessary internship for their students in the future.

Field trip was also ranked as important learning activities with low satisfaction level. Most students hoped to enhance their knowledge by field-based learning, but there is a gap between their expectation and performance. It has been found that field trips had a signification positive impact on students' learning satisfaction . Moreover, field trip not only enhances student learning, but also provides precious practical experience for educators. Therefore, teachers should pay more attention to implement relevant field trips so as to enhance students' knowledge and experience.

Apart from internship and field trip, overseas attachment program, practice-oriented education, project-based learning, and role play should also been further developed. These activities are typically relevant to the enhancement of work experience, and can prepare business ready graduates with the employability skills. It is obvious that there is much room to be enhanced in terms of experiential learning in China.

Table 1. Performance analysis of teaching activities

Teaching Activities	I	S	Gap
1.Internship	4.35	3.17	1.18
2.Field trip	4.25	3.13	1.12
3.Overseas attachment program	3.86	2.91	0.95
4.Practice-oriented education	4.26	3.37	0.89
5.Project-based learning	4.00	3.25	0.75
6.Role play	3.96	3.25	0.71
7.Lecture	4.09	3.39	0.70
8.Guest speaker	3.92	3.25	0.67
9.Case study	4.10	3.46	0.64
10.Bilingual teaching	3.88	3.28	0.60
11.Mutural interaction	4.12	3.53	0.59
12.Teachers' motivation	4.06	3.54	0.52
13.Class presentation	3.98	3.50	0.48
14.Group works	4.00	3.53	0.47
15.Library	4.16	3.76	0.40
16.Skills competition	3.51	3.14	0.37
17.Project	3.66	3.37	0.29
18.Challenging cup competition	3.66	3.38	0.28
19.Virtual learning	3.74	3.47	0.27
20.Certificate-oriented learning	3.37	3.15	0.22

* I= Importance, S= Satisfaction.

5 Conclusions

This study indicated that students highly regarded internship, practice-orientated education as important teaching methods. These activities can provide a process in which students can learn from their work experiences and improve their employability. Experiential education benefited students with higher abilities and has the potential to help the students to obtain the soft skills needed in today's market. It is therefore important for teachers to arrange relevant internship to offer students the opportunity to acquire transferable skills and the specific detailed knowledge necessary in today's workplace.

The results showed that most students were dissatisfied with face-to-face lecture or boring classroom teaching, they were looking for an enhanced learning experience from vivid teaching activities. It is therefore important for teachers to reform the traditional teaching, and adopt experiential learning, which may focus more on "knowing how" instead of "knowing what", problem-solving instead of problem-making, and task-based instead of knowledge-based. It is necessary for educators and employers work together to provide students with tailor-made course curriculum, and help them to cope with the challenges they may encounter in the real work world.

This study suggested that it was not sufficient for the teachers to only transmit information to the students on the class, education should be grounded in experience. Many skills can be learnt from practical experience and students may have more

chances to attain this experience in the field before their graduation. In order to improve students' potential for success in their employment, teachers should adopt various experiential learning activities, such as internships, field trips, and practice-oriented programs, to offer work experience and work-related learning. Therefore, a good teaching method should combine the common teaching methods with experiential learning activities.

Acknowledgment. This paper is supported by Shandong Social Science Planning Research Grant # 10CJGJ60, and Teaching Reform Grant of Shandong University at Weihai.

References

1. Zhang, Y.: Social occupational classes and higher-education opportunities in contemporary China: A study on the distributional of a scarce social capital. *Front, Educc., China* 1, 89–99 (2006)
2. Guo, Z., Van der Heijden, B.I.J.M.: Employability enhancement of business graduates in China. Reacting upon challenges of globalization and labour market demands. *Education and Training* 50(4), 289–304 (2008)
3. Cooper, S., Bottomley, C., Gordon, J.: Stepping out of the classroom and up the ladder of learning: An experiential learning approach to entrepreneurship education. *Industry & Higher Education* 18(1), 11–22 (2004)
4. Pownall, D., Jones, M., Meadows, M.: Bridging the gap between tourism teaching and industry: A review of the first year of the tourism and leisure post graduate teaching certificate at Liverpool John Mouress University, HK. *Journal of Teaching in Travel & Tourism* 7(4), 85–101 (2007)
5. Huerta-Wong, J.E., Schoech, R.: Experiential learning and learning environments: The case of active listening skills. *Journal of Social Work Education* 46(1), 85–101 (2010)
6. Tse, T.S.M.: What do hospitality students find important about internships? *Journal of Teaching in Travel & Tourism* 1(3), 251–264 (2010)
7. Hawtrej, K.: Using experiential learning techniques. *Journal of Economic Education* 38(2), 143–152 (2007)
8. Tse, T.S.M.: What do hospitality students find important about internships? *Journal of Teaching in Travel & Tourism* 1(3), 251–264 (2010)
9. Wong, A., Wong, S.: Useful practices for organizing a field trip that enhances learning. *Journal of Teaching in Travel & Tourism* 8(2-3), 241–260 (2008)
10. Beidatsch, C., Broomhall, S.: Is this the past? The place of role-play exercises in undergraduate history teaching. *Journal of University Teaching & Learning Practice* 7(1), 1–20 (2010)
11. Gentry, J.W.: What is experienced learning? In: Gentry, J. (ed.) *Guide to Business Gaming and Experiential Learning*, pp. 9–20. Nichols Publishing, East Brunswick (1990)

A Robust Control Scheme for Nonlinear Systems

Limin Liu

Institute of Embedded Systems
IT School, Huzhou University
Huzhou, Zhejiang, 313000, China
liulimin@ieee.org

Abstract. The control for nonlinear systems, especially for nonlinear MIMO industrial processes, is an important solution of intelligent control. Robustness of a nonlinear control system is an expected target for most designers. In this paper, a method of robust intelligent control for nonlinear systems is discussed. The scheme is built through some robust intelligent model. On the other hand, the control is concerned to an immune model with self-updating.

Keywords: Robust intelligent control, non-linear systems, immune model, fuzzy systems, MIMO.

1 Introduction

A control system's design is usually based on some mathematic model. In reality, the system may behave differently than the model indicates, or the system parameters may vary with time, especially for some nonlinear systems [1]. The Nyquist stability test was developed in 1932 as a graphical way to determine the stability of a linear single-input-single-output (SISO) feedback control system. This test shows how robust a control system is to a system's uncertainties. That is, given a stable control system, how much can the plant gain change before the system becomes unstable. But, today, most real control systems are multiple-input multiple-output (MIMO) ones. Hence, they are not amenable to robustness analysis via the Nyquist criterion (even though it continues to be useful) [2]. Also, the Nyquist stability criterion can be used only with specific types of perturbations (gain and phase perturbations). Thus, perturbation analysis based on intelligent control was developed. Intelligent control can be applied in the presence of any type of perturbation. As a result, intelligent technique's popularity is increasing as a general way of analyzing and designing the robustness of MIMO systems.

Fuzzy control, neural network and knowledge-based system are three main branches of intelligent control system [3]. The idea of fuzzy system was first applied to control system in 1974. Since that time, there have been a lot of successful applications based on fuzzy systems [4]. In recent years, more application systems are developed in different fields [5], such as industrial processes, electromechanical systems, instruments, automobiles, communications and so on. Actually, fuzzy control, neural network and knowledge-based techniques can be hybrid in one application [6]. For most nonlinear MIMO industrial processes, intelligent control has become a common approach [7].

Stability and reliability are most important for most industrial systems. In order to guarantee the stability and reliability of an industrial process, the robustness of its control system is required [8]. The requirement, in fact, is to ask some robust intelligent control systems.

2 A Scheme of Robust Intelligent Control

There are various solutions of nonlinear MIMO systems. Decoupling control is common traditional method for MIMO system. The key of the solution is to find the decoupling matrix. Sometimes, the matrix is a high order reverse matrix, and may be solved hardly. Singular value technique is another method to performance analysis and design robust control system for MIMO processes. The structured singular value can be used to determine the stability of a system that is subject to structured perturbations. The key is to find a suitable perturbation. For some MIMO system, the perturbation cannot be got easily.

An intelligent control scheme of MIMO systems, however, is easier to be fetched than other MIMO control ways.

2.1 A Universal Model of Nonlinear Intelligent Control

A smooth nonlinear control system is as the following form.

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) \tag{1}$$

where

$$x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T \in R^{n \times 1}$$

denotes the state vector.

$$u(t) = [u_1(t), u_2(t), \dots, u_m(t)]^T \in R^{m \times 1}$$

denotes the input variables.

It is well known that a Takagi-Sugeno model can be a universal approximation of a smooth nonlinear dynamic [9]. The nonlinear system is represented as follows.

Rule i:

If $z_1(t)$ is M_{i1} and ... and $z_n(t)$ is M_{in}

Then $\dot{x}(t) = (A_i + \Delta A_i(t))x(t) + (B_i + \Delta B_i(t))u(t)$

$$y(t) = C_i x(t) \quad i=1, 2, \dots, g \tag{2}$$

where M_{ij} is the fuzzy set, $A_i \in R^{n \times n}$, $B_i \in R^{n \times m}$, and g is the number of rules. $z_1(t), \dots, z_n(t)$ are premise variables. A_i, B_i, C_i are known constant matrixes to

describe the nominal system. The structure of the time-varying parametric uncertainties represented by $\Delta A_i, \Delta B_i$ is as follows.

$$[\Delta A_i(t), \Delta B_i(t)] = H_i F_i(t) [E_{1i}, E_{2i}] \quad i=1, 2, \dots, g \tag{3}$$

where H_i, E_{1i}, E_{2i} are constant matrix appropriate dimensions. $F_i(t)$ is an unknown matrix which is bound as:

$$F_i(t) \in \Omega_{F_i} := \{F_i(t) : \|F_i(t)\| \leq 1, \text{ the elements of } F_i(t) \text{ are Lebesgue measurable}\}$$

The overall fuzzy system is inferred as follows.

$$\begin{aligned} \dot{x}(t) &= \frac{\sum_{i=1}^g \mu_i(x(t)) \{(A_i + \Delta A_i(t))x(t) + (B_i + \Delta B_i(t))u(t)\}}{\sum_{i=1}^g \mu_i(x(t))} \\ &= \sum_{i=1}^g h_i(x(t)) \{(A_i + \Delta A_i(t))x(t) + (B_i + \Delta B_i(t))u(t)\} \end{aligned} \tag{4}$$

$$y(t) = \sum_{i=1}^g h_i(x(t)) C_i x(t) \tag{5}$$

where

$$x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$$

and

$$\mu_i(x(t)) = \prod_{j=1}^n M_{ij}(x(t)) \quad ; \quad h_i(x(t)) = \frac{\mu_i(x(t))}{\sum_{i=1}^g \mu_i(x(t))}$$

$M_{ij}(x(t))$ is the grade of membership of $x(t)$ in M_{ij} . It is assumed for all t as follows.

$$\mu_i(x(t)) \geq 0, \quad i=1, 2, \dots, n \quad ; \quad \sum_{i=1}^n \mu_i(x(t)) > 0$$

Then the following conditions for all t may be obtained.

$$h_i(x(t)) \geq 0, \quad i=1, 2, \dots, g \quad ; \quad \sum_{i=1}^g h_i(x(t)) > 0$$

The (4) is a typical Takagi-Sugeno fuzzy system [10].

2.2 Robust Intelligent Controller and Immune Model

In order to guarantee the global asymptotic stability of the controlled Lyapunov's direct method, the sufficient conditions are discussed here.

For the system shown in (2), a robust fuzzy controller with the state feedback law is described as follows.

Rule i:

$$\text{If } x_1(t) \text{ is } M_{i1} \text{ and } \dots \text{ and } x_n(t) \text{ is } M_{in}$$

$$\text{Then } u(t) = K_i x(t), \quad i=1, 2, \dots, g \tag{6}$$

The final output of the controller is as below.

$$u(t) = \sum_{i=1}^g h_i(x(t)) K_i x(t) \tag{7}$$

Refer to (1) and (2), the closed-loop system becomes

$$\dot{x}(t) = \sum_{i=1}^g \sum_{j=1}^g \mu_i \mu_j (A_i + \Delta A_i + (B_i + \Delta B_i) K_j) x(t) \tag{8}$$

$$y(t) = \sum_{i=1}^g h_i(x(t)) C_i x(t) \tag{9}$$

The closed-loop eigenvalues can be chosen according to the performance requirement for $A_{1+} B_1 K_1, A_{2+} B_2 K_2, \dots, A_{g+} B_g K_g$. Then the controllers K_1, K_2, \dots, K_g can be obtained. These controllers may guarantee the good robust performance of controlled nonlinear systems. For some large-scale systems, the following theorem can offer an asymptotically stable control.

Theorem: If there exist a symmetric and positive definite matrix P, and some scalars λ_{ij} ($i, j= 1, 2, \dots, g$) such that the following LMIs, Linear Matrix Inequalities, are satisfied, the fuzzy system (4) is asymptotically stable via the fuzzy model-based state feedback controller (7) as (10).

Although the theorem may guarantee the robust control for some nonlinear processes, it cannot supply the immune solution when the control scheme is asked updating. Therefore, an immune algorithm is required. For many applications, the immune system means that the system can update itself. An immune algorithm in the intelligent control may be considered as self-updating model[11].

In order to build the self-updating model for an intelligent control system, the hierarchical rule sets may be designed. The rules for control, such as rules of robust control, are low-level rules. There are rules in higher level. They can modify or update low-level rules. They may be called system rules. For some intelligent systems, the self-updating or immune models are based on system rule sets.

$$\begin{bmatrix} A_i^T P + PA_i + K_i^T B_i^T P + PB_i K_i & X & X \\ E_{1i} + E_{2i} K_i & -\lambda_{ij} I & X \\ H_i^T P & 0 & -\lambda_{ij}^{-1} I \end{bmatrix} < 0, \tag{10}$$

(1 ≤ i ≤ g)

$$\begin{bmatrix} A_i^T P + PA_i + A_j^T P + PA_j + K_j^T B_i^T P + PB_i K_j + K_i^T B_j^T P + PB_j K_i & X & X & X & X \\ E_{1i} + E_{2i} K_j & -\lambda_{ij} I & X & X & X \\ E_{1j} + E_{2j} K_i & 0 & -\lambda_{ij} I & X & X \\ H_i^T P & 0 & 0 & -\lambda_{ij}^{-1} I & X \\ H_j^T P & 0 & 0 & 0 & -\lambda_{ij}^{-1} I \end{bmatrix} < 0 \tag{11}$$

(1 ≤ i ≤ g)

where X denotes the transposed elements in the symmetric positions.

3 The Robust Intelligent Control for a MIMO Industrial Process

A complex industrial process often deals with the problems of multivariable control in which there are more than two controlled variables and acting variables in a system, or to be called as Multiple Input and Multiple Output (MIMO) systems. Most of the systems are nonlinear and closed coupling. In a practical application, the design of a robust intelligent controller is not very complicated as description in section 2. It is because that the variables of an industrial MIMO nonlinear control system are not more than three inputs and three outputs normally. On the other hand, the implementation of automation for the system is often based on intelligent model, but not only mathematic algorithm.

In fact, operators do not make complicated calculation and still manage the system better in some MIMO industrial processes. This shows that the system can be controlled by non-mathematics way. Based on an imitated control, intelligent fuzzy controller for the ball mill can be designed. The core of control algorithm is fuzzy rule sets.

In some case, deviation of variables are quantized into nine fuzzy levels as: PV, Positive Very large; PL, Positive Large; PM, Positive Medial; PS, Positive Small; ZO, ZerO; NS, Negative Small; NM, Negative Medial; NL, Negative Large; NV, Negative Very large..

4 Conclusions

A robust intelligent control can be analyzed and explained through mathematic method. But for some practical applications, the scheme of robust intelligent control

does not depend on the mathematical model. The intelligent control is suitable for the practical applications requested to operate simply, maintain conveniently and modify easily.

The robustness of intelligent control for nonlinear MIMO systems is guaranteed by some control rules from operating experts in the industrial fields. The immune model can update control scheme and guarantee the available system operation when the process shows some uncertain status.

Acknowledgments. This research was supported in part by the National Natural Science Foundation of China under grant 60872057, by Zhejiang Provincial Natural Science Foundation of China under grants R1090244, Y1101237, Y1110944 and Y1100095. We are grateful to NSFC, ZJNSF and Huzhou University.

References

1. Chang, Y.: Intelligent Robust Tracking Control for a Class of Uncertain Strict-Feedback Nonlinear Systems. *IEEE Transactions on Systems, Man, and Cybernetics* 39(2), 142–155 (2009)
2. Liu, L.: A Control Based on Rule Updating for Non-Linear Systems. In: *Proc. of 2011 CCDC, Mianyang, China*, vol. 5, pp. 3094–3097 (2011)
3. Nazir, M.B., Wang, S.: Optimum robust control of nonlinear hydraulic servo system. In: *Proc. of IECON 2008, Orlando*, vol. 11, pp. 309–314 (2008)
4. Haber, R.E., Alique, J.R., Alique, A., Jiménez, J.E.: Embedded Fuzzy Control System: Application to an Electromechanical System. In: Sloot, P.M.A., Abramson, D., Bogdanov, A.V., Gorbachev, Y.E., Dongarra, J., Zomaya, A.Y. (eds.) *ICCS 2003. LNCS*, vol. 2658, pp. 812–821. Springer, Heidelberg (2003)
5. Chiu, C.S.: Robust adaptive control of uncertain MIMO non-linear systems - feedforward Takagi-Sugeno fuzzy approximation based approach. In: *IEEE Proceedings of Control Theory and Applications*, vol. 152(2), pp. 157–164 (2005)
6. Villacci, D., Bontempi, G., Vaccaro, A.: An adaptive local learning-based methodology for voltage regulation in distribution networks with dispersed generation. *IEEE Transactions on Power Systems* 21(8), 1131–1140 (2006)
7. Chu, M., Jia, Q., Sun, H.: Robust Tracking Control of Flexible Joint with Nonlinear Friction and Uncertainties Using Wavelet Neural Networks. In: *Proc. of ICICTA 2009, Changsha, China*, vol. 10, pp. 878–883 (2009)
8. Szabat, K.: Robust control of electrical drives using adaptive control structures — a comparison. In: *Proc. of ICIT 2008, Chendu, China*, vol. 4, pp. 1–6 (2008)
9. Hua, C., Wang, Q., Guan, X.: Robust Adaptive Controller Design for Nonlinear Time-Delay Systems via T–S Fuzzy Approach. *IEEE Trans. on Fuzzy Systems* 17(8), 901–910 (2009)
10. Liang, Y., Xu, S., Ting, L.: T–S Model-Based SMC Reliable Design for a Class of Nonlinear Control Systems. *IEEE Trans. on Industrial Electronics* 56(9), 3286–3295 (2009)
11. Ling, P., Wang, C., Lee, T.: Time-Optimal Control of T-S Fuzzy Models via Lie Algebra. *IEEE Trans. on Fuzzy Systems* 17(8), 737–749 (2009)

A Research for Bilingual Courses of Embedded Systems

Limin Liu

Institute of Embedded Systems
IT School, Huzhou University
Huzhou, Zhejiang, 313000, China
liulimin@ieee.org

Abstract. Embedded systems are the most popular computer application systems. With the larger market, education of embedded systems is becoming a hot point in universities. Since the newest documents for embedded systems, such as papers, reports and data sheets, are mostly issued in English, some graduated students feel difficulty to collect new information for updating their knowledge. It is because their native language and course books are not in English. In order to improve their professional English in study and applications, a new way, bilingual education, is developing in some countries. This paper focuses on the organization, objects, and contents of bilingual education for courses of embedded systems.

Keywords: Embedded systems, bilingual education, computer engineering, IT education.

1 Introduction

The embedded systems are influencing our life in many fields every day. The embedded systems have taken an irreplaceable position in many countries [1] and China. In order to continue pace of technological innovation, many Asian countries are making long-term commitments to increase their engineering, such as electronic and computer engineering, workforce [2]. However, good teaching is at the heart of good engineering and prosperous societies. Therefore, the relevant education is important and useful for high quality workforce in embedded systems [3][4].

Since the contribution of the United States for electrical engineering and IT science, twentieth century is called “ An American Century” [5]. The newest documents of embedded systems, such as research papers, reports, specifications, data sheets and so on, are almost issued in English. A qualified specialist in embedded systems is required with better English ability, especially in reading. However, the native language and course books for most students in some countries, such as China, are not English. Although they have English courses and tests, the students rarely learn their professional knowledge with English. Sometimes, the expression differences between two languages can make misunderstanding or trouble. For example, MCU, Micro-controller Unit[6], in English is discussed as SCM, Single Chip Microcomputer, in Chinese. If a student wants to collect useful information about SCM from English documents, he/she has to be disappointed. It is because

SCM just can be found in Chinese books but not in English professional references. Therefore, some graduated students often have problems when to update their knowledge with foreign language.

In order to work well, many students have to take more time to improve their professional English in their technical positions after graduations. In fact, it is responsibility of educational institutions for engineering students to be trained the English. As a skill to collect new concept and information, professional English is helpful for students in their technological innovation career [7]. Hence, bilingual education for lifelong learning of engineering students, especially in electronic and computer engineering, is significant. What is the bilingual education in a university? How is to take bilingual education in computer engineering? A solution in Chinese universities may be discussed bellow.

2 The Outline of Bilingual Education for Embedded Systems

The bilingual education in China is to take education in both of Chinese and English.

2.1 The Car of a Bilingual Education

If to assume bilingual education as a car, it is the body with two wheels, objects and procedure, such as Fig.1. Organization of a bilingual education includes two main parts, objects and contents. The organization can be indicated as Fig.2.

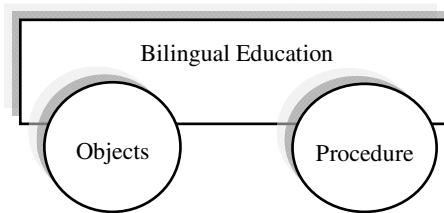


Fig. 1. The car of a bilingual education

The objects in bilingual education are faculty and students. What kinds of teachers and students are more satisfied for the education? Maybe there are various answers. But, at least, their English is ok for the education.

In Chinese universities, for example, the English level of students in one class may be different. Sometimes, the difference is bigger. The reason is that English language education is various for different schools or different area. For instance, some students start to learn English from kindergartens, some from primary schools, and some from secondary schools. Since the English foundation is not similar, students' ability, especially for students in year one and year two, to understand English contents of a course is more different. Therefore, the bilingual education for students of year one or two in most Chinese universities is more difficult.

Is it possible to estimate students' English level in China? The answer is positive. There is a rule in many universities for English. It is that a student in most tertiary institutions must pass CET4, College English Test 4, a national test. Otherwise, the student cannot obtain a Bachelor degree. In order to get their degrees, most students have to pass the CET4 in year 2 or 3. Hence, the students of year 3 or 4 are more suitable for bilingual education. Since students are passive in a lecture, they can be call passive objects in the education.

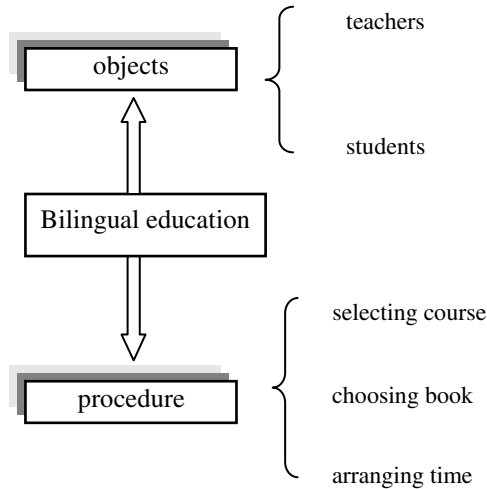


Fig. 2. The organization of a bilingual education

Opposite to passive objects, the active objects in bilingual education are teachers. After seventies of last century, thousands upon thousands Chinese go oversea to study, research and work. Some of them have gone back and taken important teaching positions in Chinese universities. Many of teachers with oversea experience are qualified as lecturers for bilingual education. It is because they studied their professional knowledge in Chinese and English, they familiar to bilingual expression in their fields.

The contents of bilingual education involve plan, practice and review.

2.2 Contents of Bilingual Education

There are three stages for contents or a procedure of bilingual education.

The first one is plan. There are three main sectors, selecting course, choosing book and arranging time, in this stage.

Selecting course means to choose a suitable subject according to the education objects, teachers and students. This is very important for the education to success. The course should be easier to be understood by students and should be familiar to

teachers. If it is difficult to be studied in native language, it should be more difficult to be learned with both native and foreign languages. On the other hand, the course had better to concern design or applications rather than theory. It is because design and applications with English are more helpful for the engineering students in their future jobs.

Choosing book or reference is an important work for bilingual education. For most courses, the sector is not easy. Actually, the course frame in China, for example, is different from some English countries in titles and contents. For instance, in China a course is called Single Chip Microcomputer Principle and Its Applications [8]. The similar course is named Embedded Systems [9] in other countries. The content of this course is mainly concerned with MCU-MCS51 in most Chinese universities [10]. But in some English countries, the contents are more different. Maybe microprocessors or other MCU are discussed more. Therefore, to find a suitable book or reference is not easy. Sometimes, for bilingual education, an accessible way is to use English data sheets of products as references[11].

Arranging time is to talk about lecture time for a selected course. Since the course is taught with two languages, it is sure to take more time for the similar contents. Normally additional 1/3 to 1/2 teaching time for a bilingual course is acceptable. If to keep same teaching time, the contents of the course may be decreased 1/3 at least. On the other hand, in order to get a more efficient result, the proportion of native and foreign language in a lecture should be arranged better. At the beginning, the lecture with foreign language takes less time, about 1/5 to 1/4. When students are used to the discussion in the class through the foreign language, the teaching by foreign language can take 1/2 class time.

Practice of bilingual education is composed of lectures, homework, experiments and tests.

Review is important to estimate and improve the bilingual education. This is a feedback procedure. In this stage, administration of school collects information from objects, faculty and students. In fact, an available estimation should be got from the graduated students yet. But it is more difficult. So far, the work can not to be done well.

3 The Bilingual Education for Embedded Systems

Embedded systems are the most popular computer applications [12], The relevant courses are compulsory to students for undergraduate or postgraduate degree in most Chinese universities. The bilingual education for computer applications or embedded systems is a kind of engineering education. To train ability of design, development and applications is a main target in an engineering institution. A practical procedure of bilingual education for a course of embedded systems is described as follows.

The plan of bilingual education for engineering consists of selecting course, choosing book and arranging time.

Selecting course is the first stage of education plan. The basic contents of selecting course are shown in Fig.3.

According to the analysis for students in above section, bilingual courses in engineering had better to be opened in year 3 or year 4 at Chinese universities. The reason is that most students have passed CET4 in year 3. Their English level is good enough for bilingual education. Normally there are 8 terms for students in universities for Bachelor degree. The final design is in the term 8. A course of bilingual education may be better taken in term 6 and 7.

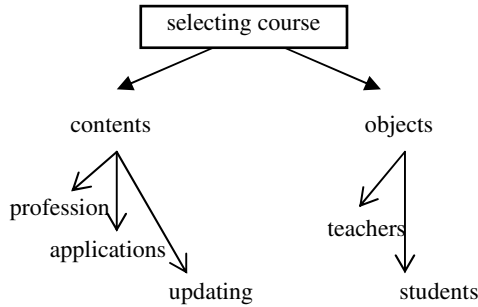


Fig. 3. The organization of a bilingual education

For example, the course, MCU and Its Applications, in term 7 is selected as a bilingual course. There are several reasons for the selecting. Firstly, the teacher is familiar to MCU with bi-language. It is easy to teach for the lecturer. Secondly, it is a professional course. Some basic knowledge has been introduced in the pre-courses. Its content is not very hard to be learned. Thirdly, its main teaching point is in design and applications. The English references can be got more easily. Finally, some MCU chips are often in updating. MCU system design is to refer to newest information in English frequently. The bilingual education of MCU is more useful and helpful for students.

Arranging time for MCU bilingual education is not very difficult. The course is issued in term 7. The teaching time of MCU for bilingual education is similar to single language. But its content may be cut down. It focuses on training skills of design.

Choosing book for MCU in fact SCM, course in Chinese is very convenient. There are many choices for MCU Chinese books, twenty to thirty versions at least. In order to learn easier, the contents of English books should be similar to Chinese. But the relevant English book is hardly to be found. Although some books concerned with MCU, the content about MCS51 almost cannot be found. In order to get a solution for the book choosing, to collect some references is necessary. As a replacement, product data sheet of 89C51 MCU is used as an English reference for the MCU bilingual education.

Review is important to estimate and improve the bilingual education. Assessment of experts and feedback of students are main parts of a review for bilingual education.

The experts' assessment is the opinions of education experts after listening some lectures.

There are two important points from students' feedback. One is aspiration. Most students expect to study more and improve their professional English through bilingual education. Another one is difficulty. Some of students feel hard in learning with poor English.

4 Conclusions

The bilingual education in China is to take education in both Chinese and English. The education involves education objects and contents. The objects are people. They are teachers and students. The contents include plan, practice and review. In plan, there are three main contents, selecting course, choosing book and arranging time. Practice of bilingual education is composed of lectures, homework, experiments and tests. Review is important to estimate and improve the bilingual education. It is a feedback procedure.

A course, MCU and Its Applications, in term 7 for a Bachelor degree, is selected as a bilingual course. According to assessment of experts, the attempt is positive and significant.

Acknowledgments. This research was supported in part by the National Natural Science Foundation of China under grant 60872057, by Zhejiang Provincial Natural Science Foundation of China under grants R1090244, Y1101237, Y1110944 and Y1100095. We are grateful to NSFC, ZJNSF and Huzhou University.

References

1. Sifakis, J.: Embedded systems design - Scientific challenges and work directions. In: Proceedings of DATE 2009, Nice, France, vol. 4, p. 2 (2009)
2. Liu, L.: Advanced Manufacturing of China and Embedded Systems. In: Proc. of CCIE 2010, Wuhan, China, vol. 6, pp. 81–84 (2010)
3. Pan, Z., Wells, B.: Hardware Supported Task Scheduling on Dynamically Reconfigurable SoC Architectures. *IEEE Trans. VLSI Sys.* 16(11), 1465–1474 (2008)
4. Wallner, S.: A configurable system-on-chip architecture for embedded and real-time applications: concepts, design and realization. *Elsevier Journal of Systems Architecture* 51, 350–367 (2005)
5. Nooshadadi, S., Garside, J.: Modernization of teaching in embedded systems design-an international collaborative project. *IEEE Transaction on Education* 49(5), 254–262 (2006)
6. Claasen, T.M.: An industry perspective on current and future state of the art in system-on-chip (SoC) technology. Proceedings of the IEEE 94(6), 1121–1137 (2006)
7. Karanian, B.A., Chedid, L.G.: 21st Century Trends That Influence Constructing Creative Classroom Environments. *IEEE Transaction on Education* 47(5), 157–159 (2004)
8. Wang, Y.: MCU Principle and Embedded Applications. Xiwang Press, Beijing (2002) (in Chinese)
9. Marwedel, P.: Embedded System Design. Springer, NY (2004)
10. Zhang, Y., Xie, W.: Single Chip Microcomputer Principle and Its Applications. Mechanical Industry Press, Beijing (2004) (in Chinese)
11. Moallem, M.: A Laboratory Testbed for Embedded Computer Control. *IEEE Transactions on Education* 47(8), 340–347 (2004)
12. Liu, L.: A Survey for Concept of Embedded Systems. *Advanced Materials Research* 63-64(7), 73–76 (2011)

Research on Negative Inventory Balances Management of Modern Manufacturing Industry under ERP Environment

LinKun Li¹, KaiChao Yu^{1,*}, and YiMing Li²

¹ Faculty of Mechanical and Electronic Engineering,
Kunming University of Science and Technology, Kunming 650093, China

² Kunming YUNNEI Power CO., LTD, Kunming 650024, China
glimmer_lin@live.cn

Abstract. In the manufacturing industry, each enterprise will encounter in the negative inventory balances problem in the warehouse management. The problems of negative inventory balances management, has become the bottleneck of processing flexibility of storage services. Allowing negative balances, it will facilitate delivery of cargo from storage rather than management. Does not allowing negative balances, it will facilitate management rather than delivery of cargo from storage. Aiming at this problem, this paper discusses the causes of negative inventory balances and the corresponding solutions, and we proposed some suggestions about treating the phenomenon of negative inventory balances.

Keywords: Negative Inventory Balances, Inventory Management, Manufacturing Industry, ERP.

1 Introduction

No matter early intervention in ERP, or business of inventory have entered the right path, negative inventory problem will occasionally appear in various stock affairs. Therefore many ERP systems offer a set of negative inventory parameters option to solve this problem. For example, the option of negative inventory balances in Oracle ERP determines whether inventory transactions can drive the inventory balances of an item negative [1]. Because of negative balances, sometimes, it is more difficult to face this problem for many enterprises in business process of inventory. And if negative inventory in bad control, it will cause aspects of logistics, information flow and funds flow disorder.

Some research has been carried on in the negative inventory balances field. But most of them are about retail business, and the methods by strict management are used to avoid negative balances, moreover, reducing the possibility of data error or in disorder in system for negative inventory balances [2]. A few scholars offered improved algorithm to deal with a series of problems brought by negative balances for

* Corresponding author.

minimizing such impact. On the contrary, there is an argument about enlarging the influence of the negative inventory balances. But the viewpoint is mainly at the enterprise by supply chain downstream superior position, using the fund flow from negative inventory balances for the goal of reducing the stock costs [3]. So in the following, the paper will be discussed about the reason of negative balances and method should treat the negative balances actually.

2 Negative Inventory Balances Phenomenon

Inventory is the physical stock of any item or resources used in an organization [4]. In the ERP, the system can record the every transaction and business data. But the hysteresis quality and incorrect with process of affairs record by clerk will be able to cause the stock for the negative situation. For example, in some urgent shipment case, warehouseman have marked the outbound order, but godown entry hasn't done by pruchasing department or manufacture department. In other words, material delivery charge has done before its receiving charge caused negative inventory balances. If following the provision, warehouse entry first, it would reduce the flexibility of material moving to sales departmen and production department, and it would not contribute to increased departmental performance in some case, then the performance of staff's work enthusiasm could be reduced. But increasing the elasticity of system, it will increase the risk of inventory inaccuracy. Therefore, whether permitting the negative balances or not, system operators must accord to actual conditions and causes of different negative inventory balances.

3 Negative Inventory Balances Caused by Convenient Operation

In order to facilitate production and sales, it could simplify the workflow, but in some case, the negative inventory balances will be happened or appear in ERP system, which leads to many unknown troubles for the operators.

3.1 Advance Material Issue Notes

In quantity production, material delivery from storage is according to the BOM. In order to reduce procedure batch and statistical information, material for delivery are all marked in one issue notes. And it will lead to that, in some case, the material in inventory balance which number of inventory count less than delivery become to emerge negative. These kind of negative balances are produced in this case — not-exhausted, advance-issue-tally, postpone-receipts-tally, not-practical-logistics.

3.2 Postpone Material Receipts Notes

Deal with first-out formalities, last-in formalities, negative inventory balances will inevitably arise in the stock affairs. For instance, as production of emergency, the material (purchased) without quality checking was delivered to production field. Because there was not quality proving, material receipts notes could not be marked. Warehouse keeper had to mark material issue notes first, and mark material receipts

notes afterwards. These kind of negative balances are produced in this case — exhausted, advance-issue-tally, postpone-receipts-tally and practical-logistics.

The negative inventory balances caused by two situations above, will show some deviations from the cost accounting in varying degrees. As price changing, the amount of inventory will be deviated in some case. Because of negative balances, the value of profit and loss of cost accounting will be fluctuated by quantity of storing item. Therefore, solving this negative inventory balances caused by convenient operation, there are two choices to enterprise managers:

1 For agreeing item negative, appending a new algorithm in ERP to eliminate the negative balance's effect and improve the flexibility of inventory transactions.

2 For forbidding item negative, the more management measures shall be carried out in general affairs to avoid negative balances and their troubles.

If agreeing item negative, new algorithm could solve the problem of cost accounting of stock from negative balances and ensure the accuracy of the calculation for stock affairs. It will be in background of system and flexible in standard procedure to solve specific cases like accounting management system as above but encouraging relaxing management. On the contrary, a serious of management measures as inventory transactions cannot drive the inventory balances of an item negative, can keep inventory balances active. But warehouseman and other department must accord to procedure of stock, and what reduce the convenience of work deal with any emergency.

4 Negative Inventory Balances Caused by Artificial Inaccuracy Operation

4.1 Wrong Bar Codes

When finding, getting code and pasting the bar code of warehousing material, warehouse keeper might make a mistake caused that the material need registered is different from the item represented by the bar code, what lead to the receiving material is changed. The number of original receiving material might be added to other coding by error, and when original material is issued, system of inventory management could show negative inventory balances up or error. To avoid such things happen, the warehouse keeper should get code from purchase orders or receipts planed orders instead of finding code by recognizing material.

4.2 Fault Storing Check

In some case, because of subtle distinctions of some material like name, specification or type, supplier might provide wrong item list, and warehouse keeper does not find it out when checking storing. The inventory data maybe inaccurate, and lead to negative inventory balances when picking. In addition, counting error, specifications calibration error and other check storing can also cause negative balances. Therefore, when storage department receive goods, enhancing the second acceptance check, warehouse keeper need to check the specifications and quantity of material and the consistency of order and other information of storing carefully. Once the difference between goods and

planned orders turns up, warehouse keeper should respond to the purchasing department, traced differences strictly for improving the accuracy of material acceptance.

4.3 Out-of-Warehouse Error

When warehouse keeper make the issue notes, and if he swipes the wand across the wrong bar or type the wrong material information, it might cause the negative balances in system. In order to avoid this case, it is necessary to improve the management measures and staff training. The warehouse keeper must check the corresponding relationship between the original material information and issue notes especially exchange purchase material, and increasing approve process if necessary.

4.4 Material of Different Suppliers Disordered

The material and suppliers are many-to-many relationships. The material need be distinguished by suppliers when purchase storing, storage and delivery if they track suppliers in system. In the event of material of different suppliers disordered, it will appear inaccurate inventory phenomenon which causes the error quantity of material belong to different suppliers. For instance, the quantity of material A of supplier A was 10 which would be delivered to other department, because of warehouseman's fault, the material A of supplier B was delivered. But at that moment, system deducted 10 materials A of supplier A, the ending inventory in system is 0. Actually, the inventory of material A of supplier A had not changed. When material A of supplier A delivered again, the ending inventory in system would be -10 which show the negative balances. For this problem, it is usually due to material does not set out by standard, and it is not easy to distinguish the difference of material between various suppliers. The treatment of this problem is to set different area marks as suppliers, and keep the accuracy of original delivery material to avoid inventory quantity storage error.

4.5 Stock Taking Error

The mistake in stock taking, book value is different with actual inventory in system, will lead to negative balances also. So warehouse keeper should record the stock in& out timely and accurately, ensure the smooth operation of stock checking daily and monthly, keep account, product and card in the same state. It is useful to take small range and high frequency inventory methods to reduce mistakes.

5 Deal with Negative Inventory Balances

All the adjustment of inventory information in ERP or other MIS would cause negative inventory balances. To deal with the problem caused by artificial inaccuracy operation, it is necessary to intensify staff training and strengthen management for reducing the incidence of these mistakes which should be eliminated because it's unreasonable to negative balances. To deal with the problem caused by convenient operation, the decision maker needs to determine whether to allow the negative inventory balances by actual inventory business. If take prohibitive delivery measures, strengthen supervision on warehousing, supervise on stock receiving in ERP system, especially when

inventory of some material in system down to zero, and according to per plan to arrange operation, it will delay production and sales in a certain extent, and it is disadvantageous for the urgent supply material to production.

Therefore, at the beginning of operation of new ERP or MIS, the method to allow negative inventory balances by increasing the initial algorithm and temporary manual bookkeeping, will be propitious to inventory management and moreover production and sales, not be decided by kinds of system operation process and rigid regulation, which can be used to weaken or prohibit the negative inventory balances if necessary by the time system has a good stability and reliability. Negative inventory balances are not exist in realistic inventory but only inventory information in ERP or account book if permitted. In case of inevitable negative inventory balances, it is necessary to improve the rationality and validity of initial algorithm and correctly understand the negative inventory balances.

Negative inventory balances are a double-edged sword, which can be used equally for inventory management better or worse. Used correctly, they can simplify the process and improve the flexibility of inventory business. Instead, they can complicate the inventory information processing in ERP and lead to inaccurate stocks between actual inventory and account, moreover inventory costs estimate combined effects of holding costs, setup costs, ordering costs, and shortage costs [5].

References

1. Oracle Inventory User's Guide Part No. A83507-10, [DB/OL] (February 2011), http://download.oracle.com/docs/cd/B25516_18/current/acrobat/115invug.zip
2. Wang, Z.: Analysis of Negative Inventory. *China Computer & Communication* 4 (2004)
3. Chen, G.: Crossing Zero inventory and Create Negative Inventory. *Fujian Quality Management* 5 (2005)
4. Verma, R., Boyer, K.K.: Operations & Supply Chain Mangement: Word Class Theory and Practice. In: Cengage Learning, 1st edn., vol. 196 (2009)
5. Robert Jacobs, F., Chase, R.B.: Operations and Supply Management: The Core, pp. 311–312. The McGraw-Hill Companies (2009)

Design of Automatic Control System of Fructose Crystallization

Jian Chu and Qing Lu

School of Autocaton and Electrical Engineering
Tianjin University of Technology and Education Collage,China
Chujian6@126.com, bjxyluqing@126.com

Abstract. At present, the production processes of fructose mainly have realized automatic, such as press, lustration, but except crystallization[1]. Fructose crystallization process control operations are achieved by experienced operatives operating under the personal skills and experience fructose, reducing the product yield and quality. Therefore, the crystallization process has become the decisive process of fructose production processes, the quality of work results process directly determines the quantity and quality of fructose, which also determines the sugar industry in economic benefits. Whereby the system study how to achieve automatic control of fructose crystals.

Keywords: fructose crystallization, temperature control, PID control.

1 Introduction

1.1 Background Research

In China, the average consumption of sugar is the lowest in Asia, even the world[2], the sweeteners market has very broad space for development, which expected that sweetener industry is bound to usher continued and rapid growth high tide while rapid economic development of our country. At present, China's sugar demand, strong prices will not ease the situation in the short term, food regulations and consumer bias limit the efficient application of sweeteners, People's pursuit on health and natural foods is growing strongly. Therefore, the crystallization of sugar and other natural nutritive sweeteners will be greatly respected and popular. In the sugar industry, the forefront of the world are Denmark, the United States, France, Australia etc., these developed countries, sugar production have automated largely, which are also big consumers of sugar[3]. China is far behind these developed countries in this area. China's sugar industry is still in a state of manual operation. Therefor, it is increasingly important to achieve automatic control of sugar industry.

1.2 Recent Research at Foreign and Domestic

1.2.1 Study the Situation of Foreign Production

In the early 1920s, the United States began to use the technology of acid large-scale to prepare glucose syrup and fruit glucose syrup etc[4].

In 1959, enzymatic production of glucose achieved success, which was not only a major innovation in glucose industry, but also the first leap in the history of fruit glucose production, and promoted the great development of the sugar industry.

In 1960, Marshall won the patent that transformed glucose to fructose using xylose isomerize.

In 1974, United States immobilized isomerize firstly, which reduced the costs in the fruit glucose syrup production largely.

In 1980, the company UOP of U.S. issued patents, using a molecular sieve to separate fructose, and combining with the simulated mobile bed, so that the separation of fructose and glucose achieved industrialization[5].

1.2.2 Study the Situation of Domestic Production

China began to develop the fruit glucose syrup until 1970s, owing to backward production technology, high product cost, its development was limited. Until the mid-90s, the new sugar technology and equipment continued to improve, making the product cost greatly reduced; at the same time, as Chinese sugar policy adjusting in 2000, sugar prices began to rise, glucose instead of sucrose being used in food gradually revealed the advantages, the development of fruit glucose syrup in China ushered a rare opportunity[6].

In 1992, the Research Institute of Daqing Petrochemical Company and Nanning Cassava Technology Development Center developed the simulated moving bed that was using inorganic molecular sieve to separate and prepare the high purity fructose, which filled the domestic blank of crystalline fructose industrial production technology.

Into the twenty-first century, the development of fruit glucose syrup was relatively rapid in china, In 2001, a 10 million ton / year production line of high-fructose corn syrup was established in Anhui; In 2002, the Dacheng Group of Changchun and Cargill of the U.S.have built joint venture in Shanghai, that 10 million tons level of fructose production plant was built, of which products and quality met international standards, Mainly used in main beverages and other food production in international and domestic[7].

1.3 Research

The study employs a combination of cooling and dilution to prepare crystalline fructose, intending to resolve the existing problem fructose crystallization process, such as low yield, poor product quality etc. To get an optimum process of cooling crystallization of fructose.

This study includes:

(1)Introduces the research status of fructose in foreign and domestic, indicates the content, purpose and research significance of fructose research.

(2)Though experiments, in-depth study the impact that all aspects of automatic control system have had on this system, familiar with the process of experimental control.

(3) Optimize the process of powerful glucoamylase to degrade oligosaccharides and fructose cooling - solvent crystallization process.

2 Crystalline Fructose Overall Design of Automatic Control System

The control block diagram of crystal automatic control system as shown in Figure 1. Where A / D conversion, PID regulator, D / A conversion operations are performed inside the PLC, as well as providing available data for PC to display. PC can also be assigned to the appropriate address, to set the relevant parameters.

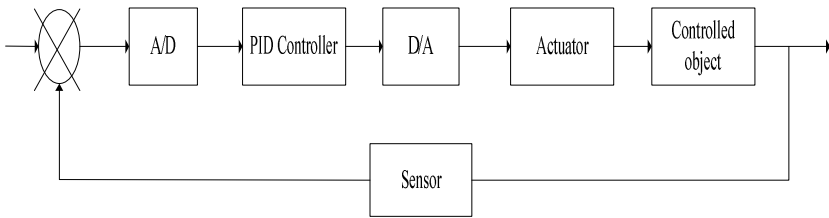


Fig. 1. Crystal automatic control system

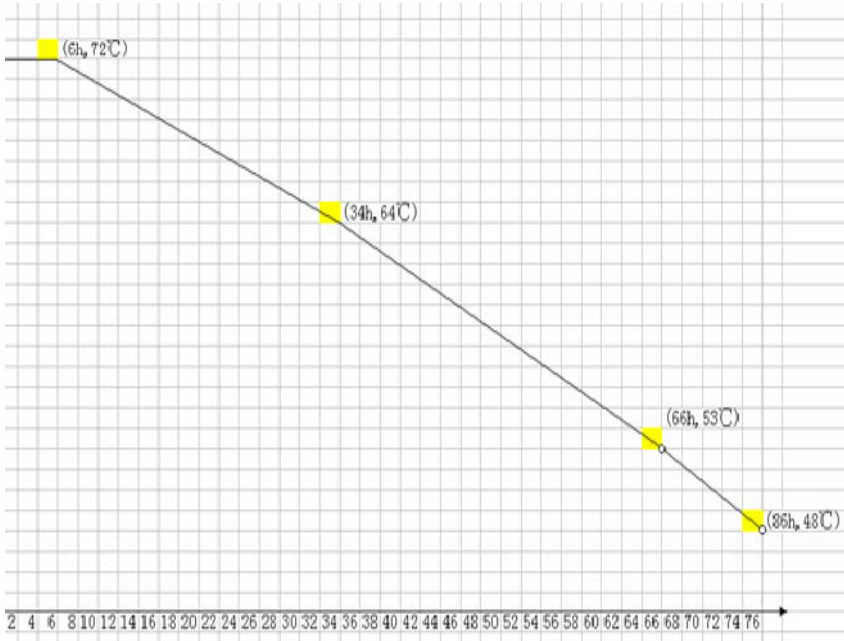


Fig. 2. Temperature experience curve

Input signal of the system is circulating water temperature, and temperature in accordance with the experience curve (Figure 2) to control the temperature of the system in order to control the crystallizer temperature. Crystallizer temperature measurement uses the Pt100, its resistance - temperature relationship is very good linearity and have become very close to the linear line in the temperature range -200°C ~ 650°C .

3 The Determination of Control Scheme

This chapter systematically studies that adding ethanol to reduce the viscosity of fruit syrup, confirms the amount of ethanol in the operation of crystalline fructose; determines the basic data of metastable zone width that fructose crystallized in water and ethanol solution; optimizes the process that the powerful glucoamylase degrade oligosaccharide of fruit syrup, reduces the viscosity of high-fructose corn syrup; optimizes various factors that affect the process of fructose crystallization.

On the one hand, fructose is heat-sensitive material, poor thermal stability, YANG Rui-jin and other studies have shown that fructose solution maintained at 65°C for 6 hours, a greater increase in color value[8]. On the other hand, the solubility of fructose in water with temperature changes significantly and the solubility of fructose in organic solvents such as ethanol, is much smaller, these characteristics determine the fructose is not suitable for crystallization by evaporation but to the cooling crystallization or solvent crystallization. From the perspective of safety and cost, this study chose ethanol as solvent, to adopt cooling method that dilution aided to prepare crystallization process of crystalline fructose.

As the experiment is based on visual observation mutations that fructose corn syrup from transparent to turbid as a judge the basis of arising spontaneous nucleation, it is inevitable lags. Moreover, the literature[9]reported that crystalline fructose with two or half of crystal water separate out more easily than the anhydrous crystalline fructose, but their metastable zone boundaries blur with anhydrous crystalline fructose, the higher degree of supersaturating may make the operation into the unstable zone of crystal fruit sugar crystals that makes a lot of precipitation. This phenomenon should be avoided, therefore, the operation of crystalline fructose is suitable under low supersaturating. The supersaturating generally are controlled in the $1/2$ of metastable zone width turn to solubility curve place in the industrial crystallization. Viscosity of saturated solution from 0 to 20% (v / v) of ethanol solution decreases very rapidly, then continue to improve the ethanol ratio, that decreases slowly, while the solubility decreases uniform, there is no turning point. In order to avoid entering the unstable areas and out of consideration for saving ethanol and reducing the cost, the preliminary view is that selecting 20% (v / v) ethanol levels is more appropriate.

Acknowledgments. This work is supported by the project of number KJ11-24.

References

1. Wei, L., Fei, H., He, Y.: A New Type of Sugar Source Crystallization Fruit Glucose. *China's Food Industry* (10), 52–53 (2006)
2. Zhang, L.: The Development of Fructose Product. *Food and Fermentation Industries* 24(1), 51–55 (1998)
3. Zhang, L., Lin, X.: China's Current Development of High Fructose Corn Syrup. *Food Science & Technology* 18(3), 48–49
4. Jurgens, H., Haass, W., Castaneda, T.R.: Consuming fructose-sweetened beverages increase body adiposity in mice. *Obesity Research* 13(7), 1146–1156 (2005)
5. Jin, S., Ying, L.: Development of High-purity Fructose and Crystalline Fructose. *Starch and Starch Sugar* (4), 17–20 (1999)
6. Yang, H.: Properties and Applications of High Fructose Syrup. *Food Science* 23(2), 154–156 (2002)
7. Yue, H.: High Fruit Syrup and other Sweetener in Food Processing of Choice. *Food Industry* 3, 15–17 (2005)
8. Yang, R.: The Nature of Pure Crystalline Fructose and Its Application. *Cold Drinks and Quick-frozen Food Industry* (1), 27–29 (1997)
9. Gao, Z., Yue, T., Yuan, Y., et al.: Fructose Production Technology and Application of Research Progress. *Northwest A & F University Academic Papers* 31, 187–190 (2003)

Text Clustering Method Based on Improved Genetic Algorithm

ZhanGang Hao, Tong Wang, and XiaoQian Song

Shandong Institute of Business and Technology, Shandong Yantai 264005, China

Abstract. In this paper, we put forward a new genetic algorithm called GA-K algorithm by putting k-medoids into the genetic algorithm, then we form a local Optimal Solution with multiple initial species group, strategy for crossover within a species group and crossover among species groups, using the mutation threshold to control mutation.

Keywords: Text clustering, K-medoids algorithm, genetic algorithm.

1 Introduction

Genetic algorithm is a random optimization algorithm based on natural selection and genetics. [1] It has been tried to use genetic algorithm to solve partitioning problem, [2-4] resulting in either undesirable outcome or failure to solve the sloe point problem. Someone has achieved some achievement using combination of k-medoids and genetics algorithm to make partitioning and text partitioning, however, the isolated points problem remains unsolved. This paper proposes a new GA-K algorithm that puts k-medoids into the genetic algorithm and improves the crossing factor and changing factor of the genetic algorithm. This method can better solve both the problem of local optimization and outlier.

2 Brief Introduction to k-Medoids Algorithm

The primary idea of the k-medoids algorithm is that it firstly needs to set a random representative object for each clustering to form k clustering of n data. Then according to the principle of minimum distance, other data will be distributed to corresponding clustering according to the distance from the representative objects. The old clustering representative object will be replaced with a new one if the replacement can improve the clustering quality. A cost function is used to evaluate if the clustering quality has been improved. The function is as follows:

$$\Delta E = E_2 - E_1$$

where ΔE denotes the change of mean square error; E_2 denotes the sum of mean square error after the old representative object is replaced with new one; E_1 denotes the sum of mean square error before the old representative object is replaced with new

one. If ΔE is a minus value, it means that the clustering quality is improved and the old representative object should be replaced with new one. Otherwise, the old one should be still used.

3 Text Clustering Method Based on Improved Genetic Algorithm

Although genetic algorithm has such shortcomings as slow convergence, can not converge to global optimization etc, it has been approved to be a good optimization algorithm, which has been widely used since its introduction. While this paper puts k-medoids algorithm into the genetic algorithm and improves the crossover operator and mutation operator. By using the genetic algorithm, we can solve the problem of slow convergence, can not achieving the global optimization, and the outlier of text clustering.

3.1 Encoding

Suppose we divide dataset n into k subsets, then we may consider two denotation ways. The first one is that we can classify a certain text into a certain category, after all the texts have been classified, the denotations representing text file and the category it belongs to form a chromosome. For example, a chromosome is denoted as

$r = \{r_{12}, r_{23}, \dots, r_{ij}, \dots, r_{nk}\}$, r_{ij} denotes i articles that belong to the j th category, $1 \leq i \leq n, 1 \leq j \leq k$. The second way is to form the chromosome by the center of

each clustering. For example, a chromosome is denoted as $r = (r_1, r_2, \dots, r_k)$, r_i denotes the i th clustering center is r_i . In general, there are many text files as clustering, if the first way is used, the chromosome will be too long, resulting in increased difficulty for crossover and mutation. So the second way is adopted by this paper.

There are many encoding methods associated with genetic algorithm, such as Binary coding, Grey coding, Real number coding, Notation coding, Parameter coding, Crossover coding and coding based on genetic formulation[6], etc. According to the characteristics of the second denotation way, this paper chooses real number coding as the way to encode.

3.2 The Formation of Initial Species Group

Initial species group can be generated by random function to form an initial group matrix. However, this matrix is so random that the quality of chromosome of the whole species group can't be ensured. Therefore, this paper adopts k-medoids algorithm to optimize the species group generated randomly, resulting in a new species group matrix as the initial species group matrix of the genetic algorithm.

For example, we perform clustering VI category against 100 text files to generate initial species group with real number coding. The species group matrix generated by random function is shown as follows:

$$\begin{pmatrix} 1 & 3 & 5 & 2\ 6 & 4\ 0 & 5\ 0 \\ 2 & 4\ 1 & 2\ 3 & 4 & 6\ 5 & 7\ 6 \\ 3 & 7 & 9\ 0 & 8\ 4 & 3\ 2 & 6\ 7 \\ 8\ 2 & 4\ 5 & 7\ 0 & 9 & 8 & 1\ 5 \end{pmatrix}$$

This is a species group matrix comprising 4 rows and 6 columns, which denotes that 4 rows mean there are 4 individuals in this species group and 6 columns mean each chromosome contains 6 genes, which is the category number that need to be clustered. Optimizing this initial matrix by k-medoids algorithm result in the following species group matrix

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{pmatrix}$$

This is the initial species group matrix. To prevent the genetic algorithm from converging into partial optimization, diversity of the species must be strengthened. This paper uses the idea of small biological environment genetic algorithm to generate a number of initial species groups, the number of which is the same as the clustering number. That is if 4 categories are clustered, 4 initial species groups are generated. The initial species group of each generation is generated by this method.

3.3 Fitness Function

This paper uses mean square error as fitness function and defines

$$E = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2$$

E is the mean square error sum of all data objects and corresponding clustering center, p is a point in the space representing objects and m is the mean value of clustering C_i . The fitness function used by this paper meets the major conditions required for designing fitness function.

3.4 Genetic Operator

This paper uses roulette selection, the basic idea of which is that the probability of each individual is selected is positively proportional to its fitness. The specific operation is expressed as follows:

$$p(a_j) = \frac{f(a_j)}{\sum_{i=1}^n f(a_i)}, \quad j = 1, 2, \dots, n$$

$p(a_j)$ denotes the probability the j th individual is selected, $f(a_j)$ denotes the value of fitness function of the j th individual and n denotes the number of total individuals.

In this paper, the genetic algorithm adopts multiple initial species group strategy, the crossings of which include crossover within a species group and crossover among species groups. For crossover within a species group, to retain sound gene fragment, this paper uses single point crossover, which represents a crossover point randomly set in an individual coding string, and then portion of the genes of pairing individuals is exchanged at this point. For crossing among species groups, for every 50 generations evolved, random pairing crossing among species groups are performed (for odd groups, the group left after all the other even groups have finished pairing go into the next cycle), single point crossing is also adopted by crossing among groups. Crossing rate normally takes 0.4-0.9[7].

In order to retain the diversity of species groups, mutation operators are needed. However, mutation might destroy valuable genes. Hence, this paper sets a mutation threshold ∂ . Before mutation, a random number should be produced. If this number is greater than ∂ , mutation happens; if this number is not greater than ∂ , then the genes are retained without mutation occurring. Crossing rate normally takes 0.001-0.1[7].

3.5 Criteria to Stop the Algorithm

The first one is to fix the maximum genetic algebra. The algorithm stops as the maximum algebra appears. The second one is according to the degree of convergence. The algorithm stops as the mean fitness of the species group undergoes no change after a few generations.

4 Experimental Analysis

This paper picks up 505 articles in 6 categories from CQVIP as experiment data. The first 5 categories contain 100 articles each and the last category contains 5, which form the isolated points. The first 500 articles for the experiment are sourced from <http://www.lib.tju.edu.cn>. The 5 categories are Finance, Aviation, Environment Science(ES), Construction Materials(CM) and Metallurgical Industry(MI) respectively. The last category is current affair and news sourced from <http://www.baidu.com/>. After having undertaken basic treatment and dimension reduction to these files, k-medoids algorithm and GA-K algorithm are used for clustering analysis.

4.1 Experiment 1

First, k-medoids algorithm is used for clustering analysis. The results are shown in Table1.

Table 1. Results from k-medoids algorithm

	CM	Aviation	ES	MI	Finance	CAN
Wrong articles	59	56	49	56	53	2
Correct articles	45	41	51	47	42	4
Percentage of correct ones	43	42	51	46	44	67
Time(second)	35.3					

4.2 Experiment 2

Then, GA-K algorithm is used for clustering analysis. The results are shown in Table 2.

Table 2. Results from GA-K algorithm

	CM	Aviation	ES	MI	Finance	CAN
Wrong articles	12	13	8	9	9	0
Correct articles	89	87	94	91	88	5
Percentage of correct ones	88	87	92	91	90	100
Time(second)	20468					

Comparison of the results from 2 methods indicates that the clustering effect of k-medoids is poor despite of a short time used. While GA-K algorithm proposed by this paper shows an excellent clustering effect with strengthened capability in searching ideal targets in spite of a lot of time consumed and its capability in searching isolated point is also very good.

5 Summary

Text clustering is widely used in real world and an important subject for data mining. Both k-medoids and genetic algorithms can be used to study it although each method has shortcomings. This paper embeds k-medoids algorithm into genetic algorithm,

proposing new tactics for initial species group, crossover and mutation as well as a new algorithm GA-K. This algorithm increases the diversity of species groups, enhances genetic algorithm's capability to search ideal targets and improves clustering accuracy and its capability to acquire isolated points.

Acknowledgements. This paper is supported by the National Natural Science Foundation of China (Grant No.70971077), Shandong Province Doctoral Foundation(2008BS01028), Natural Science Foundation of Shandong Province(Grant No.ZR2009HQ005, ZR2009HM008).

References

1. Fogel, D.B.: An introduction to simulated evolutionary optimization. *IEEE Trans. Neural Network* 5(1), 3–14 (1994)
2. Jones, D.R., Beltramo, M.A.: Solving partitioning problems with genetic algorithms. In: *Proc. 4th Int. Conf. Genetic Algorithms*, pp. 442–457. Morgan Kaufman, San Mateo (1991)
3. He, T.-T., Dai, W.-H., Jiao, C.-Z.: Research of Text Clustering Based on Hybrid Parallel Genetic Algorithm. *Journal of Chinese Information Processing* 21(4), 55–60 (2007)
4. Qin, X., Yuan, C.-A.: Text clustering method based on genetic algorithm and SOM network. *Computer Applications* 28(3), 757–760 (2008)
5. Bhuyan, J.N., Raghavan, V.V., Elayavalli, V.K.: Genetic algorithm for clustering with an ordered representation. In: *Proc. 4th Int. Conf. Genetic Algorithms*, pp. 408–420. Morgan Kaufman, San Mateo (1991)
6. Yuan, C., Tang, C., Wen, Y., et al.: Convergence of genetic regression in data mining based on gene expression programming and optimized solutions. *International Journal of Computer and Application* 28(4), 359–366 (2006)
7. Wang, X.-P., Cao, L.-M.: *Genetic algorithm: Theory, application and software realization*. Xi'an Communication University Press, Xi'an (2002)

Layout Optimization Method of Traffic Information Data Acquisition System Based on Linear Programming

Wei Li

School of Electronic and Control, Chang'an University, Xi'an, Shaanxi, 710064, China
zurish1975@163.com

Abstract. The purpose of the research is scientifically and accurately determining layout scheme of traffic data acquisition system. The paper proposes layout optimization method of traffic information data acquisition system based on linear programming using mathematical statistics and linear programming theory, and determines the value method of stationing coefficient. Accordingly layout scheme of traffic data acquisition equipment on various road sections is obtained. The method not only can achieve accurately and quantitatively optimal layout, but also reduce the large number of acquisition equipments, save human power and material resources, and reduce the cost. The research leads to the further development and improvement of intelligent transportation system used on national road and provincial road.

Keywords: traffic information, data acquisition system, layout optimization, linear programming.

1 Introduction

Traffic problems are becoming increasingly common with the increase of highway mileage. Intelligent transportation system is an effective way to solve traffic problem. Traffic data acquisition system is a key subsystem of intelligent transportation system and the foundation of the development of intelligent transport system and the premise of implementation of intelligent transportation management [1]. The research at home and abroad on traffic data acquisition system mainly focused on building various data acquisition system platforms and proposing data acquisition methods, such as traffic information acquisition system based on sensor networks [2], data acquisition system based on floating car [3], and circular coil, ultrasonic, microwave, infrared, video vehicle inspection equipment and so on [4-5]. Most of the current study of intelligent transportation is for urban road, while the research on national road and provincial road around city is poor. Meanwhile the research on layout optimization is also little. China's road transportation coverage area is wide, mileage is long, and road network density increases. If acquisition equipments are laid on each node, it will be a costly project and consume large amounts of materials, and bring forth great waste. Therefore, the paper finds correlation between network nodes by correlation theory of mathematical statistics and linear programming method, so as to achieve optimal layout of traffic data acquisition system.

2 Layout Optimization Model Building of Traffic Information Data Acquisition System Based on Linear Programming

2.1 Correlation Coefficient

Traffic flow shows similarity. The correlation coefficient (ρ_{xy}) [6] can be used to determine which road sections show similarities, how are similar. The similarity of traffic flow can be shown by single-variable equation or multiple linear equations. It found that single-variable linear equation is the most accurate for expression of traffic flow characteristics according to observation and statistics of traffic flow on national road and provincial road in X city. The equation of Correlation coefficient (ρ_{xy}) is:

$$\rho_{xy} = \frac{Cov(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \quad (1)$$

There,

$$Cov(X, Y) = E(XY) - E(X)E(Y); D(X) = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}; D(Y) = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2}.$$

In correlation coefficient (ρ_{xy}), X represents the traffic volume of road section X; Y represents the traffic volume of road section Y. Traffic volume data of road section X and Y must be obtained from corresponding to the data in the same time period. And traffic volume data must be large enough sample to support, otherwise the calculated result will be disperse and not accurate and so on.

2.2 Threshold Value of Similarity (α)

Threshold value of similarity (α) is that traffic flow of two road sections can be considered linear correlation when how much does correlation coefficient (ρ_{xy}) reach. The greater threshold value of similarity (α), the higher correlation accuracy requirement is. The data without data acquisition system is deducted by detected road section, and this data has greater accuracy. The road sections needing to lay data acquisition system are more. "1" expresses correlation, "0" expresses no correlation. When the correlation is identified as "1", it means the traffic flow can be calculated through linear regression equation between the two road sections. That is only laying acquisition system on any one of two road sections.

2.3 Similar Matrix

Similar matrix consisting of "0" and "1" of all road sections can be obtained according to the result of correlation coefficient (ρ_{xy}) and threshold value of similarity (α).

2.4 Linear Programming Model

Linear programming is an optimization model of operations research. It is aimed at seeking maximum or minimum problems of a linear function under meeting a set of linear equations or inequalities [7-9]. The standard form is:

$$\begin{cases} \min & c^T x \\ \text{s.t.} & Ax = b, x \geq 0 \end{cases} \tag{2}$$

There, $c = (c_1, c_2, \dots, c_n) \in R^m$ is a constant vector; $x = (x_1, x_2, \dots, x_n)^T$ is a column vector; A is an $m \times n$ -dimensional matrix $[a_{ij}]$; $b = (b_1, b_2, \dots, b_n)^T$ is a column vector.

According to the above theory, linear programming model of traffic data acquisition system can be described as follows:

$$\begin{cases} \min & z = \sum_{i=1}^n C_i X_i \quad i = 1, 2, \dots, n \\ \text{s.t.} & \sum_{j=1}^n a_{ij} X_j \geq 1 \quad i, j = 1, 2, \dots, n \end{cases} \tag{3}$$

There, n is the number of road sections; i, j is road section i , road section j , respectively. X_i is the variable of section model; C_i is comprehensive stationing coefficient of section i ; a_{ij} is similar matrix.

Stationing situation: “ $X_i=1$ ” expresses stationing; “ $X_i=0$ ” expresses no stationing;

Similar situation: “ $a_{ij}=1$ ” expresses similar; “ $a_{ij}=0$ ” expresses dissimilar.

Equation 3 indicates the optimal number of stationing when traffic flow of each road section is detected by one data acquisition equipment at least. There, C_i reflects the importance of stationing on the road section. C_i is associated with many factors, such as grades of highway, traffic volume, road network topology, the sensitivity of road, traffic composition and complexity of intersection, traffic management measures, application of collected data, available comprehensive condition of data acquisition equipment stationing and so on. For critical road sections, data acquisition equipments are directly laid to improve the accuracy of data. It will help control the accuracy of overall data. The road sections with relatively well-equipped facilities also should directly lay data acquisition equipment. It can make full use of existing conditions. The influence factors on C_i are numerous, and the correlation is high, so C_i is comprehensively determined by classification assignment method. The main factors divided into four categories, shown in Table 1.

Table 1. Influencing factor and value of C_i

1				2				3			4		
Road grade				Traffic volume(20min)				Sensitivity			Traffic complexity		
first	second	third	fourth	50~150	150~250	250~350	350~450	excellent	medium	inferior	great	middle	little
4	3	2	1	4	3	2	1	3	2	1	3	2	1

Note: The influence of other factors on C_i is relatively weak, so it can be ignored.

3 Application of Layout Optimization of Traffic Information Data Acquisition System Based on Linear Programming

For instance, layout optimization is executed in data acquisition region consisting of a few intersections, shown in Figure 1. Arrow indicates driving direction. Each transverse line represents a road section and marked with number. There are 45 road sections in Figure 1.

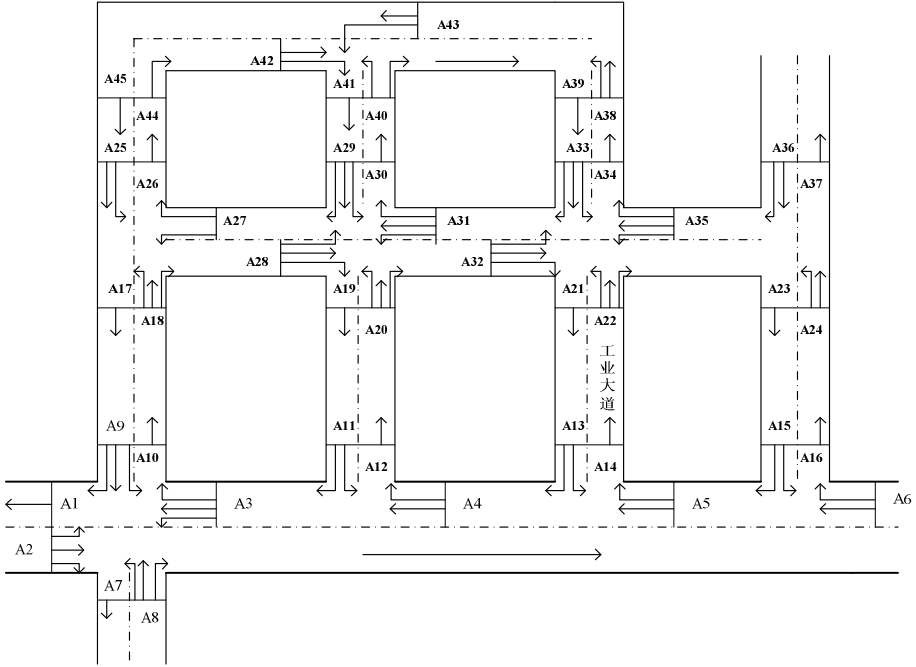


Fig. 1. Sketch map of layout optimization of traffic acquisition system

(1) Correlation coefficient is calculated according to Equation 1, taking section A4 and A12 for example. The calculated correlation coefficient (ρ_{xy}) is 0.936. The calculation method of correlation coefficient of other road sections is the same.

Table 2. Daily traffic volume data in 12 hours of A4 and A12 road sections (veh)

Time bucket	A4	A12	Time bucket	A4	A12	Time bucket	A4	A12	Time bucket	A4	A12
8:00-8:20	575	107	11:00-11:20	621	119	14:40-15:00	731	119	18:20-18:40	412	60
8:20-8:40	651	134	12:00-12:20	684	120	15:00-15:20	694	123	18:40-19:00	495	76
8:40-9:00	682	144	12:20-12:40	649	107	16:00-16:20	626	110	19:00-19:20	568	104
9:00-9:20	716	155	12:40-13:00	615	129	16:20-16:40	577	104	11:20-11:40	548	97
9:20-9:40	840	170	13:00-13:20	783	166	16:40-17:00	641	114	11:40-12:00	608	114
9:40-10:00	795	150	13:20-13:40	716	149	17:00-17:20	585	106	15:20-15:40	564	119
10:00-10:20	748	155	13:40-14:00	769	141	17:20-17:40	523	90	15:40-16:00	579	93
10:20-10:40	635	116	14:00-14:20	724	138	17:40-18:00	462	94	19:20-19:40	535	104
10:40-11:00	585	110	14:20-14:40	791	142	18:00-18:20	429	70	19:40-20:00	509	91

(2) Threshold value of similarity (α) is taken as 90% in order to make the data more accurate. If the correlation coefficient is greater than 90%, threshold value of similarity is taken as "1", or taken as "0". 45*45-dimensional similar matrix can be obtained. Because 45*45-dimensional matrix contains 2025 elements, 16 intersections (A1~ A16 road sections) are taken for calculating in order to facilitate calculation. Similar matrix is shown in Table 3.

Table 3. Similar matrix when threshold value of similarity as 90%

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0
4	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0
5	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0
6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
7	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
10	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
11	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
12	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
13	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
15	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0
16	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1

The C_i values of 16 road sections are shown in Table 4 according to influence factors of C_i value in Table 1.

Table 4. C_i value

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16
2	4	4	3	3	3	1	3	2	1	2	1	3	2	2	1

C_i values in Table 4 and similar matrix in Table 3 are substituted into Equation 2, and the result is obtained through programming using MATLAB. Minimum of z is 20, stationing number N is 9. Stationing is located in A1, A2, A4, A5, A7, A8, A10, A11, A16 road sections. It shows that laying data acquisition equipments in only 9 sections of 16 sections can meet the requirement after optimization. It nearly reduces 50% of the workload and saves a lot of equipment investment. It can be described that there is no need laying data acquisition equipment in each road section by scientific and reasonable research. Otherwise, it will have complicated data duplication, and increase workload and increase cost.

4 Conclusion

Layout optimization model of traffic information data acquisition system is built by mathematical statistics theory and linear programming method, according to traffic flow characteristics. The value method of stationing coefficient is determined. Finally layout scheme of traffic data acquisition equipment in various road sections is scientifically obtained. Layout basis of data acquisition system is quantitatively studied. The research ensures the integrity and accuracy of data, while reduces the layout number of data acquisition equipment, saves a lot of human power and material resources, and effectively prevents construction waste.

References

1. Huang, Y., Wang, J., Cai, B.: Traffic Information Collection System Based on Wireless Sensor Network. *Modern Electronics Technique* (23), 158–160 (2010)
2. Qin, A., Li, S., Nie, S.: Traffic Data Collecting and Processing Technology Based on CDMA. *Journal of Chongqing Jiaotong University (Natural Science)* 29(2), 276–279 (2010)
3. Jiang, G., Zhang, W., Chang, A.: Data Organization Method for Traffic Information Acquisition System based on GPS-equipped Floating Vehicle. *Journal of Jilin University (Engineering and Technology Edition)* 40(2), 397–401 (2010)
4. Gao, M.: Introduction of Non-intrusive Traffic Data Collection Technology. *Transport Standardization* (23), 31–33 (2010)
5. Yu, Y.: Study on Traffic Evaluation and Intelligent Control Technology for Arterial Highway. Master Thesis of Chang'an University, Xian (2010)
6. Han, X., Yang, R.: Probability and Mathematical Statistics. Beijing University of Posts and Telecommunications Press, Beijing (2003)
7. Cheng, L.: Operational Research Model and Method Tutorial. Tsinghua University Press, Beijing (2000)

Application of MUSIC Algorithm to Scenario-Based Training Database Construction Method of ISAR Images

Sang-Hong Park, In-O Choi, Jae-Heung Ye, and Young-Jun Kim

Department of Electronic Engineering, Pukyong National University (608-737),
Daeyeon 3-Dong, Busan, Korea
radar@pknu.ac.kr

Abstract. In this paper, the multiple signal classification (MUSIC) algorithm was applied to the efficient scenario-based training database construction method for the inverse synthetic aperture radar (ISAR) image and the performance was compared with the conventional range-Doppler (RD) ISAR algorithm. In simulations using five scatterer models derived from the CAD data of real aircraft, the performance of the conventional RD image derived using bandwidths ≥ 200 MHz was as high as that of the MUSIC algorithm. MUSIC was more efficient especially when the bandwidth was narrow.

Keywords: ISAR, Training Database, Scenario-Based Method.

1 Introduction

Inverse Synthetic Aperture Radar (ISAR) imaging is a technique to generate a high resolution two-dimensional (2D) image of a target [1-2]. It can be generated by synthesizing radar signals obtained from several observation angles. For the successful recognition of the ISAR image, the ISAR image must be matched to a corresponding image in a training database. However, constructing the training database composed of ISAR images similar to the test image is very difficult due to the 2D nature of the image. Mismatch between the training and the test images can yield serious performance degradation.

To solve this difficulty, we recently proposed an efficient method to construct the training database of the ISAR image of the aircraft [3]. The basic idea of the proposed method is to construct the training database using ISAR images derived from the assumed flight scenario of the target. The test ISAR image of an unknown target derived from the similar flight scenario were compared with the training image and high classification results were yielded. However, because the system performance is generally affected by the system bandwidth, it is required to study the effect of bandwidth reduction and the high resolution methods.

In this paper, the performance of the scenario-based construction of training database was studied in terms of the system bandwidth. The multiple signal classification (MUSIC) algorithm [4] was applied to further improve the performance and compared with the conventional range-Doppler algorithm (RDA). In simulations using five scatterer models derived from the CAD data of real aircraft, the performance of RDA derived using the wide bandwidth larger than or equal to

200MHz was better than that of MUSIC. MUSIC yielded higher classification results than those of RDA.

2 Principles of the ISAR Image and Proposed Method

2.1 RDA, MUSIC and Classifier

For the ISAR imaging, we adopted the conventional RDA. This method is composed of three steps: range compression, translational motion compensation (TMC) and azimuth compression [5]. Range compression yields high resolution range profiles (HRRPs) at each aspect angle by compressing wideband chirp radar signals. Because the target moves during the observation time, TMC composed of range alignment and phase adjustment compensates for the motion of the target. For range alignment, the method that minimizes the following 1D entropy cost function is used:

$$H_{G_m, G_{m+1}} = - \sum_0^{N-1} \bar{G}(\tau, n) \ln \bar{G}(\tau, n), \quad (1)$$

where $\bar{G}(\tau, n)$ is an average of two HRRPs with one shifted by τ . N is the total number of range bins. τ that minimizes the 1D entropy is found and used for the alignment of two RPs. For phase adjustment, the phase error vector to compensate for the phase error of each HRRP is searched for using the 2D entropy cost function defined by

$$Ent = \sum_{i=1}^M \sum_{j=1}^N \left| \bar{I}(i, j) \right|^2 \ln \left| \bar{I}(i, j) \right|^2, \quad (2)$$

where $\bar{I}(i, j)$ is the (i, j) component of the ISAR image normalized by the total power of it. M is the number of down-range bins and N is the number of cross-range bins. This method finds the phase error for each m that minimizes the 2D entropy. Then, azimuth compression locates each scatterer at its corresponding location using their Doppler frequencies (see [5] for the detailed procedure).

The MUSIC algorithm further improves the resolution of the ISAR image derived using RDA. This algorithm utilizes the fact that the eigenvector of the noise in the covariance matrix is orthogonal to the direction vector of the scatterer on the ISAR image. The noise matrix \mathbf{E}_n^H is constructed using the smallest b eigen values of the estimated covariance matrix and the ISAR image is given as

$$I_{MUSIC}(x, y) = \frac{\mathbf{e}(x, y)^H \mathbf{e}(x, y)}{\mathbf{e}(x, y)^H \mathbf{E}_n^H \mathbf{E}_n \mathbf{e}(x, y)}, \quad (5)$$

where $\mathbf{e}(x, y)$ is the direction vector at (x, y) position on the image and H is the complex conjugate transpose. (see [4] for the procedure).

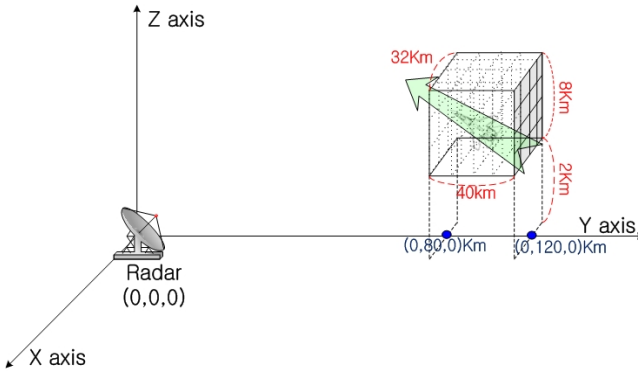


Fig. 1. Construction of scenario-based training database (45° flight angle $[-1 -1 0]$).

For the classification of ISAR images, we used the efficient polar mapping classifier [6].

2.2 Scenario-Based Method for the Training Database and the Test Image

The recently proposed scenario-based construction method constructs ISAR training database based on the assumed flight scenario of the aircraft [3]. We uniformly sampled the 3D space (training space) and assumed that a target moved at a given velocity in a given direction starting from each given grid point. Then, the derived ISAR image was stored in the training database. Three flight scenarios were used: $[x \ y \ z] = [-1 -1 0]$ (45° horizontal flight angle), $[0 -1 0]$ (flying directly at the radar), and $[-1 0 0]$ (90° horizontal flight angle) (Fig. 1). Very high classification results were obtained for the scenario at an angle to the radar viewing direction.

3 Simulation Results

Simulations were performed using five targets consisting of isotropic point scatterers extracted from the 3D CAD data of real B737, F18, F14, Su35, and Rafale jets (Fig. 2). For the radar system, we used a monostatic chirp radar with pulse repetition frequency = 2 kHz, center frequency = 9.15 GHz, sampling frequency = 512 MHz and $\tau = 30 \mu s$. Three flight scenarios were used for the classification as in [3] ($[x \ y \ z] = [-1 -1 0]$, $[0 -1 0]$, and $[-1 0 0]$).

For the $[-1 -1 0]$ and $[-1 0 0]$ flight trajectories, the training space was $80 \leq y \leq 120$ km, $-16 \leq x \leq 16$ km, and $2 \leq z \leq 10$ km. However, because the angular variation was very small for $[0 -1 0]$ flight, the training space was $6 \leq y \leq 10$ km, $-16 \leq x \leq 16$ km and $2 \leq z \leq 10$ km to obtain the required aspect angle variation. Then 125 training samples per target were stored by dividing the training space into $5 \times 5 \times 5 = 125$ grid points in each of the three axes ($125 \times 5 = 625$ training images). Because the flight trajectory was known in the training phase, TMC was not applied. Instead, the aircraft was rotated using the known azimuth and elevation angles.

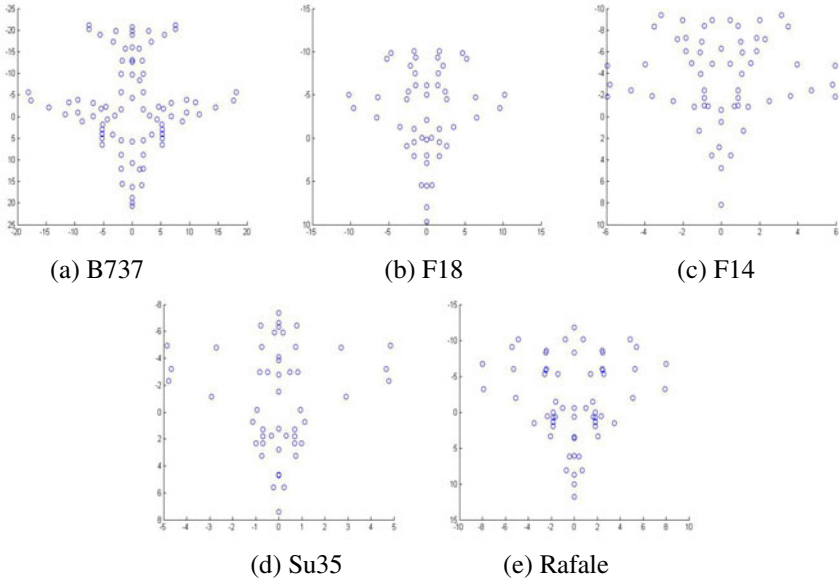


Fig. 2. Targets used for simulation

For the test image, 50 test images of each target were derived for each scenario. To make the down-range resolution Δr and cross-range resolution Δr_c equal, the aircraft flew until the aspect angle variation $\Delta\theta$ satisfied the relationship $\Delta r_c = \lambda_c/2\Delta\theta = \Delta r = c/2B$, where λ_c is the wavelength of the center frequency, c is the speed of the light and B is the bandwidth. In this procedure, TMC was applied because the motion of the target was unknown. Then, the derived image was enhanced using MUSIC algorithm and performance was compared for the $B_s = [50\ 100\ 150\ 200\ 250\ 300]$ MHz at the signal to noise ratio (SNR) equal to 10 dB. Using the relationship above, the down-range resolutions used were [3 1.5 1 0.75 0.6 0.5] m. Resolution of the ISAR image derived using 100MHz bandwidth at the SNR = 10 dB was considerably improved for the three scenarios (Fig. 3 and Fig. 4).

The scenario-based method was very sensitive to the bandwidth; the correct classification ratios (Pcs) significantly decreased in the three flight scenarios (Fig. 5);

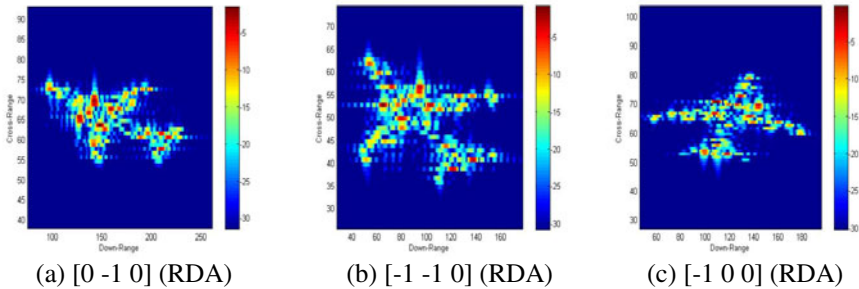


Fig. 3. ISAR images of B737 derived by RDA (SNR = 10 dB)

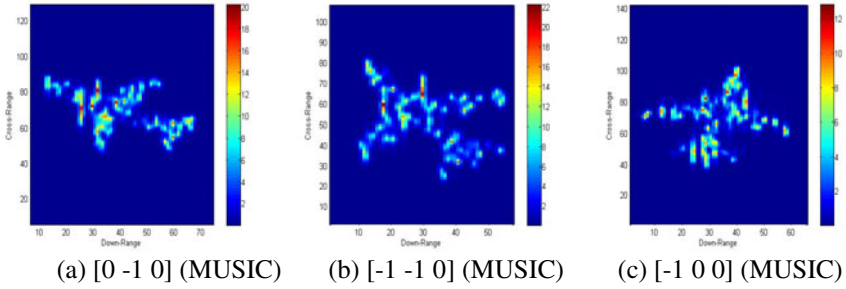


Fig. 4. ISAR images of B737 enhanced by MUSIC (SNR = 10 dB)

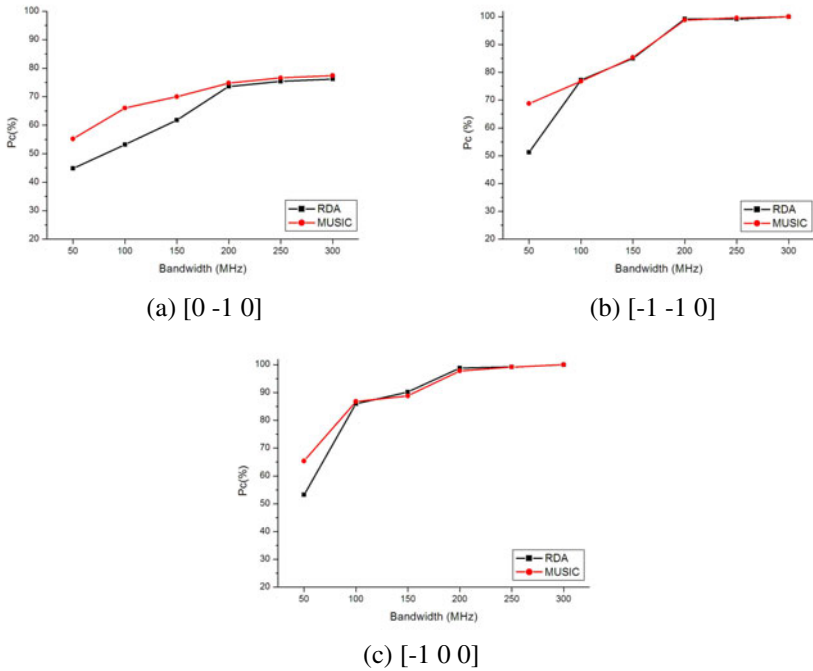


Fig. 5. Classification results for the three flight scenarios

P_c s reduced by approximately 32% ~ 49 % for RDA and 32 ~ 35 % for MUSIC algorithm. Because the ISAR image in [0 -1 0] flight was derived by the elevation angle variation, P_c s were low even at wide B_s . This proves that ISAR image is not adequate for the target recognition when the target directly closes toward the radar. Because of the low P_c s, the performance reduction in this scenario was smaller than in the other two scenarios. MUSIC provided higher P_c s for all B_s , however, at $B_s \geq 200$ MHz, the performance improvement was negligible (Fig. 5a). Considering generally much time is consumed for MUSIC, RDA is more beneficial than using MUSIC at these wide B_s . At narrow $B_s \leq 200$ MHz, MUSIC should be applied to raise P_c s.

$[-1 \ -1 \ 0]$ and $[-1 \ 0 \ 0]$ directions, the performance degradation was much larger than that in $[0 \ -1 \ 0]$ flight because P_{cs} were very close to 100 % (Fig. 5b and c). As in the $[0 \ -1 \ 0]$ flight, performance of RDA and MUSIC was approximately equal for the $B_s \geq 200$ MHz. At $B = 50$ MHz, P_{cs} decreased significantly with the amount of degradation in RDA much larger than in MUSIC. Therefore, MUSIC should be applied to increase P_c at this bandwidth.

4 Conclusion

In this paper, we applied the MUSIC algorithm to the efficient scenario-based construction method for the training database and its effect was analyzed compared with the conventional RDA algorithm. For the three flight scenarios, the performance of the two methods degraded considerably depending on B_s and the amount of degradation of MUSIC was smaller than that of RDA. At wide $B_s \geq 200$ MHz, the performance of MUSIC didn't improve considerably. Considering much computation time for MUSIC and the development of the modern radar system, we can conclude that using RDA is much more beneficial than using MUSIC in the real-time military operation.

Acknowledgments. This work is the result of the "Human Resource Development Center for Economic Region Leading Industry" Project, supported by the Ministry of Education, Science & Technology(MEST) and the National Research Foundation of Korea(NRF).

References

1. Chen, C.C., Andrews, H.C.: Target-motion-induced radar imaging. *IEEE Trans. Aerosp. Electron. Syst.* 16(1), 2–14 (1980)
2. Sang-Hong, P., Kang-Kook, P., Joo-Ho, J., Hyo-Tae, K., Kyung-Tae, K.: ISAR imaging of multiple targets using edge detection and Hough transform. *J. Electromagn. Wav. Appl.* 22(2-3), 365–373 (2008)
3. Sang-Hong, P., Moon-Gab, J., Kyung-Tae, K.: Construction of ISAR Training Database for Automatic Target Recognition. *J. Electromagn. Wav. Appl.* 25(11-12), 1493–1503 (2011)
4. Odendaal, J.W., Barnard, E., Pistorius, C.W.I.: Two-Dimensional Superresolution Radar Imaging Using the MUSIC Algorithm. *IEEE T. Antennas Propagat.* 42(10), 1386–1391 (1994)
5. Li, X., Liu, G., Ni, J.: Autofocusing of ISAR images based on entropy minimization. *IEEE Trans. Aerosp. Electron. Syst.* 35(4), 1240–1251 (1999)
6. Kyung-Tae, K., Dong-Kyu, S., Hyo-Tae, K.: Efficient classification of ISAR images. *IEEE T. Antennas Propagat.* 53(5), 1611–1621 (2005)

The LFT Formulation of Dynamical Systems with the Uncertainties of Parameter Variations

Jieh-Shian Young

Institute of Vehicle Engineering, National Chang-Hua University of Education
#1, Jin De Road, Chang-hua 500, Taiwan
jsyoung@cc.ncue.edu.tw

Abstract. This paper mainly studies the problems of the formulation of dynamical systems with the parameter uncertainties. The uncertainties in the inertial term of a dynamical system are studied and dealt with since they are necessary but troublesome in applications. These uncertainties can be treated with an inverse transfer function in a block diagram. These problems can be formulated as the general linear fractional transformation which is regarded as the standard problem. The proposed technology can be easily extended to other standard control problems such as the robust 2-parameter compensator problems. An example of an inverse pendulum is also presented by this proposed technology.

Keywords: Dynamical system, linear fractional transformation, parameter variation, robustness, uncertainty.

1 Introduction

Most systems in industry have severe nonlinearity and uncertainty, which may usually make the work of the synthesis harder or even infeasible. The robust problem with the H^∞ -bounded uncertainty is a kind of the standard control problems in H^∞ control synthesis (see [1,2] and the references therein). The H^∞ approach has provided some promising results for the robust stabilization of the plant with unstructured uncertainty. The plant uncertainty could be modeled as a transfer function from the nominal plant such as the multiplicative uncertainty, the additive uncertainty, the coprime factor uncertainty, the parametric state-space uncertainty, etc. The robust stabilization under unstructured perturbations was fulfilled by Glover [3] and Verma et al. [4], etc. The robustness for plants with coprime factor uncertainty has been studied in [5-7]. Robust stability analysis for control systems with parameter uncertainties is one of the significant issues in systems theory and in applications. A number of tests are available to evaluate robustness and performance such as the Kharitonov theorem and related studies [8,9], quadratic stability tests [10], μ and K_m analysis [11], integral-quadratic-constraint (IQC) techniques [12,13], linear matrix inequality (LMI) [14], etc. A comprehensive achievement from the quadratic stability has been presented by LMI approach in [15]. The affine parameter-dependent Lyapunov functions which depend on the real uncertain parameters are used to make

the robust stability criteria [16]. They are less conservative than quadratic stability. A sufficient condition for existence of them has been presented in terms of LMIs [17]. Yang and Lum have compared the results from the robust stability criterions mentioned above for linear systems with affine parameter uncertainties [18].

In the physical system, the inverses of inertial terms will lose the authenticity the as analysis or synthesis of systems. For example, a system is governed by the equation below.

$$\begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix} \dot{x} = \begin{pmatrix} 0 & 1 \\ -K & -B \end{pmatrix} \dot{x} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} F, \tag{1}$$

where M is ranged from 0.125 to 0.625, or $M = 0.375 + \Delta$ with $\|\Delta\| \leq 0.25$. However, in case that M is invertible, (1) becomes

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ -M_i K & -M_i B \end{pmatrix} \dot{x} + \begin{pmatrix} 0 \\ M_i \end{pmatrix} F, \tag{2}$$

where $M_i \equiv M^{-1}$ is range from 1.6 to 8, or $M_i = 4.8 + \Delta_i$ with $\|\Delta_i\| \leq 3.2$. (1) and (2) are the same actually but with different uncertainty bounds. In case M is a constant probability distribution, M_i will not be the same probability distribution. In addition, the analysis or synthesis will become more complicated if K or B varies. This paper is concerned with the problems of the formulation of the dynamical systems with the parameter uncertainties especially for those of the inertial parameters in system equations of motion. The plant uncertainty formulated by the linear fractional transformation (LFT). The uncertainties in the inertial terms of physical systems have been discussed original in this paper. An example of the inverse pendulum will be utilized to demonstrate the proposed approach.

The notations in this paper are as follows: $x = [a, b]$ denotes that x is ranged from a to b and x can be any number within this range. Δ_x is the uncertainty varying form

$$X_0 \cdot G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] \text{ means } G(s) = C(sI - A)^{-1} B + D.$$

2 Mathematical Preliminaries

A general LFT of a control problem can be shown as Fig. 1, where v_c is the exogenous input vector, or the command input vector; u is the control signal input vector; v_Δ is the pseudo uncertainty input vector; z_c is the controlled output vector; y is the measured output vector; and z_Δ is the pseudo uncertainty output vector. In

addition, $\begin{pmatrix} z_\Delta \\ z_c \\ y \end{pmatrix} = P \begin{pmatrix} v_\Delta \\ v_c \\ u \end{pmatrix}$, where $P = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}$.

Lemma 1. The transfer function of a control problem with the general LFT from v_c to z_c in Fig. 1 can be

$$z_c = \{ [P_{21} + P_{23}K(I - P_{33}K)^{-1}P_{31}]\Delta(I - [P_{11} + P_{13}K(I - P_{33}K)^{-1}P_{31}]\Delta)^{-1} \cdot [P_{12} + P_{13}K(I - P_{33}K)^{-1}P_{32}] + [P_{22} + P_{23}K(I - P_{33}K)^{-1}P_{32}] \} v_c.$$

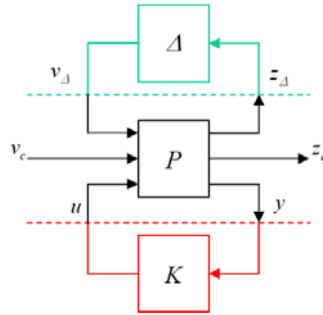


Fig. 1. General LFT's of control problems

The proof of Lemma 1 is straightforward. In case that $\Delta = 0$, the problem is degraded to the so-called standard control problem. It could be the model matching problems, tracking problems, and so on. The objective of the standard problem is to find an appropriate controller K to stabilize the plant P internally as well as to satisfy the specified requirements of the overall system performances. The general LFT of a control problem may become a typical robust control problem when $v_c = 0$ without the specified controlled output. By the small gain theorem, $\| [P_{11} + P_{13}K(I - P_{33}K)^{-1}P_{31}]\Delta \|_\infty$ must be smaller than 1 such that the system maintains stable as $\Delta(s)$ varies in stable region. The objective of the typical robust control problem is to find a robust controller K which stabilizes the plant P internally with $\Delta(s)$.

Let a typical dynamical system be $M\ddot{x} + B\dot{x} + Kx = F$, where \underline{x} is the variable vector of the displacement term, M is the parameter matrix of the inertia term, B is the parameter matrix of the damper term, K is the parameter matrix of the spring term, and F is the force term. The state vector is defined as $x = (x_1 \ x_2)^T \equiv (\underline{x} \ \dot{\underline{x}})^T$, the control signal input vector is $u=F$, the command input vector is denoted as v_c , the controlled output vector is $z_c \equiv C_c x + D_z v_c + D_c F$, and the measured output vector is $y \equiv Cx + D_y v_c + DF$. The equations of the dynamical system could be rewritten in the form of the state and output equations as follows.

$$\begin{cases} J\dot{x} = Ax & + Bu \\ y = Cx + D_y v_c + Du \\ z_c = C_c x + D_z v_c + D_c u, \end{cases} \tag{3}$$

where $J = \begin{pmatrix} I & 0 \\ 0 & M \end{pmatrix}$, $A = \begin{pmatrix} 0 & I \\ -K & -B \end{pmatrix}$, and $B = \begin{pmatrix} 0 \\ I \end{pmatrix}$. In general, (3) can be

$$\begin{cases} (J_0 + \Delta_J)\dot{x} = (A_0 + \Delta_A)x + (B_0 + \Delta_B)u \\ y = (C_0 + \Delta_C)x + (D_0 + \Delta_D)u + D_y v_c \\ z_c = C_c x + D_c u + D_z v_c, \end{cases} \tag{4}$$

where $J_0, A_0, B_0, C_0, D_0, \Delta_J, \Delta_A, \Delta_B, \Delta_C,$ and Δ_D are the nominal values and the uncertainties of matrices $J, A, B, C,$ and $D,$ respectively. These uncertainties could be obtained from the physical values directly. For instance, if (4) is a 2nd order system, M varies from 4 to 6. Then $J_0 = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$ and $\Delta_J = \begin{pmatrix} 0 & 0 \\ 0 & [-1, 1] \end{pmatrix}$.

3 Main Result

(4) is the dynamical equation with the parameter uncertainties of a physical system in general. Some modern control techniques can be applied to a general LFT of a control problem. The objective is how to construct a general LFT of a control problem from a dynamical equation, or a physical system.

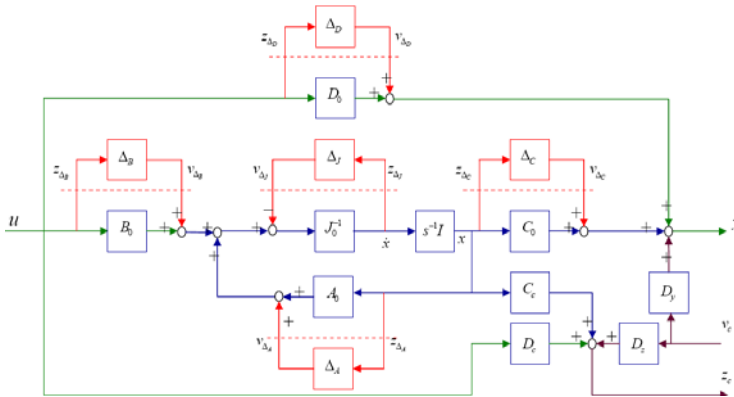


Fig. 2. The block diagram of the dynamical equation with uncertainties

Theorem 1. Let (4) be the dynamical equation of the system. The parameter uncertainty matrix can be defined as $\Delta \equiv diag\{\Delta_J \ \Delta_A \ \Delta_B \ \Delta_C \ \Delta_D\}$. $v_\Delta = (v_{\Delta_J}^T \ v_{\Delta_A}^T \ v_{\Delta_B}^T \ v_{\Delta_C}^T \ v_{\Delta_D}^T)^T$ and $z_\Delta = (z_{\Delta_J}^T \ z_{\Delta_A}^T \ z_{\Delta_B}^T \ z_{\Delta_C}^T \ z_{\Delta_D}^T)^T$ are the uncertainty pseudo input vector and the uncertainty pseudo output vector, respectively. Then the state space description of a general LFT of the control problem in (4) will be

$$\begin{pmatrix} z_\Delta \\ z_c \\ y \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} \begin{pmatrix} v_\Delta \\ v_c \\ u \end{pmatrix}, \tag{5}$$

where

$$\begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} = \left[\begin{array}{cccccc|ccc} J_0^{-1}A_0 & -J_0^{-1} & J_0^{-1} & J_0^{-1} & 0 & 0 & 0 & J_0^{-1}B_0 \\ J_0^{-1}A_0 & -J_0^{-1} & J_0^{-1} & J_0^{-1} & 0 & 0 & 0 & J_0^{-1}B_0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \\ \hline C_c & 0 & 0 & 0 & 0 & 0 & D_z & D_c \\ C_0 & 0 & 0 & 0 & I & I & D_y & D_0 \end{array} \right].$$

Proof: The block diagram of (4) is shown in Fig. 2. Remove the blocks of Δ -terms. v_Δ and z_Δ are regarded as a pseudo input vector and a pseudo output vector, respectively. From Fig. 2, the state equation is

$$\dot{x} = J_0^{-1}A_0x - J_0^{-1}v_{\Delta_j} + J_0^{-1}v_{\Delta_A} + J_0^{-1}v_{\Delta_B} + J_0^{-1}B_0u,$$

and $y = C_0x + v_{\Delta_c} + v_{\Delta_D} + D_yv_c + D_0u$, i.e.,

$$y = \begin{pmatrix} P_{31} & P_{32} & P_{33} \end{pmatrix} \begin{pmatrix} v_\Delta \\ v_c \\ u \end{pmatrix} = \left[\begin{array}{cccccc|ccc} J_0^{-1}A_0 & -J_0^{-1} & J_0^{-1} & J_0^{-1} & 0 & 0 & 0 & J_0^{-1}B_0 \\ C_0 & 0 & 0 & 0 & I & I & D_y & D_0 \end{array} \right] \begin{pmatrix} v_\Delta \\ v_c \\ u \end{pmatrix}.$$

It is intuitive to construct the state space descriptions of z_Δ and z_c from $(v_\Delta^T \ v_c^T \ u^T)^T$, i.e.,

$$z_\Delta = \begin{pmatrix} P_{11} & P_{12} & P_{13} \end{pmatrix} \begin{pmatrix} v_\Delta \\ v_c \\ u \end{pmatrix} = \left[\begin{array}{cccccc|ccc} J_0^{-1}A & -J_0^{-1} & J_0^{-1} & J_0^{-1} & 0 & 0 & 0 & J_0^{-1}B_0 \\ J_0^{-1}A & -J_0^{-1} & J_0^{-1} & J_0^{-1} & 0 & 0 & 0 & J_0^{-1}B_0 \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \end{array} \right],$$

and

$$z_c = \begin{pmatrix} P_{21} & P_{22} & P_{23} \end{pmatrix} \begin{pmatrix} v_\Delta \\ v_c \\ u \end{pmatrix} = \left[\begin{array}{cccccc|ccc} J_0^{-1}A_0 & -J_0^{-1} & J_0^{-1} & J_0^{-1} & 0 & 0 & 0 & J_0^{-1}B_0 \\ C_c & 0 & 0 & 0 & 0 & 0 & D_z & D_c \end{array} \right].$$

4 An Illustrated Example

An inverted pendulum mounted on a motor-driven cart is shown in Fig. 3. θ and \underline{x} are defined as the angle of the rod from the vertical line and the displacement of the cart, respectively. v and ω are the velocity terms, or the derivatives, of \underline{x} and θ , respectively. Let $x = (\underline{x} \ v \ \theta \ \omega)^T$. Assume that \underline{x} and θ are the measured output, i.e., the measured output vector is defined as $y = (x \ \theta)^T$. The control signal input is u . Let the controlled output vector be $z_c = (\underline{x}_d - \underline{x} \ \theta_d - \theta)^T$, and the command input vector be $v_c = (\underline{x}_d \ \theta_d)^T$, where \underline{x}_d and θ_d are the desired values of \underline{x} and θ , respectively. θ_d is equal to 0 in most cases. Assume $M = [1.9, 2.1]$, $m = [0.095, 0.105]$, and $l = [0.99, 1.01]$. The dynamical equation is as follows.

$$\begin{cases} (J_0 + \Delta_J)\dot{x} = Ax + Bu, \\ y = Cx, \\ z_c = -Cx + v_c, \end{cases}$$

where

$$J_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 2.1 & 0 & 0.10005 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 9.8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ v \\ \theta \\ \omega \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

and

$$\Delta_J = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & [-0.01, 0.01] \\ 0 & 0 & 0 & 0 \\ 0 & [-0.105, 0.105] & 0 & [-0.006, 0.006] \end{pmatrix}.$$

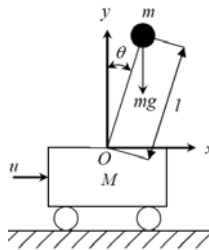


Fig. 3. The inversed pendulum system

Let the uncertainty pseudo input vector and output vector be v_Δ and z_Δ , respectively. The parameter uncertainty is defined as $\Delta = \Delta_J$. From Theorem 1, the state space description of the general LFT of the control problem for this inversed pendulum is $(z_\Delta^T \ z_c^T \ y^T)^T = P(v_\Delta^T \ v_c^T \ u^T)^T$, where

$$\begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} = \left[\begin{array}{c|c|c|c} J_0^{-1}A & -J_0^{-1} & 0 & J_0^{-1}B \\ \hline J_0^{-1}A & -J_0^{-1} & 0 & J_0^{-1}B \\ \hline -C & 0 & I & 0 \\ \hline C & 0 & 0 & 0 \end{array} \right] \cong \left[\begin{array}{cccc|cccc|cccc} 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.4903 & 0 & 0 & 0 & 0 & 0.0500 & 0 & -0.5000 & 0 & 0 & 0 & 0.5000 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 10.2903 & 0 & 0 & 0 & 0 & -1.0500 & 0 & 0.5000 & 0 & 0 & 0 & -0.5000 \\ \hline 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.4903 & 0 & 0 & 0 & 0 & 0.0500 & 0 & -0.5000 & 0 & 0 & 0 & 0.5000 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 10.2903 & 0 & 0 & 0 & 0 & -1.0500 & 0 & 0.5000 & 0 & 0 & 0 & -0.5000 \\ \hline -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

In this example, the parameter uncertainties of the inertial term are highlighted since to characterize them is tough indeed in control synthesis. However, the parameters of the inertial term usually vary. The state space description of the general LFT representation of a control problem proposed in this paper can achieve not only this kind of uncertainties but the uncertainties of other terms in dynamical equation.

5 Conclusions

In this paper, the general LFT representations of control problems with all possible parameter uncertainties are presented. The uncertainties in the inertial term of a dynamical system are original discussed and formulated in general LFT representations. The typical 2-parameter compensated system is an extension of the proposed technology and can be easily fulfilled. Other standard problems can follow the same procedure. A mechanical example of an inverted pendulum is also illustrated. This proposed technique will facilitate the analysis and the synthesis of the robustness stability and performance by some developed technologies such as H^∞ robust control, μ -synthesis, mixed-norm H^2 / H^∞ control, etc. Furthermore, it can be extended to the formulations of the affine parameter-dependent uncertainty.

Acknowledgments. This work was supported in part by the National Science Council under Grand NSC 100-2221-E-018-017 and NSC 100-2623-E-018-003-D.

References

1. Francis, B.A.: A course in H^∞ control theory. Springer, New York (1987)
2. Francis, B.A., Doyle, J.C.: Linear control theory with an H^∞ optimality criterion. SIAM Journal of Control and Optimization 25, 815–844 (1987)
3. Glover, K.: Robust stabilization of linear multivariable systems: relations to approximation. International Journal of Control 43, 741–766 (1986)
4. Verma, M., Helton, J., Jonckheere, E.: Robust stabilization of a family of plants with varying numbers of right half-plant poles. In: Proceedings of American Control Conference, Seattle, USA, pp. 1827–1832 (1986)

5. Glover, K., McFarlane, D.: Robust stabilization of normalized coprime factor plant descriptions with H_1 -bounded uncertainty. *IEEE Transactions on Automatic Control* 34, 821–830 (1989)
6. Young, J.-S.: Synthesis of Decoupling Controller for Non-minimum Phase Plants of Different Pole Numbers on RHP within Uncertainties. *International Journal of Systems Science* 42(6), 939–950 (2011)
7. Young, J.-S.: Robust compensator synthesis for antiwindup design with coprime factor uncertainties. *International Journal of Innovative Computing, Information and Control* (to be appeared 2012)
8. Barmish, B.R.: A generalization of Kharitonov's four polynomial concept for robustness stability problem with linearly dependent coefficient perturbations. *IEEE Transactions on Automatic Control* 34, 157–165 (1989)
9. Kharitonov, V.L.: Asymptotic stability of an equilibrium position of a family of systems of linear differential equations. *Diferentsial'nie Uravnenia* 14, 2086–2088 (1978)
10. Khargonekar, P.P., Petersen, I.R., Zhou, K.: Robust stabilization of uncertain linear systems: Quadratic stability and H_∞ control theory. *IEEE Transactions on Automatic Control* 35, 356–361 (1990)
11. Fan, M.K.H., Tits, A.L., Doyle, J.C.: Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics. *IEEE Transactions on Automatic Control* 36, 25–38 (1991)
12. Rantzer, A., Megretsky, A.: System analysis via integral quadratic constraints. In: *Proceedings of the 33rd IEEE Conference on Decision and Control*, pp. 3062–3067 (1994)
13. Jonsson, U., Rantzer, A.: Systems with uncertain parameters: Time variations with bounded derivatives. In: *Proceedings of the 33rd IEEE Conference on Decision and Control*, pp. 3074–3079 (1994)
14. Gahinet, P., Apkarian, P.: A linear matrix inequality approach to H_∞ control. *International Journal of Robust Nonlinear Control* 4, 421–448 (1994)
15. Boyd, S., El Ghaoui, L., Feron, E., Balakrishnan, V.: *Linear matrix inequalities in systems and control theory*. SIAM, Philadelphia (1994)
16. Gahinet, P., Apkarian, P., Chilali, M.: Affine parameter-dependent Lyapunov functions and real parametric uncertainty. *IEEE Transactions on Automatic Control* 41, 436–442 (1996)
17. Feron, E., Apkarian, P., Gahinet, P.: Analysis and synthesis of robust control systems via parameter-dependent Lyapunov functions. *IEEE Transactions on Automatic Control* 41, 1041–1046 (1996)
18. Yang, G.H., Lum, K.Y.: Comparisons among robust stability criteria for linear systems with affine parameter uncertainties. *Automatica* 43, 491–498 (2007)

The Design and Implementation of Wireless Network System for Temperature Control Based on ARM

Tianhao Hu^{1,*}, Kunchang Chen¹, Xinli Wu¹, Dichong Wu²,
Canlin Mo¹, and Renwang Li¹

¹ Advanced Manufacturing Institute, Zhejiang Sci-Tech University, Hangzhou 310018

² Zhejiang University of Finance & Economics, Hangzhou 310018

hu_tianhao@foxmail.com

Abstract. Regarding to the problem that existing temperature control system with low precision, bad adaptability and complex operation, this paper proposed one kind of wireless network temperature control system based on the ARM(Advanced RISC Machines) platform. This system adopted embedded Linux2.6 core RedHat operating system and 32bits Samsung-S3C2440 processor, nRF903 single-chip RF transceiver and DS18B20 digital temperature sensor, etc. to group the wireless temperature control device, and used the bias in proportion, integral and differential data processing of PID to control algorithm, finally reach the temperature of remote control. The test results showed that the system has the advantages of high sensitivity, good stability, easy to maintain and manage, has a very good application prospect.

Keywords: ARM, Linux2.6, nRF903, wireless network, temperature control.

1 Introduction

In the field of agriculture, industry and daily life, temperature is one of the most commonly used parameter. However, in the temperature precision, controllability, remote control and other high performance requirements, temperature control is not a simple thing, and so it become a difficult problem for a lot of enterprises in the production process. Therefore, a simple, high control precision, strong adaptability and high stability of the temperature control system emerge as the times require.

This paper designs a wireless network temperature control system based on embedded platform, Mainly are applied in the field of remote control, which is the main control mechanism and a control mechanism from the phase separation occasions, The main control mechanism through and from control mechanism for wireless data exchange, real-time display the environmental temperature and control from the control mechanism model for temperature regulation, with abnormal temperature alarm, data record, data query graph display and other functions.

* In the Master of Zhejiang Sci-Tech University, Mainly engaged in intelligent control technology.

2 Overall Design of The System

This paper designed is a set of field data collecting, data receiving and sending, data processing, data storage ,display and control, this five functions in one remote wireless network control system, According to the main technical parameters and functional requirements of the system, among the system of control terminal, use the embedded ABMA 2 ARM-S3C2440 bus controller as the main control chip, among the system of executive terminal, use the serial communications capabilities of 80C51 processing chip which developed by ATMEL company, and attaching a plurality of peripheral circuit module, Including reset module (Reset), memory expansion module, control / display module, wireless communication module , temperature acquisition module and the actuator module and so on. When the system runs, temperature acquisition module is responsible for real-time collection of external temperature information, and convert the analog signal into a corresponding digital signal, Execute terminal receives the data information which temperature acquisition module transmitted, after the simple treatment, through wireless transmitting module to the control terminal, then the control terminal receives and sends the control instructions to awaken the actuator action; control terminal for receiving data and data disposal and sends control instructions, including the display of whole system real time information, stored data, providing an alarm signal and the exchange interface of man-machine, which is convenient for on-site staff manual intervention. The overall structure of the system is as shown in Figure1.

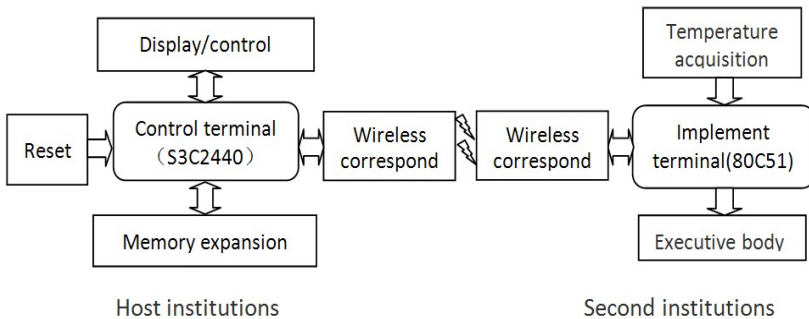


Fig. 1. The overall structure of the system

3 The Design of Key Module

3.1 Control Terminal

ARM microprocessor, because of its small size, low power consumption, high performance advantages, have occupied in industrial control, consumer electronics, communication system and other types of product markets[1], with the related data rich and easy to obtain, and the performance also can meet the system requirements of various parameters, so this design block adopts Samsung-S3C2440 as main control

chip, which integrate with the 32 bit microcontroller of ARM ARM920T core , support for Linux2.6, Windows CE and other operating system, the frequency can up to 533MHZ. In addition, the chip also incorporates the following modules: 16Kb instruction Cache, and data Cache,memory management unit (MMU), external memory controller (SDRAM control and chip select logic), LCD controller and so on. In the core plate, integrated with 64Mb SDRAM, 256Mb Nand Flash, 2Mb Nor Flash; While the expansion circuit board provide with the design of the full IO port resources: 1 4 wire resistive touch screen interface, one 100Mb Ethernet RJ-45 interface, the three serial (UART0, 1, 2),one SD expansion card interface, one USB interface and one 34 pin 2.0mmGPIO interface. The expansion board and peripheral interface structure as shown in Figure.2[2]:

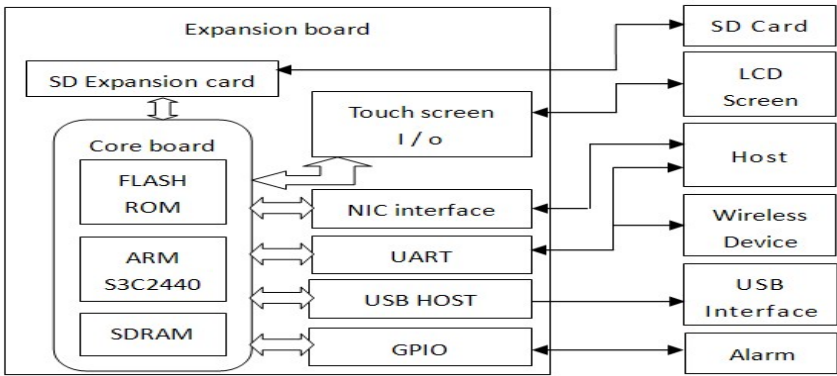


Fig. 2. The expansion board and peripheral interface

In picture2, the touch screen input / output interface are used to connect LCD touch screen, as the system man-machine interface, real-time display system of spot data information (including temperature values and curves), and provide relevant input options which control system operation ; SD expansion card interface is used to connect the attached storage devices, increase the system's storage space (the actual application on different occasions storage space may be insufficient) ; Serial UART0 and host connect to output debugging information, UART1 used to connect wireless communications equipment for data input and output ; USB interface can achieve data download; GPIO interface is used to output alarm signal.

The process access the serial port driver through the API (Automatic Program Interrupt) during the normal operation time, thereby according to the predetermined frequency to receive serial data, the system serial select the of D type 9 pin RS232 interface which provides 9 asynchronous communication signal, data transmission rate of 19200bps, meet the data transmission rate requirements. Then according to the temperature control algorithm, completes the data processing. The module designed as, data processing of temperature continues to use the traditional temperature control algorithm: PID control algorithm, namely according to the deviation ratio (Kp), integral (KI) and differential (KD) control method, controller formula[3] is as follows:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{d e(t)}{d t}$$

$$u(t) = K_p [e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{d e(t)}{d t}] \quad \left(\begin{array}{l} T_i = \frac{K_p}{K_i} \\ T_d = \frac{K_d}{K_p} \end{array} \right)$$

Type Kp is proportional amplification factor; KI is integral coefficient; KD is for differential coefficient. PID control based on feedback theory among the automatic control technology, mainly including three parts: measurement, comparison and Implementation (measured variables related information, and compared against expected values, with error correction control system response), the principle of the algorithm is shown as Figure.3 shown[4]:

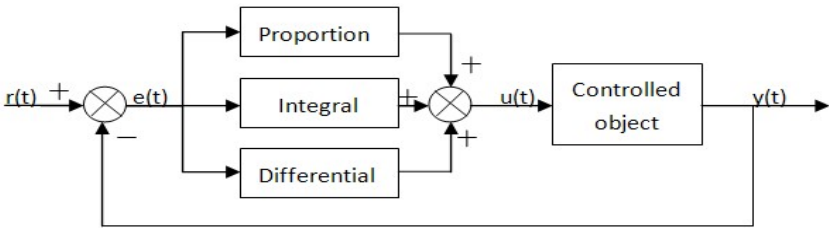


Fig. 3. Schematic diagram of PID algorithm

The design idea of software for controller terminal is: the selection of core board operating system is based on Linux2.6 kernel RedHat operating system, which has a fast response and kernel interaction advantages, and the kernel source code is open source code. Considering the operating system kernel function module is independent with the actual application, thus causing resource waste problem, During the process of design, the operating system kernel tailoring transplanted into the hardware platform after adjustment; To ensure the integrity of the operating system, reliability and meet the functional requirements of the system ,on this basis , delete kernel which included MIPS and PowerPC, multiple processor structure code, remove IPX and AppleTalk and other network protocols and delete redundant block device driver, retain the architecture of ARM source code, TCP / IP network protocol code and the necessary block device driver, but also add the wireless communication module of the driver, so that the kernel code system structure is more compact, more streamlined.

System application includes interactive interface, data transmission and data processing procedures, use the crossed compile tool, procedures will be transplanted to the target machine to run on the running process of the system software debug information from the serial output.

3.2 The Wireless Correspond Module

The wireless communication module is the important composition part of the temperature control system, this part select single chip RF transmitter based on the nRF903 chip, with wireless receiving, wireless transmission, PLL synthesis and other functions, the working

frequency of generic ISM band : 433 / 868 / 915MHz; and have the following properties : 1) using FSK / GMSK modulation and demodulation technology, strong anti-interference ability ; 2) with advanced frequency synthesis, frequency stability, the effective bandwidth in the 155.6KHz maximum data transmission rate up to 76.8Kbps; 3) high sensitivity, long transmission distance and transmission distance can be up to 300 ~ 500m;

nRF903's working mode is divided into receive mode, transmission mode, standby mode and a power down mode. When keeping the STBY and PWR_DWN pin low, by controlling the level of TXEN pin, can be achieved nRF903 in the receive mode and sent to switch between modes, the switching time is about 1.5ms. The state control table as follows[5]:

Table 1. nRF903 state control table

No.	Work Mode	STBY	PWR_DWN	TXEN
1	Receive	0	0	0
2	Transmit	0	0	1
3	Standby	0	1	X
4	Brown	1	0	X

According to the nRF903 working mode to write the corresponding driver under Linux system, to control the received and transmit signals. The driver consists of the following functions: 1) in the register_nRF function complete the registration. 2) in the init_nRF function initializes the device working mode and other related parameters. 3) in the send_rec_nRF function to realize data transmission and reception. 4) in the test_nRF function to detect the current data transmission or reception state. 5) in the change_mode_nRF function of the equipment in the working state conversion. Communication subprogram flow as shown below :

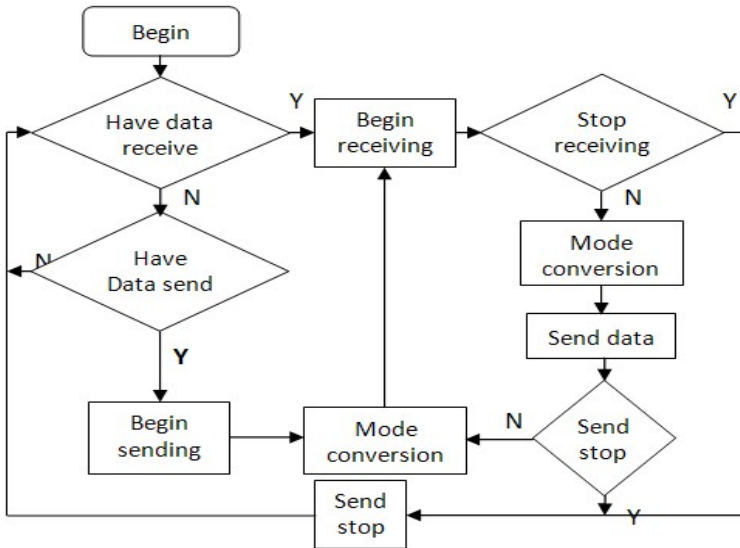


Fig. 4. Communication subprogram flow

3.3 The Implement Terminal

The implement terminal using 80C51 which is produced by ATMEL company based on the MCS51 core 8 processor chip, chip storage space for 4K byte. The main function is responsible for the field temperature data acquisition and transmission of temperature data to the control terminal, and receives from the control terminal of the control signal and to take appropriate action.

MCU interface resource allocation as follows: P0 as a data transmission interface and a wireless communication module, P1 is set to the signal input interface is connected with a temperature acquisition device, P2 is set to the signal output port is connected with the actuator module, Hardware configuration as shown in figure 5:

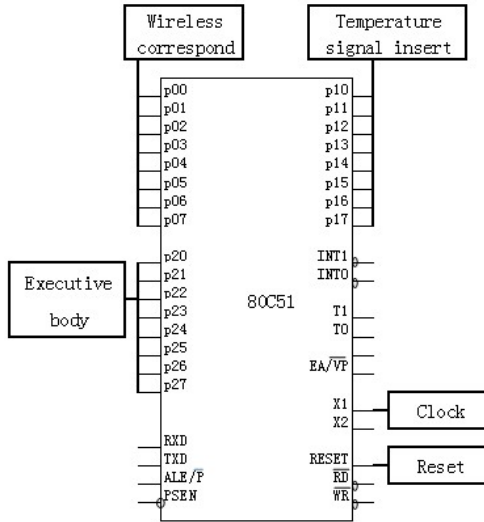


Fig. 5. The executive terminal structure

4 Conclusion

This paper design and implement a kind of remote wireless network temperature control system based on ARM platform, together with the industrial field data acquisition, data processing and data transmission and other functions into one, while use the PID fuzzy control idea, to overcome the ordinary temperature control system which has short control distance, complex operation, controllable low problem. Considering the versatility of the system, part of the module is according to the actual site of the different environment and use different equipment. Through testing, the system has stable performance, and has good application prospect.

Acknowledgments. The work is supported by: Zhejiang Province Natural Science Foundation (Grant No. R6080403, No. Z6090572, No. Y6110568), Zhejiang Province

Key Science and Technology Innovation Team (Grant No. 2011R09015), Zhejiang Province Key Science and Technology Innovation Team (Grant No. 1102326-E: Modular CNC design and selection).

References

1. Liu, Y., Shen, X., Wang, Y.: ARM based design and Realization of intelligent monitoring system. *Journal of Jilin University (Information Science Edition)* 29(2), 158–163 (2011)
2. Liu, S., Mu, C., Zhao, M.: The embedded system based on ARM humanoid robot controller design. *Tsinghua Univ. (Sci. & Tech.)* 48(4), 482–485 (2008)
3. Zhang, Q.: The design of missile control system based on Fuzzy-PID. *Journal of Projectiles, Rocket, Missiles and Guidance* 23(3), 10–12 (2003)
4. Xing, R., Solemn, Liang, G.: Based on PID control illumination automatic control system. *Computer & Digital Engineering* 38(5), 70–73 (2010)
5. Qin, J.: Wireless communication chip nRF903 and 89C51 interface design. *Electronic Engineer*. 30(9), 52–54 (2004)

Research of Interactive Product Design for Virtual Tourism

Zhongyan Hu^{*}, Zaihui Cao, and Jinfa Shi

Department of Art and Design, Zhengzhou Institute of Aeronautical Industry Management
450015 Zhengzhou, China
huzohngyan@zzia.edu.cn, czhhn@126.com

Abstract. As a supplementary for traditional tourism, virtual tourism is more and more welcomed by netizens. Virtual tourism created an economical, unlimited and safe tourism environment. But as a way that combines the physical building and the mental relaxing together, traditional tourism is still irreplaceable. The virtual tourism interactive products research mainly spots at the new interactive interfaces, which could benefit from the advantages of both virtual tourism and traditional tourism. So with the new kind interfaces, we can enjoy physical exercises and virtual tourism at the same time.

Keywords: virtual tourism, interactive product, interface, physical exercise.

1 Introduction

Since the 1990s, Williams-Hobson (1995) to form the first theme park will virtual reality technology into tourism, many scholars after by virtual technology, network technology, GIS (geographic information system), 3 d visualization and multimedia technology, the virtual tour influence factors and the future application research [1]. In 2005, Google Earth, the launch of a virtual tourism development based on the Internet boom, accelerate the virtual tour development situation [2].

2 Virtual Tour Development Present Situation

Entering the new century, a virtual tour of the research at home and abroad has gradually developed into the application stage. Luis A. Hernandez (2001), Sandau, R. (2003), Guo Xi Rong (2009), Pingan Feng (2009) and from 360-degree panoramic video systems, aerial photography, shooting, space technology, three-dimensional three-channel surround screen projection technology and other aspects to explore specific ways to obtain real space [3-6]. Ali A. POUYAN (2007), Craig Saper (2008), Christoph Grün (2008) use the network platform to showcase tourism and other content for the exchange interaction, ecological protection of the environment [6-7]. Frew AJ (2000), Lo A, Chetmg C, Law R (2004), Law R. Huang T (2005),

* Corresponding author.

Marcussen CH (2005), Huang Xueying (2008) from the information browsing, retrieval, service marketing point of view of the virtual tour has been studied .

The United States, the most famous virtual travel site "Second Life" (www.secondlife.com), for example, now has more than 1,000 spots of the virtual scene, visit the web site visitors to download software and procedures, and then registered, and design their own network identity, can be provided on the site has more than 1,000 tourist attractions in the virtual world, has attracted more than 460 million members. According to the American people and specializes in Internet research institutions Pew Internet Americanlife published a survey report showed that among U.S. Internet users, 45% of adult American Internet users have had in the online virtual tour experience. Some older users prefer to participate in virtual tourism to 40-49 age group, for example, 52% participated in the Virtual Tour. Opening of the domestic online travel sites are "Ctrip website" (www.ctrip.com), "China Tourism Network" (www.ctn.com.cn), "swim skin Network" (www.youpitrip.com), etc.. In addition, Britain, Canada and China's Articles Glenn Napier University Xing Xing Digital Technology Co., Ltd. jointly developed the Virtual Summer Palace; Dunhuang Academy is developed by the "Digital Dunhuang virtual roaming system" for our virtual tour laid the foundation for the development of applications.

3 The Analysis of Advantages and Disadvantages of Traditional Tourism and Virtual Tour

3.1 The Superiority of the Virtual Tour and Limitations

Virtual tour using virtual reality technology, through the simulation reality or surreal landscape, "never leave home" can "see the world", with great superiority, see table 1.

Table 1. The superiority of the virtual tour

Superiority	Main Feature
perceptibility	Vision, hearing, smell, taste, touch, and other aspects of real experience
TaraLabs	The freedom to choose the past, present, and future and fantasy virtual scene
economies	Avoid traditional tourism in the many obstacles, such as cost, transportation, time, etc.
convenience	Can be anywhere, at any time, tourism
comfortable、 safety	Don't leave familiar environment, experience, avoid traditional tourism in all kinds of accidents can happen
universality	No matter whether physical strength, all can participate in limited; Especially suitable for the disabled and the elderly
sustainable development	Avoid the traditional tourism caused environmental degradation

It is because of virtual tour with these many advantages, by many Internet users, but the praise highly of virtual tourism also has its obvious limitation (see table 2).

Table 2. Virtual tour limitations

Limitations	Display
The authenticity	For technical reasons, simulation and the traditional tourism completely consistent tourism scene it difficult, no matter how fine making 3 d simulation and the real environment, still have nuances.
Physical harm	The long hours in the online travel will cause the body of tourists discomfort
Social harmfulness	Easy to make teenagers addict network, escape from society, realize.
Activities limited scope	But only in the location of the network offers travel, lack of tourism autonomy.
The limitation of the users	The user must understand the Internet, have a certain network operating ability.

For most old people on the computer network is not familiar, to learn and have certain difficulty, natural were excluded from the network virtual tourist activities outside. Next to the network computer primarily virtual tour operation mode, use for a long time can cause all kinds of disease, cause body all sorts of discomfort. So the virtual tour still does not have the traditional tourism pleasurable double characteristics, completely replace the traditional tourism.

3.2 The Advantages and Limitations of Traditional Tourism

As people living standard rise, consumption level also increase. Tourism this widen our sight, relieve pressure, pleasurable consumption by more and more DuoRen's favorite form. According to the national tourism administration 2007-2009 China tourism statistical bulletin material, in 2007, the number of domestic tourism 1.61 billion people, and the income is 777.062 billion yuan RMB, a 15.5% increase from the previous year respectively and 24.7%; Chinese citizens outbound toll to 40.954 million people, and a 18.6% increase from the previous year; Tourism income reached 1.0957 trillion yuan RMB, a 22.6% growth from last year. In 2008, the number of domestic tourism 1.712 billion people, and the income is 874.93 billion yuan RMB, a 6.3% increase from the previous year respectively and 12.6%; Chinese citizens outbound toll to 45.8444 million people, and a 11.9% increase from the previous year; Tourism income reached 1.16 trillion yuan RMB, a 5.8% growth from last year. In 2009, the number of domestic tourism 1.902 billion people, and the income is 1.018369 trillion yuan RMB, a 11.1% increase from the previous year respectively and 16.4%; Chinese citizens outbound toll to 47.6563 million people, and a 4.0% increase from the previous year; Tourism income reached 1.29 trillion yuan RMB, a 11.3% growth from last year of [7]. Visible tourism has become part of the lives of people, its advantage sex mainly in the following aspects, see table 3.

Table 3. The advantages of the traditional tourism

Superiority	Main Feature
relieve stress	Through the tourism, release pressure, adjust the mood, life in the more fierce social competition
Enhancing understanding	In the journey, make new friends, expand the scope of social communication, and create more opportunities
Feel nature	Close to nature, see FeiLiu waterfall, the birds singing, smelt smell flowers and plants, product brocade show native land, experience of natural beauty.
Strengthen body	The mountains, wading play, breathe fresh air, and promote the healthy body.
Increase knowledge	Understand the local geological and geomorphic folk custom, history legend, fun, celebrity open field of vision
cuisine	Taste the local specialty, understand local culture, in the entertainment of humanities culture influence, and accept the charm of feeling

The traditional tourism in to people happiness at the same time, it has also exposed a some limitations. Travel expenses climbed, but travel services and tourism quality but sell at a discount greatly. Tourist traffic exhaust gas accelerated the deterioration of the environment; The expanding tourism scenic spots and to avoid wildlife habitat growth environment; People like, not only make dispatches overshadowed the natural beauty, but also accelerated the precious natural heritage [7] the loss (see table 4).

Table 4. The traditional tourism limitations

Limitations	Display
Economic burden	The traditional tourism must have corresponding economic support, want to get high quality tourism services, will need to pay the higher economic cost.
Accident risk	The controllable by many factors, tourism will inevitably assume certain accident travel risk
Excessive fatigue	For saving travel expenses, usually travel arrangements, hoping compact. Not enough rest.
Worry about	Worry transportation, accommodations, meals, play in the process of safety; The body adapt to the environment and overworked ability
Damage the environment	Travel garbage, not civilization behavior to the scenic area, the natural environment of the great cause damage.

4 Virtual Tour Interactive Product Design Research

For tourism will and natural protect double demand, exploring a collection of virtual tour and the traditional tourism advantage as one of the new tourism way both satisfy people relieve pressure, strengthen the body, the closest natural desire, and no natural ecological unduly influence. Improve network virtual tour interactive way, the new interactive transmission ways.

4.1 Interactive Product Design Ideas

Use modern network virtual tourism research achievements, to change the current use of mouse and keyboard for virtual tourist mode of operation that sports fitness equipment and virtual tourist imaging, and the combination of through the motion control virtual scene change, realize movement and virtual tour synchronization.

4.2 The Virtual Tour Interaction Design Product

Fitness equipment and other electronic equipment of interaction design with the comparatively successful example. If the revolution and the TV connection, dancing machine and electronic display and control equipment of connection, etc. The working principle, research is will present the hand of electronic equipment control by foot to transfer induced control movement. Using this approach, through the study of the mouse and the relative motion working principle, will face the sports fitness equipment as the mouse face, foot movement hand moving the mouse action substitution, use the existing mouse working principle capture fitness equipment working direction and speed movement, change the cursor movement and the relative positions, achieve fitness equipment and computer virtual scene interaction movement.

Specific transmission principle chart 1-3. Control method for: people in the running, drive running on sports bring transmission, simulation mouse moved back and forth, the relative motion of the virtual scene before and after the image transformation direction. Support rod inserted into a fixed set of tube, turns the rotary transmission through virtual scene around the image transformation direction. The above activities are placed in the base by light signals to optical mouse, and through the standard chip USB interface or PS / 2 interface and computer, TV or other can output image output devices connected, realize the interaction of the virtual scene and exercise, can be in motion at the same time, control the scene of the virtual tour changes and meet tourism pleasurable dual characteristics. Each parts name see table 5.

Table 5. serial number and name mapping table

Parts serial number	name	Parts serial number	name
1	key switch	11	Optical engines of circuit board
2	hand shank	12	cover board
3	bracing piece	13	Running with
4	foundation	14	Running down side
5	Bottom of bracing piece	15	Lens component
6	active stem hole	16	light hole
7	connecting rod	17	luminous diode
8	hinge pin	18	Image inductor
9	follower lever	19	Photoelectric mouse chip
10	turntable		

5 Conclusion

Virtual tourism product design research, based on the interaction of traditional tourism demand, people from avoid excessive tourism on the natural environment of damage, application of modern means of science and technology and virtual tourism research achievements, improve long-term network virtual tour against health limitations. More from the human point of view, for young and old will be easy movement way and tourism experience as a whole, for virtual tourism development in the future to provide a new way of thinking. By leveraging virtual tourist interactive products, not only can realize everyone tourism dream, and economic material benefit without any risk, it is people pursue entertainment of a kind of new way choice.

Acknowledgments. This paper supported by the department of education of HeNan province natural science research program(No.2011A630062),the department of education of HeNan province humanities and social science research project(No/2011-ZX-233),Henan government decision-making tender (No.2011A630062),supported by the Innovation Scientists and Technicians Troop Construction Projects of Zhengzhou City (No.10LJRC183).

References

1. Huang, X., Zhang, W.: Modern tourism service marketing new ways of virtual tourism phenomena analysis. *Shanghai Scientific Management* (1), 89–90 (2008)
2. Peng, F.: Virtual tour application and development. *The Modern Market*, 114–115 (2010)
3. Hernandez, L.A., Taibo, J., Seoane, A.J.: Immersive Video. And for country. In: *t Video, as for Multimedia is A marketing method*, Denver, USA, August 21, pp. 63–73 (2001); ISSN: 0277-786 X, ISBN: 0-8194-4244-5
4. Sandau, R.: [EB as country. / OL]. *International Symposium on the Role of the Restructuring in the New Economy*. In: *Judge not According as Agronomy Infra-, Benefit*, Damascus, Syria, December 15., vol. 17 (2003), <http://elib.dlr.de/10783/>
5. Guo, X., Miao, F., Wang, H., et al.: Based on G/S model structure of the digital tourism service platform study. *Remote Sensing Technology and Applications*, 190–195 (2009)
6. Feng, A., Wang, Y.: Three-channel HuanMu stereo projection technology in the application of tourism education. *Science Square* (07), 149–151 (2009)
7. Pouyan, A. A., Beigi, A.H., Kadkhoda, M.: An Agent-Based Model for country. *Object Petri Nets and put*. In: *5th WSEAS International on Circuits, Systems, 't Control Signal-may Well (CSECS 2006)*, Dallas, Texas, USA, pp. 39–44 (2006); ISSN: 1790 5117, 960 8457. ISBN 55-6

Rapid Implementation of Communication-Related Simulation Equipment on the Open Web-Based Virtual Lab

Yue Zhao¹ and Fuan Wen²

Beijing Key Laboratory of Network System and Network Culture
Beijing University of Posts and Telecommunications
Beijing, P.R. China
uuyueyue@gmail.com, fawen@bupt.edu.cn

Abstract. It mainly describes how to easily realize a piece of new equipment needed in the communication-related simulation experiments on the open web-based virtual lab (OWVLab). This lab supports users to do circuit simulation experiments via the Internet, and we expand the range of experiments to communication-related area by using the visual circuit editor. Through building the circuit diagram of a certain equipment to simulate its function and packaging the circuit as a piece of new equipment, we have implemented much communication-related equipment by calculating the circuit. And users can log in the platform using the visual circuit editor to customize equipment. Therefore, users can easily expand the equipment libraries for communication-related experiments.

Keywords: Simulation equipment, circuit visual editor, rapid implementation, communication.

1 Introduction

Rapid development of communications and Internet has greatly promoted the progress of science around the world, also made people's lifestyles and learning has undergone profound changes. At the same time, colleges and universities are lack of experimental teaching resources, especially the large number of real experimental equipment which has large costs of purchase, maintenance and loss. And also, United States, as today's technological power, virtual laboratory has been included in its strategic planning of research and development, in order to maintain its leading position in science and technology [1].

Through investigating circuit simulation software in the world, we have found that lots of popular software for circuit simulation supports users to expand equipment library and edit equipment [2], but all these functions are too professional whose focus is on how to package equipment rightly and using spice statement strings to describe the device model to fit for experiment courses. Virtual experiment platforms in China are still stay the stage that users cannot edit and extend equipment library.

When mentioned web-based experiment platform of communication, it is mainly based on Matlab Web Server which cannot support users to customize components and expand the equipment libraries. And it saves the .m file of the experiment on the server which also contains the html file to input parameters and show results. The shortcoming is that it cannot reuse the realized experiment, and we must write separate .m file for every experiment.

Relying on technological progress and lifestyle changes, we designed and implemented a virtual experiment platform based on Web. In order to implement the virtual equipment easier, also avoid repeating progress of development, we developed the visual circuit editor to combine existing components and package them as a piece of new equipment. All the new equipment can be added to the equipment bar. Therefore, in addition to using fixed equipment, users can log in the web-based platform to create equipment meeting their needs via Internet, which increases the scalability of OWVLab, reduces the developers' workload and difficulty of equipment implementation, and allows users to remote login to create or do experiments.

2 An Overview of the Implementation and Design

Based on the virtual web-based platform, we implemented a visual circuit editor to expand the libraries of simulation equipment. We have an equipment bar containing libraries of equipment which provides resistors, capacitors, inductors, transistors, voltage sources, current sources, FET and other basic equipment, also provides the adder, flip-flops, counters, digital chips and other digital components, even provides an integrated operational amplifier, bridge rectifiers and other complex combination of equipment. Different equipment can be stored in different libraries, and users can extract the needed equipment from the equipment bar to the platform to build a circuit in order to complete the experiment. The visual circuit editor has great scalability, supports dynamic customization and expansion of simulation equipment, therefore, users can add and delete equipment dynamically, change any equipment configuration and properties at any time. Due to the circumstances of the virtual platform, the visual circuit editor, and building the internal circuit of the new equipment packaged as a piece of new equipment is the same, it is easy to understand the implementation for users. Using the visual circuit editor, users can add the appearance of the equipment and simulation models, modify a variety of attributes of the basic equipment, and edit the internal logical structure of the combinative relationships and components in visual environment on OWVLab.

The visual circuit editor uses Java language for the development, Java Swing for the interface, Java2D technology for graph drawing, "shell" of models defined by SPICE using XML file as the storage carrier [3].The editor mainly contains four key parts which describe the main external settings or internal process [4].

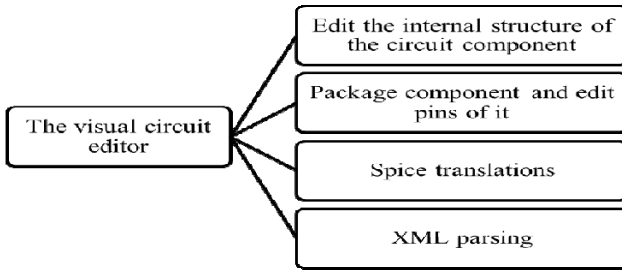


Fig. 1. Key parts of the visual circuit editor

2.1 Edit the Internal Structure of the Circuit Component

Define the internal logical structure of the circuit component. For example, use several diodes and transistors to build the internal circuit of series of 74LS chips.

2.2 Package Component and Edit Pins of It

Add the image of the component, pins and sensor area of pins. Editing pins can ensure positions of them, and also define input and output characteristics.

2.3 Spice Translations

Spice is a recognized circuit description language, and the visual circuit editor also follows this standard. This module is mainly responsible for translating the internal structure of the component to the spice statement strings.

2.4 XML Parsing

Describe the internal logical structure of components, generate and parse predefined XML syntax.

2.5 The Overall Process

The simulation equipment created by the visual circuit editor can be converted into the circuit description statement of spice because of using X-spice simulation program as the kernel. Simulation equipment must be combined at any levels, and the circuit module or complex equipment with their own certain functions is built and connected with basic components, which will be frequently reused in other larger circuit also including many basic components.

3 Detailed Implementation

Through the investigation into experiments of signals & systems and communication theory, we know that the key equipment includes multipliers, signal generator, low pass filter, dual trace oscilloscope, adder, clock source, D flip-flop, Butterworth filter, spectrum analyzer, the baseband signal, Gaussian noise source, PN code

generator, differential encoder, samplers, band-pass reconstruction device, zero-crossing comparator, the power spectrum analyzer and so on.

All these equipment and devices are obtained by the real circuit in reality. We build its internal circuit and package it as a piece of equipment or a device. In order to illustrate how we implement new components by visual circuit editor, we take pseudo-random sequence (PN code) generator for example due to frequently being used as signal source. The specific process is as follows:

3.1 Ensure the Circuit Diagram of the New Equipment

First of all, we should find the real circuit of generating pseudo-random sequence. We decide to use the diagram of generating m sequence which is one kind of pseudo-random sequences used frequently.

M sequence is the brief name of longest linear feedback shift register sequences usually generated by using shift register and logic gate. We take five shift register which period is thirty-one as an example, and choose the feedback function is $Q_2 \oplus Q_5$. According to the principles of m sequence and the simulation circuit diagram, the sequence must be “1110 1100 0111 1100 1101 0010 0001 010”.

Then we make a simulation of that circuit on circuit simulation software to verify the correctness in order to sure the accuracy of the circuit diagram. The pseudo-random sequence generator consists mainly of components: D flip-flop or chip 74LS74 (contains two D flip-flop), XOR gate, or gate, the pulse source, VCC and Ground.

The actual sequence simulated by the circuit simulation software of Multisim is “1110 1100 0111 1100 1101 0010 0001 010” which is same as the theoretical sequence, so we just verify the diagram.

3.2 Ensure All the Needed Equipment of the Diagram

Split the internal circuit into individual components: D flip-flop or chip 74LS74 (contains two D flip-flop), XOR gate, or gate, the pulse source, VCC and Ground. And the chip 74LS74 consists of two D flip-flops. We have developed most of basic components and commonly used devices (equipment), including all we needed components in the diagram of generating m sequence.

The D flip-flop is a basic component described by circuit description statements meeting spice standard, and the specific parameters of D flip-flop can be referred to the X-spice manual for free. Following is example of D flip-flop’s final translation in X-spice:

Example SPICE Usage:

```
a7 1 2 3 4 5 6 flop1
.model flop1 d_dff(clk_delay = 13.0e-9 set_delay = 25.0e-9
+          reset_delay = 27.0e-9 ic = 2 rise_delay = 10.0e-9
+          fall_delay = 3e-9)
```

Fig. 2. Example of D flip-flop in X-spice manual

The first line describes how the elements connect in the circuit. The meaning of each parameter is as follows: .model is show where is the beginning of circuit model; flop1 is the name of current trigger (component); d_dff is the model name of D flip-flop in X-spice; clk_delay is the clock delay parameter; set_delay is the set delay parameter; reset_delay is the reset delay parameter; ic is output initial state; rise_delay is the rising edge delay parameter; fall_delay is the falling edge delay parameter.

The chip 74LS74 is made of two D flip-flops, so we need to create this chip by using D flip-flop, but in actual, we have provided 74LS74 on our virtual experiment platform.

3.3 The Process of Building and Packaging the Needed Components

Now we need to use D flip-flop and other basic equipment to build chip 74LS74. After verifying that the result on OWVLab is the same as the theoretical results, we create chip 74LS74 according to its simulation diagram. Firstly, we create a new component and set some parameters.

The specific process of creating a new component is as follows: Firstly, select the picture of the component. Select a real picture of 74ls74 chip. Secondly, set the type of the component: choose which type the component belong to, for example, a three-state gate, analog chips, digital chips, and so on. 74LS74 select the type of digital chip. Thirdly, set Model of the component: No need temporarily. Fourthly, instructions of the component: 74LS74 consists of two independent components of D flip-flops, CLK pins connect the clock pulse, CLR pins connect clear pulse, PR pin is an enable pin, D pins are inputs, Q and Q' pins are output. Fifthly, model properties of the component: include internal circuit diagram and truth table. Sixthly, set pins: set pins according to the actual pin number, and add sensor area on the corresponding pin of the chip image. Chip 74LS74 has a total of 14 pins, arranged counterclockwise from the lower left, add pin 1 to 14 and set the corresponding sensor position on the image of 74LS74 chip.

After setting all the items, click “OK” to submit and then get into the interface of the virtual experiment platform where we can start to build the internal circuit of 74LS74. On the top of the platform are pins for inputs and outputs. After dragging all the needed components to the bench, connecting them and the pins all, click “Files”→”Save Files” in the upper left corner of the platform to save chip 74LS74 which will appear in the equipment bar. And the process of package is over.

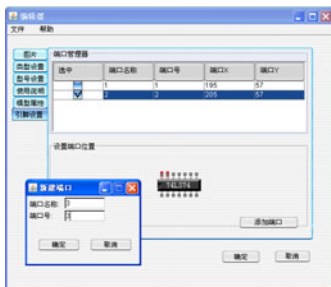


Fig. 3. Settings of chip 74LS74's implementation

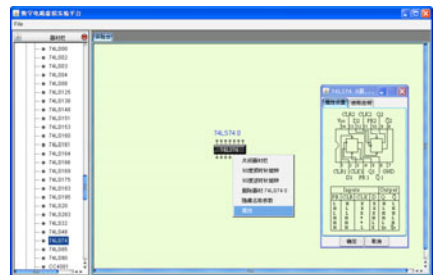


Fig. 4. New created chip 74LS74

Drag the new created chip 74LS74 to the platform, and then Right-click on the chip, it will appear property and instruction of the chip, the property includes an internal circuit diagram and truth table, the instruction includes features of pins.

3.4 Verifying the Results

Use a packaged chip and the basic equipment to build the internal circuit, and then verify the result.

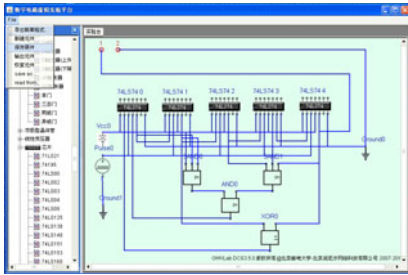


Fig. 5. Internal diagram of generating m sequence

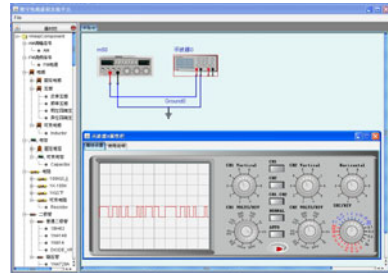


Fig. 6. Result of m sequence generator

Check the waveform generated by customized pseudo-random generator in our platform, and then compare it to the waveform calculated by principals of m sequence which is “1110 1100 0111 1100 1101 0010 0001 010”. The sequence generated on OWVLab is “1110 1100 0111 1100 1101 0010 0001 010” which is same as the results of principals. Therefore, we successfully implement components with existing equipment to expand our equipment library to meet more requests.

3.5 Using the New Equipment to Do Related Experiments

Use the new equipment of PN code generator as signal source in experiment, such as ASK, PSK, QAM, simple baseband transmission system and so on.

4 Results and Prospects

Using the visual editor we can add any necessary equipment and expand our equipment library, also users can build their own components required. This method reduces the difficulty of developing the components, so users can create their needed components, and also focus on experiments of signals & systems and communication theory because of not worrying about the lack of key component in the simulating process of experiments.

We have extended many components used in the experiments of signals & systems and communication theory, such as Butterworth bandpass filter, decoder and so on. And also we have completed many experiments of signals & systems and communication theory. Following is some realized experiments of signals & systems: zero input, zero state and complete response, signal decomposition and synthesis,

distortion-free transmission system, the analysis of analog filters, continuous time system simulation, second-order circuit transfer characteristic, second-order circuit transient response. At the same time, communication principle experiments are amplitude shift keying (ASK), binary differential encoder and decoder, the energy spectrum of rectangular pulse density, phase frequency shift keying (PSK), double-sideband suppressed-carrier amplitude modulation (DSB-AM), amplitude modulation with discrete bilateral large carrier (AM), cosine signal waveform and spectrum, single sideband amplitude modulation waveform and spectrum (SSB) and so on. Later we will implement more communication-related equipment and experiments by using the visual circuit editor on OWVLab and expand it.

Acknowledgement. Foundation: This paper is supported by the Key Project in the National Science & Technology Pillar Program of China under grant No.2008BAH29B00 and the Fundamental Research Funds for the Central Universities under grant bupt2010pt10.

References

1. Bender, A., Edman, A., Sundling, L.: A Combined Knowledge and Hypermedia System to Attain Educational Objectives. In: Proceeding of the Sixth IFIP World Conference on Computers in Education, pp. 235–242 (1995)
2. Christoph, M.H., Houstis, E.N., Rice, J.R.: SoftLab-A Virtual Laboratory for Computational. Science Mathematics and Computers in Simulation, 497–491 (1994)
3. Lu, B., Shanguan, Y.: The design and implementation of editable simulation equipment library of analog circuit. Beijing University of Posts and Telecommunications, 14–25 (2009)
4. Lu, Z., Wen, F.: The design and implementation of the visual virtual editing equipment system. Beijing University of Posts and Telecommunications, 27–31 (2010)
5. Wen, F., Suo, D., Chen, M., Li, J.: An Openlab Platform with Virtual Experiment Learning System. In: 2010 2nd International Conference on Network Infrastructure and Digital Content, 0170-303 P839-P843 (IEEE IC-INIDC 2010), Beijing, China, September 24-26 (2010)

The Principle and Application Exploration of Kademlia's Structured P2P

Xiaochun Shi, Guoqing Liu, and Jinchun Feng

Sichuan Engineering Technical College, Deyang, 618000 Sichuan, China
{Sxc,lgq}@scetc.net, catsmiling@163.com

Abstract. Kademlia is the P2P overlay Network Communications Protocol designed by Petar Maymounkov and David Mazières, which structures the distributed P2P computer network. It is the P2P information system based on the XOR, lays down the network structure and specifies the node communication and the mode of mail exchanger. Kademlia is the third generation of P2P Peer-to-Peer technology and a network mechanism of Structured P2P. It is based on the DHT (Distributed Hash Table) and XOR, and achieves Overlay Network structure, node instant allocation and the explicit search of resource, which is used widely in the new generated P2P software, such as BitTorrent and OverNet.

Keywords: Kademlia, Structured P2P, DHT, Application.

1 Introduction

According to structure overlay network in physical network, P2P technology adopts special IP route and search strategy and shares the node and CPU, which is widely applied in the field of collaboration computing, content store, content share and distribution, instant messaging etc. The study of P2P technology focuses on four aspects: overlay network structuring technology, targeting content and search technique, content downloading technique and other assistive technology. According to the structuring pattern of overlay network, P2P is usually divided into Structured P2P and Un-Structured P2P. Un-Structured P2P, basing on the randomly structuring style, adopts BFS or DFS to do the targeting content and search, the typical example such as: Gnutella [1]. But Structured P2P is mainly based on DHT (Distributed Hash Table) technical structure, which is better than Un-Structured P2P in targeting content and search, the typical example such as: Chord, CAN, Pastry, Kademlia [2] etc.

2 The Theory Base of Kademlia

2.1 DHT (Distributed Hash Table) Technology

Simply saying, Kademlia is a sort of DHT technology. While compared with other DHT implementation technique, such as: Chord, CAN, Pastry and so on, Kademlia, by unique XOR as the distance measuring basis, establishes a brand-new DHT

topology structure, and greatly improves the routing inquiry speed matched with other algorithm. DHT uses the DH Function to access (key, value) any data structure in the $O(1)$ order time[3]. DHT has three elements: I. Key, any significant identification, like IP address etc. II. Value, connected with key, like specific file content; III. DH Function hash (key), mapping key aggregation to the stored (key, value) aggregation. One of the important characters of DHT is one-way. That is to say, if given X, we can find key, and the counting of: $x = \text{hash}(\text{key})$ is infeasible. The second character is that the ideal order of access counting is $O(1)$. The third character is that DHT forms a unified mapping address space, which is simply operated, only having three basic operations: Insert, Delete and Search. Therefore it is widely used in the Computer System and database system.

DHT is the theoretical principle of structuring P2P. The resource of P2P (e.g. node, file etc.) is mapped in the united addressing space. In terms of the overall situation, the resource is distributed in a DHT and the key problem is how to structure and store it. Usually DHT is divided into many continuous partial spaces, and distributed into the nodes as the special principle.

2.2 XOR Metric

Binary XOR metric is defined the following: $1 \text{ XOR } 1 = 0$, $1 \text{ XOR } 0 = 1$, $0 \text{ XOR } 1 = 1$, $0 \text{ XOR } 0 = 0$. XOR has an important character: supposing that a, b, c are three random numbers, then: $(a \text{ XOR } b) = (a \text{ XOR } c)$, here only $b = c$, which guarantees the nodes can determine **vicinities'** nodes distance and be sequencing through XOR Metric in the Kademila.

3 Kademlia System Descriptions

In the Kademlia, every node, using the node's distance as the base, keeps the **vicinities'** nodes by k-bucket and refines the k-bucket as the message's flushing strategy so as to gain the whole network's in-plant views. Based on the analysis of it, the node orientates new **vicinities'** nodes and searches the resource information through the node _lookup. Kademlia transfers the message and data by the UDP protocols.

3.1 Overlay Network Structuring

Node and Resource Information. In the Kademlia system, the node is identified by node ID whose bitsize is 160 bits. All of the nodes are considered as the leaves of one binary tree, and every node is located by the shortest prefix of its ID. To any of the node, it can distribute the binary tree to be a series sub-trees but not including itself. The highest sub-tree is made up of the other part (not including itself); the lower layer sub-tree is made up of the rest part (not including itself) etc. until the whole tree is distributed. Fig 1 shows how the node 0011 is divided by its sub-trees:

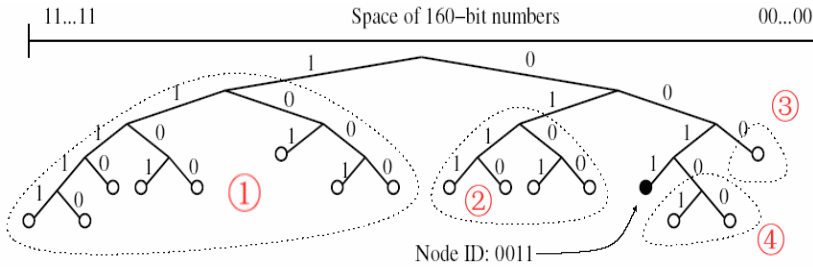


Fig. 1. Sub-trees the partition of node 0011

All the node IDs forms an address space and lets them be different and node ID is the base of Kademia system. As to how to get node ID, there is no any explanation, which has free definition in all kinds of realized software systems. Resource has the Key mark of fixed length and usually Key is gained through file by Hash Function, such as: SHA1. All of the Keys makes up of an address space. Therefore, there are 2 address spaces in kademia; all nodes form node ID; all file resources form Key address space.

Distance and K-bucket. In Kademia, distances is defined as any 2 keys or node ID to do the XOR metrica, and the result is: $d(x,y)=x \text{ XOR } y$, and distance is the base of other operation. Node uses related people table to mark other nodes, including node ID, IP and Port. Every node maintenances N related people table(Usually 160). Every table contains the related people information with the same distance to the node, in which the i table includes the distance of k related people: $2i \sim 2i+1$ (usually $K=20$). This kind of table is called k-bucket.

In each k-bucket, linkman is arranged as the time order: the earliest linkmen are at the beginning, and the latest at the bottom. K-bucket is dynamic, and it often keeps the linkman who has relation in one hour. When receiving a new-node message (2.2.2 node _lookup) , it can counts the distance and flushes k-bucket: (1) if new-node is in the opposite k-bucket, it is moved to the bottom of K-bucket; (2) if new-node isn't in the opposite k-bucket and the k-bucket isn't full(nodes <k), it is added to the bottom of K-bucket; (3) if new-node opposite k-bucket is full, send PING message to the top linkman of the k-bucket. If receiving answers, move the address of the answered linkman to the bottom, otherwise, cast it aside and add new node to the bottom .According to the flushing strategy of using information and fixing time, every node keeps different- distance online linkman's message, and gets the whole network 's in-plant views, that is Node ID space. It is only when the node isn't online that it can be thrown away, which proves the stability of in-plant views in order to prevent effectively from attacking by Dos.

Distributed Storage of (key, value). Like what said in NO.1.1 the key problem to structure P2P is how to structure and store DHT. Kademia structures and store DHT by the distance k-bucket mechanism, and then store the Key resource information on the base of it. In short, Key is stored in the nearest node through XOR metric which

realizes by the operation of kademlia's STORE (key, value) . At certain time, a few nodes keep a great many Keys in kademlia, which is most important for searching resource.

3.2 Locating and Searching

Original Language Operation. Locating and Searching is the nuclear problem in P2P system. In kademlia, it defines 4 Original Language Operation RPC: PING, STORE, FIND_NODE, FIND_VALUE, in which PING explores whether a node is online so as to search resource, FIND_NODE locates the node and FIND_VALUE searches resource. All of the RPC message has a random RPCID of 160 bits, which produces by the sender, and receiver should copy this RPCID in his returning information when responding to each message. The purpose is to avoid the address being fabricated.

FIND_NODE (or called FIND_CLOSE_NODE) uses 160 bits ID as parameter, and the receiver must return the nearest linkman information which is already known K distance ID (If all the k-bucket linkmen don't have K, it should return all the linkmen information.) .

FIND_VALUE is the same as FIND_NODE, but there is one difference that if receiver has already get a STORE RPC with this Key, it will return the other parameter Value to the sender, otherwise, the result is the same as FIND_NODE.

Node _lookup. The node _lookup of Kademila is based on the above 4 Original Language Operations whose purpose is to find the nearest node to the K distance searching ID. This method is called recursive algorithm, and its clue is the following: (1) the beginning node takes out A linkmen (usually A=3) from the nearest k-bucket as the shortlist and then does the parallel and asynchronism of FIND_* operation. If the shortlist linkman is online or responses in the given time, it will return to K or less than K linkmen, and put these linkmen into the shortlist, or else deletes the linkmen from the shortlist; (2) Choosing A linkmen whom never have any relation from the shortlist, do the process done in (1) ; (3) if it doesn't find the nearest linkmen in a round, send FIND_* operation to K nearest linkmen who haven't been communicated; (4) the condition of searching end: the returning linkmen are not nearer than them in the shortlist and k-bucket, or having already returned K online linkmen.

At last, the node gets K online linkmen being at a distance from ID or search resources (value). In this process shortlist is update; in each round, FIND_* operation will compare the returning linkmen to the ones in the shortlist in order to make the linkmen in the shortlist nearer to the ID, till it is mostly nearest or reaches K.

Here α is a parameter of system optimization, just like K. In the realization of Bit Torrent, supposing $\alpha=3$.

When $\alpha=1$, the process of inquiry is the same as the process of hop-by-hop inquiry of Chord. Like fig 2.

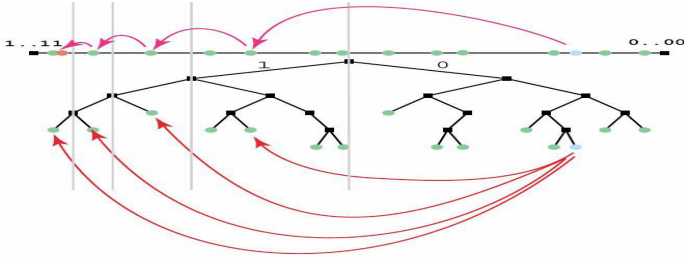


Fig. 2. The inquiry process of $\alpha = 1$

The Opting in and out of Node. When the node is added into Kademlia, it first gets the message of other linkmen and adds it to the relevant k-bucket to do node `_lookup` (its own ID) so as to flush all of the k-buckets. Up till now the node has got the partial views of the whole network. When the node exits the Kademlia, other nodes will get rid of this node as flushing k-bucket if disconnected. Or else, it sends exiting message to other nodes, and lets them flush the k-bucket .

Issue, Store and Search (key,value). On the base of Node `_lookup`, the storage of (key,value) is very simple. According to node `_lookup` (key) it can find K nodes(linkmen) nearest to Key, and sends STORE operation. The process of searching (key,value) is the same as node `_lookup`. If it succeeds, it will return the Value, and send STORE operation to the Key nearest unreturned Value node; otherwise, return K near nodes.

4 The Application of Kademlia

Established on the bases of DHT and XOR, Kademlia system provides a highly efficient distribution storage system and deposit genesis, it is applied in many p2p software such as: Overne, Emule, Revconnect, bittorrent and its open office software, retroshare, sharkypy and so on[2] since Petar Maymounkov and David Mazières put forward in 2002[4]. In view of the DHT structure, Kademlia has the P2P merits of DHT, that is , system structure has high robustness, keywords accurate searching, node efficient localization, average step length $O \cdot \log N$ and the system produces less message in the dynamic condition. Its fault is that it didn't consider the isomerism of the system, that is to say the node ability (CPU, storage and band width) is different, and has match problem of logical topology and physical topology.

In the concrete realization of P2P, Kademlia is often looked as the infrastructure layer, and on the duty of structuring overlay network and route, but the specific application like Content Delivery and downloading is set up on it. Kademlia can be used in the sharing network, so we can find the article what we need to download in the file-share network through making Kademlia -Keywords searching. For example, Kademlia realizing in BitTorrent[5], routing list realizes still consulting the standard k-bucket, and defines KRPC and DHT Queries(including PING、 FIND_NODE、 GET peers、 ANNOUNCE_PEERS) to locate and search. But the download of Bit Torrent

still uses the standard agreement and strategy, c.f. the Tir-for-Tat strategy and section First Algorithm. Using BitTorrent in Kademlia mechanism realizes Trackerless, that is to say there is no need of Tracker server to distribute quickly and download the file. The standard eMule[6] agreement makes the Kademlia mechanism come true and it defines the keywords dictionary(Key, file information list) and file index dictionary (file Key , owner list) and then store the dictionary distributed. The process of downloading the file is like this: first, the node inquires the keywords dictionary through any one keyword or its combination, and then it gets the file SHA1 proof test value of the matching keyword. After that, it gets all of the net nodes of providing the file downloading by proof test value inquiring the file index dictionary. At last, download the file in the nodes with the passage - passage method downloading.

Except widely applied in the content distribution and downloading field, Kademlia still makes progress in the other field's realistic application study, such as: streaming media [7][8] , pronunciation VoIP[9] etc.

References

1. Gnutella, <http://www.gnutella.org>
2. Kademlia, <http://en.wikipedia.org/wiki/Kademlia>
3. Xiong, J.: Study of the Routing Mechanism and Key Technology in P2P. China University of Science and Technology (April 2006)
4. Maymounkov, P., Mazières, D.: Kademlia: A Peer-to-Peer Information System Based on the XOR Metric. In: Druschel, P., Kaashoek, M.F., Rowstron, A. (eds.) IPTPS 2002. LNCS, vol. 2429, pp. 53–65. Springer, Heidelberg (2002)
5. DHT Protocol In BitTorrent BEP, http://www.bittorrent.org/beps/bep_0005.html
6. Kademlia In eMule, <http://www.emule-project.net/>
7. Wu, J.L.: KadStreaming: A Novel Kademlia P2P Network-Based VoD Streaming Scheme. In: 7th IEEE International Conference on Computer and Information Technology, CIT 2007, pp. 405–410 (2007)
8. Lu, Z.H.: Design and Implementation of a Novel P2P-Based VOD System Using Media File Segments Selecting Algorithm. In: 7th IEEE International Conference on Computer and Information Technology, CIT 2007, pp. 599–604 (2007)
9. Wu, X., Fu, C.: An Improved Kademlia Protocol In a VoIP System. In: IFIP International Conference on Network and Parallel Computing Workshops, NPC Workshops, pp. 920–928 (2007)

Automatic Color Temperature Control Circuit

Tai-Ping Sun, Chia-Hung Wang, Huei-Jyun Lin, and Kai-Yan Li

Department of Electrical Engineering
National Chi Nan University
Nantou, Taiwan (R.O.C.)

{tpps, s95323901, s98323544, s99323566}@ncnu.edu.tw

Abstract. The brightness of a Light Emitting Diode is proportional to the forward current. Constant LED forward current provides brightness. Temperature is an important factor in LED color temperature. The color temperature varies when the temperature varies. Therefore, an advanced temperature control scheme is proposed. An Automatic Color Temperature Control Circuit with color temperature compensation is proposed. It is confirmed that the proposed system provides stable color temperature within a temperature range from -30°C to 85°C .

Keywords: Light Emitting Diode, color temperature, Automatic Color Temperature Control Circuit.

1 Introduction

The development of power-saving products has become considerable through the ascension of environmental consciousness. The light emitting diode (LED) produces fantastic vivid colors, long-lasting operation, fast response and low power consumption. More industries are choosing the LED as a light source because of these advantages. LED technology is also prominent in liquid crystal display televisions. The issue of LED easily affected by ambient temperature is of great research interest. Controlling the LED light characteristics is extremely important. The positive proportion of brightness is related to the forward current. Stable brightness is achieved when the forward current is constant. Therefore, the LED must have constant output current for steady luminance [1]. The light emitting diode power decreases with the increase in light emitting diode temperature when the forward current is constant. To stabilize the light emitting diode electrical power the conventional power control circuit includes a photo diode [2, 3, 4] for detecting the emitted power and a driving unit for adjusting the amount of driving current based on the result detected from the photo diode. This photo diode provides a current output proportional to the total LED brightness. However, the proportionality of the photo diode output current is associated with the package configuration [5]. The photo diode is used to detect the emitted light at a specific wavelength. As a result, the conventional power control circuit cannot work without a photo-detector. Thus, the cost is increased. In this paper, the LED achieves automatic color temperature control system without a photodiode. When the LED is used to illuminate, the color

temperature is an important factor in presenting the perceived color. Thus, an automatic color temperature control circuit is used to control the color temperature. A micro-controller [6] is used to immediately adjust the color temperature when the ambient temperature varies.

2 Proposed System Architecture

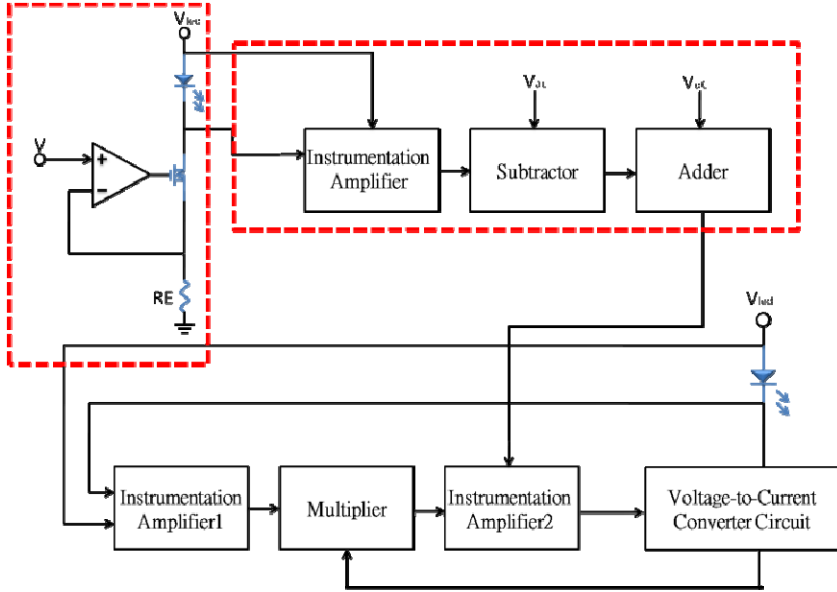


Fig. 1. Automatic color temperature control circuit

2.1 Automatic Color Temperature Control Circuit

The automatic color temperature control circuit (Fig. 1) applied to the RGB LED, replaces the photodiode circuit to correct the output current. The automatic color temperature control circuit consists of an instrumentation amplifier, a multiplier, a voltage-to-current converter circuit, a temperature detector and a color temperature compensation circuit.

2.2 Instrumentation Amplifier1

An instrumentation amplifier is used to obtain the changes in forward voltage across the LED with variations in temperature and restore the voltage value across the LED to its original value after reduction due to the voltage divider system. The variations in LED forward voltage would be reduced and the output used as inputs to the multiplier.

2.3 Multiplier

The multiplier is used to measure the LED DC power consumption and monitor the variation in conductive current and LED forward voltage when the temperature is changed by the multiplier.

2.4 Instrumentation Amplifier2

The instrumentation amplifier is a gain regulator for the color temperature compensation circuit. It restores the voltage value across the LED to its original value after the reduction in forward voltage variation. The excess voltage is output to the compensative gain (V_{ref}) color temperature compensation circuit and compared with the multiplier output voltage. The compensative rate and its outputs are modulated and used as inputs to a voltage-to-current converter circuit.

2.5 Voltage-to-Current Converter Circuit

To provide forward LED current the driver LED needs a voltage-to-current converter circuit. An operation amplifier, MOSFET transistor, and a loading resistor are applied to the voltage-to-current converter circuit with a negative feedback mode which produces a high input and output impedance configuration.

2.6 Temperature Detector

The traditional automatic power control circuits detect the LED brightness with photodiodes. The way to monitor optical power is directly or indirectly irradiating photodiodes to achieve feedback control. This increases the cost and introduces an error due to poor light detection. The photodiodes detects the luminous intensity emitted by the LED in different wavelengths. The introduced current will be different. Thus, the LED is used to achieve automatic color temperature control instead of a photodiode. A temperature detector is applied in the automatic color temperature control circuit, replacing the photodiode. Because the LED has decreased forward voltage when the temperature increases under constant current, the automatic current control circuit is completed with a Voltage-to-Current Converter Circuit (Fig. 2).

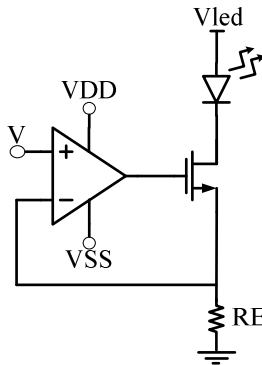


Fig. 2. Automatic current control circuit

2.7 Color Temperature Compensation Circuit

The color temperature compensation circuit includes an instrumentation amplifier, a subtractor and an adder. The color temperature compensation circuit uses an instrumentation amplifier to obtain the variations in forward voltage (Fig. 3.a). The temperature detector output is used as the input to a subtractor. Obtaining a rising V_{ac} curve (Fig. 3.b) that is calculated by the subtractor. The slope of the curve is not necessarily the same as the slope of the compensative voltage. Thus, the subtractor output is used as the adder input. The rising ($V_{ac}+V_{dc}$) curve obtained is the same as the slope of the compensative voltage (V_{ref}) (Fig. 3.c). Therefore, the output voltage is the same as the compensative gain (V_{ref}), and its output is used as the input to an instrumentation amplifier. In this way, we can adjust the current ratio flowing through RGB LEDs when the LED color temperature varies. This reduces the LED color temperature dependence on the temperature, achieving temperature compensation.

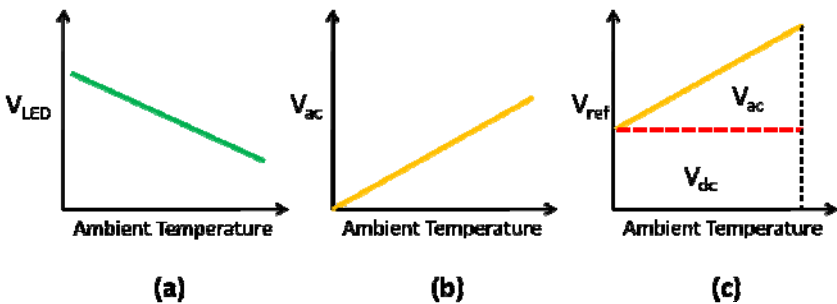


Fig. 3. (a) The variation of LED forward voltage. (b) The compensative curve of V_{ac} . (c) The output of color temperature compensation circuit of compensative gain (V_{ref})

3 Experimental Result

The color temperature control system experimental setup is illustrated in (Fig. 4) A power supply is provided for a drive voltage to the components in the color temperature control circuit. The ambient temperature is set to change from -30°C to 85°C with an interval of 10°C using a precision oven. The LED color temperature was measured using an Integrating Sphere. The full color LED NSSM016AT produced by Nichia Corporation was implemented in the driver. For color temperature control to drive RGB LEDs at the same time, three color temperature control circuits were used to drive the RGB LEDs separately at the same time for white light.

In order to simplify the color temperature control system that use a temperature detector to three color temperature compensation circuits for RGB LEDs because the color temperature control system is complex.

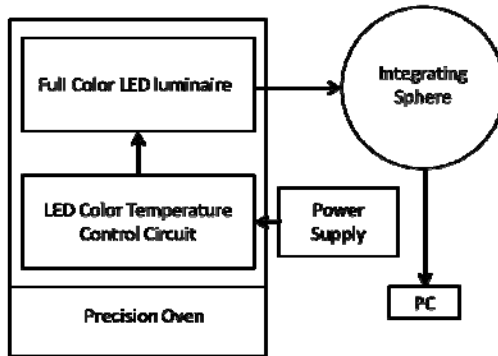


Fig. 4. The simplified illustration of the control system

Fig. 5 illustrates the relationship between the color temperature and temperature with automatic current control circuit for white light. The initial value is color temperature 6500K at -30°C and the variation of color temperature is 31.58% with the temperature range from -30°C to 85°C . Fig.6 illustrates the relationship between the color temperature and temperature with color temperature control circuit for white light. That display color temperature is stable near color temperature 6500K with the temperature ranging from -30°C to 85°C with the variation 11.72%. Table 1 illustrates the experimental results. The variation in color temperature decreased from 31.58% to 11.72%.

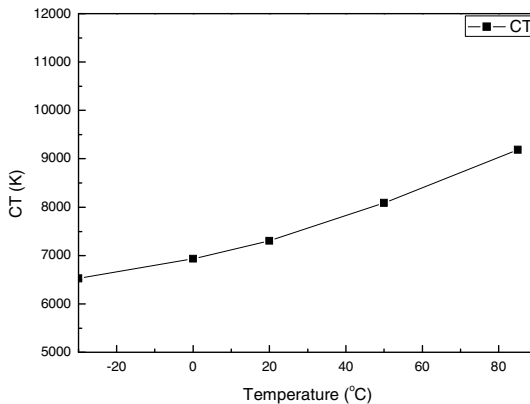


Fig. 5. The variation of color temperature when temperature is variable with automatic current control circuit

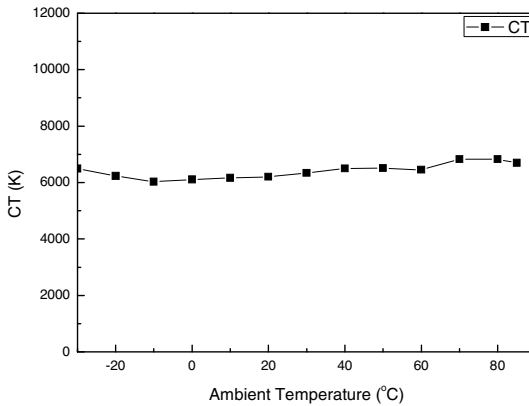


Fig. 6. The variation of color temperature when temperature is variable with color temperature control circuit

4 Conclusion

It is difficult to adjust color temperature because of the white LED which excites Blue-ray phosphors. Thus, color temperature is unchanged for white LEDs. The color temperature of RGB LEDs for white light has a characteristic that color temperature is adjustable because the drive current ratio of RGB LEDs affects the color temperature. Therefore, this characteristic was used to complete color temperature control circuit. The experimental results show the variation in color temperature decreased from 31.58% to 11.72% when the temperature was from -30°C to 85°C . The modification rate was 19.86%. It is confirmed that color temperature is stable with the proposed color temperature control circuit.

Acknowledgment. The authors wish to thank the Nation Science Council, Taiwan, R.O.C. for supporting under Grant NSC 97-2120-M-260-001.

References

1. Chen, N.C., Yang, Y.K., Lien, W.C., Tseng, C.Y.: Forward current voltage characteristics of an AlGaInP light-emitting diodes. *J. Appl. Phys.* 102, 043706 (2007)
2. Zivojinovic, P., Lescure, M., Tap-Béteille, H.: Stability of feedback laser driver with average power control. In: *Proc. Conf. Design of Circuits and Integrated Systems*, Porto, Portugal, pp. 560–563 (November 2001)
3. Zivojinovic, P., Lescure, M., Tap-Béteille, H.: Design and stability analysis of a CMOS feedback laser driver. *IEEE Trans. Instrum. Meas.* 53(1), 102–108 (2004)

4. Lin, C.H., Yao, I.C., Kuo, C.C., Jou, S.J.: 2.5 Gbps CMOS laser diode driver with APC and digitally controlled current modulation. In: 2002 IEEE Asia-Pacific Conference, pp. 77–80 (2002)
5. Claisse, P.R., Jiang, W., Kiely, P.A., Roll, M., Boughter, L., Sanchez, P., Cotney, D., Lebby, M., Webb, B., Lawrence, B.: Automatic power control of a VCSEL using an angled lid TO56 package. In: 48th IEEE Electronic Components & Technology Conference, May 25-28, pp. 203–209 (1998)
6. Ku, J.C.S., Lee, C.J.: GB LED with smart control in the backlight & lighting. In: Proc. of SPIE, vol. 6894, p. 68940W (2008)

Study on Integrated Circuit Design for Multi-color LED Pulse-Width Modulation

Tai-Ping Sun, Chia-Hung Wang, Fu-Shiung Tsau, and Jia-Hao Li

Department of Electrical Engineering
National Chi Nan University
Nantou, Taiwan (R.O.C.)

{tps,s95323901,s96323523,s98327513}@ncnu.edu.tw

Abstract. This paper takes advantage of digital dimming auto-control pulse width drive Light-Emitting Diodes to let the light in terms of brightness fade to gray by degrees and have the effect of mixture of colors. Also, Light-Emitting Diodes in such three primary colors as red, green, and blue mix with one another to yield various colors. Therefore, better stability, simpler structure and better effect of dimming. Moreover, it succeeds in utilizing red, green, and blue Light-Emitting diodes at room temperature to change Color Rendering Index of the cool white Light-Emitting Diodes. The design of the circuit makes use of the manufacturing technique of Very Large Scale Integration to make the completion of Integrated Circuit (TSMC 0.35 μ m MIXED MODE 2P4M 3.3v process).

Keywords: Light-Emitting Diodes, Color Rendering Index, Cool White.

1 Introduction

In recent years, owing to the progress of process and material, light emitting diodes enhance the brightness and its application to daily life is available. However, the varying CT(Color Temperature) causes people to have different sensations. High color temperature can make people feel as cold, as northern skies. Low color temperature can make people feel as warm, as the dusk sunset. The differing of CRI(Color Rendering Index) results in the difference of color of objects and sensations. High CRI can be applied to precision machining, restaurants, galleries, etc. Low CRI can be applied to rough plant, general lighting place, etc.

This paper proposes a color mixing control system that provides adjustable color mixing using DPWM(Digital Pulse Width Modulation) to control red-green-blue (RGB) light-emitting diodes[1-3] respectively. 2⁹ kinds of different colors are produced in a cycle. These colors can change CT and CRI[4] of the original white light emitting diode. So it can adjust the CT and can enhance CRI of the white light. Such a simple circuit modulation method is primarily, of which the white light is lack, to produce the color and spectrum so as to change the characteristics of white light.

2 Proposed Circuit

This circuit design is primary the digital one which can automatically control the duty cycle of the output. Dimming circuit is to use the counting method and then with the decoder circuit. In so doing, the output signal would change its duty cycle from time of time. Circuit designs consist of 18 states, which are controlled by the counter. Timing is shown in Figure 1. Horizontal axis stands for the unit of time. On the vertical axis, when the RGB is on the low level, light-emitting diodes are bright; however, when it is on the high level, light-emitting diode is off. Rise time indicate the gradual darkness of light-emitting diode. Fall time indicate the gradual brightness of light-emitting diode.

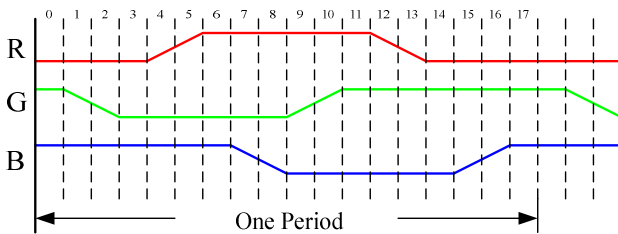


Fig. 1. The color timing diagram of one period

Figure 2 shows the overall circuit architecture diagram. The overall circuit function is described as the following.

- A. When power is turned on, light emitting diode would generate the output pattern of colorful the combination. 2^9 kinds of different colors would be yielded automatically in a cycle, as shown in Figure 1.
- B. If pushing the “KEY”, the color modulation stops functioning. However, pushing the “KEY” one more time, the color modulation continues to function.
- C. Switch S1~S8, current volume can be adjusted.

As shown in Figure 3, this circuit generates the mixed color counterclockwise, which begins with red, with regard to the CIE 1931 chromaticity diagram. The eighteen colors at the right side of Figure 3 are respectively counterclockwise illustration. We only take advantage of the circuit 17 kinds of colors among all mixed colors to adjust CRI and CT[5-6] of white light emitting diodes. White light emitting diode is driven by the constant current, which is shown in Figure 2.

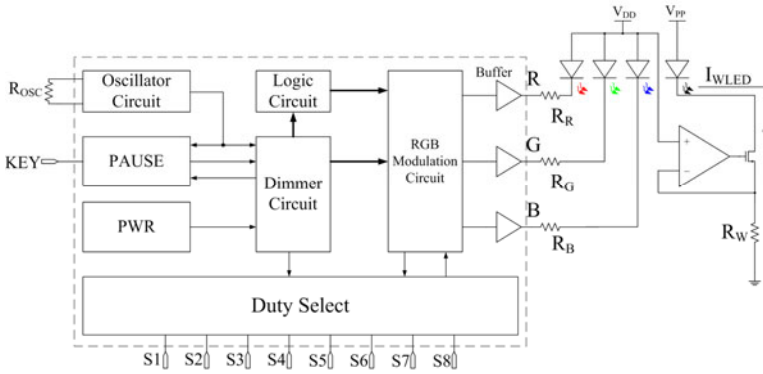


Fig. 2. The block diagram of color modulation

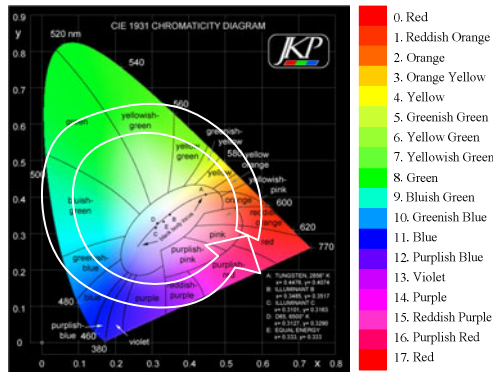


Fig. 3. The color distribution of CIE 1931

3 Experimental Results

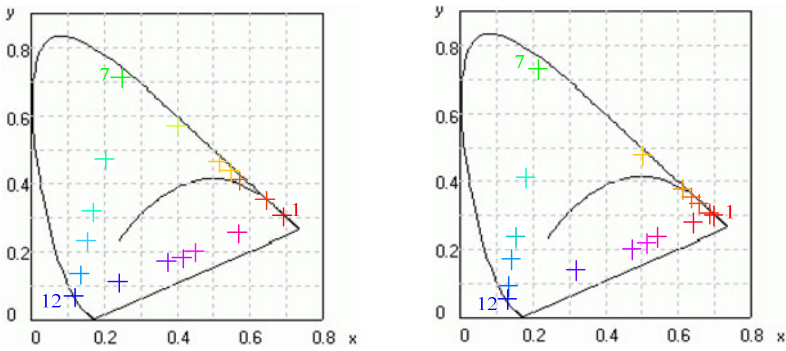
This paper mainly to use the circuit so as to make red, green and blue light-emitting diodes generate different colors mixed. Besides, these colors are complements to the spectrum of white light emitting diodes, enhance CRI, and change CT. White light-emitting diodes are composed of blue and yellow phosphor, and driven by 50mA constant current mode. At room temperature, the CRI is 72.1, and CT is 5513K °.

This paper uses 17 kinds of fixed colors to change the white light of CRI and CT. These colors are produced by 17 kinds of fixed pulse widths as shown in Table 1. The paper is to drive two different currents so to change the characteristics of white light emitting diode. First, the weak current mode is to drive red, green and blue light-emitting diode. Its maximum average currents are 3mA, 0.9mA, and 1mA respectively. Its result of color mixing is shown in Figure 4(a). The characteristic of

white light changed by each color is shown in Figure 5. Such a method improves the maximum CRI to 83.1 and CT to 5277K. The CIE1931 chromaticity diagram and spectral distribution are shown in Figure 6 (a) and Figure 6 (b) respectively. The color temperature ranges from 4977K to 5856K.

Table 1. Duty Cycle of the RGB

Color	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
R(%)	100	100	100	100	75	25	0	0	0	0	0	0	25	75	100	100	100
G(%)	0	25	75	100	100	100	100	100	100	75	25	0	0	0	0	0	0
B(%)	0	0	0	0	0	0	0	25	75	100	100	100	100	100	100	75	25



(a) Color distribution of small current drive (b) Color distribution of high current drive

Fig. 4. The CIE1931 Chromaticity Diagram

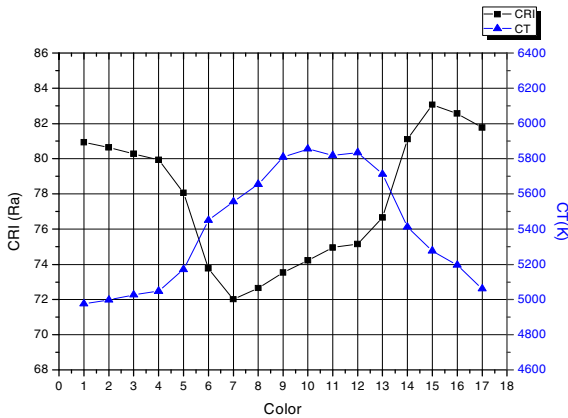


Fig. 5. CRI and CT curve of weak current drive

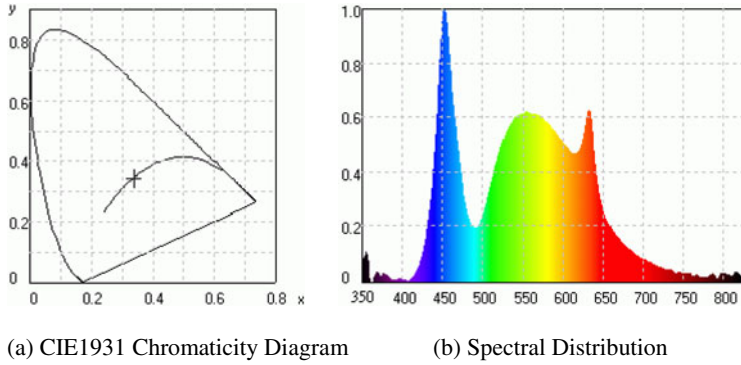


Fig. 6. The weak current drive RGB LEDs and White LED

Second, the strong current mode is to drive red, green and blue light-emitting diode. Its maximum average currents are 37.1mA, 4.7mA, and 8.2mA respectively. Its result of color mixing is shown in Figure 4(b). The characteristic of white light changed by each color is shown in Figure 7. Such a method improves the maximum CRI to 91.3 and CT to 5740K. The CIE1931 chromaticity diagram and spectral distribution are shown in Figure 8(a) and Figure 8(b) respectively. The color temperature ranges from 3310K to 6900K. Therefore, the method of the paper succeeds in changing the characteristics of white light emitting diode. The stronger the current of red, green, and blue light-emitting diode is, the larger extent to which the characteristics of white light are changed. The chip microphotograph in Figure 9 is for the block in dotted line in Figure 2.

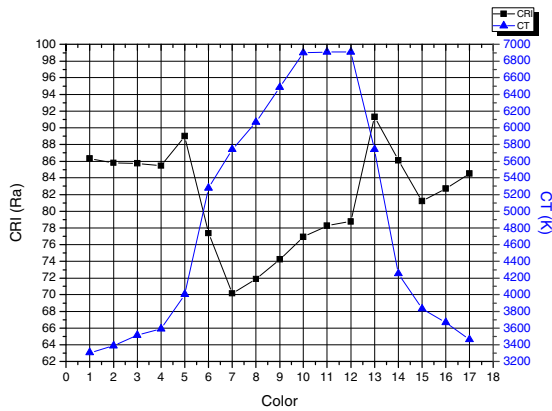


Fig. 7. CRI and CT curve of strong current drive

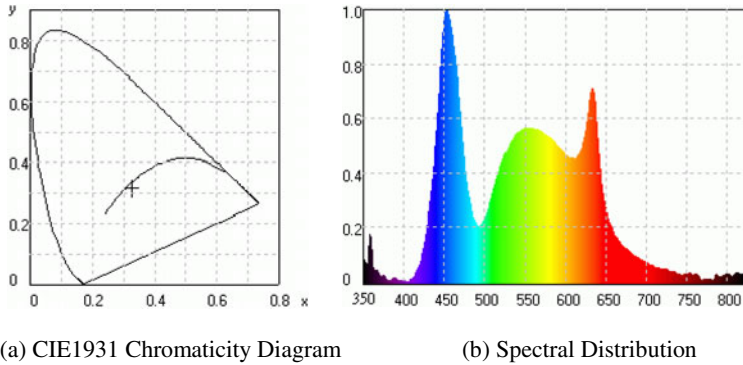


Fig. 8. The strong current drive RGB LEDs and White LED

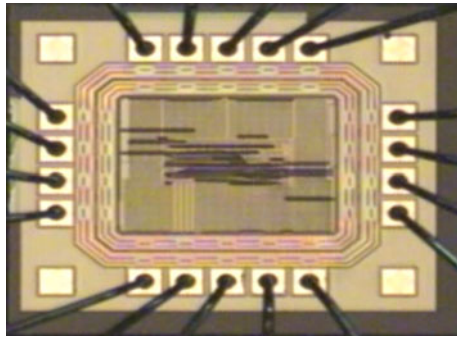


Fig. 9. Microphotograph

4 Conclusion

This paper uses a digital pulse width modulation to automatically dimming the red, green and blue light-emitting diode for the purpose of mixing multi-colors. These colors succeed in changing the characteristics of white light emitting diode. At room temperature, CRI of white light-emitting diodes is 72.1, the CT is 5513K. After these colors are complements to the spectrum of white light, CRI ranges from 70.2 to 91.3, and CT from 3310K to 6900K. The contribution of the paper is that such a technique can be used to change CRI and CT to the desirable ones in any places in the future.

Acknowledgment. The authors wish to thank the Nation Science Council, Taiwan, R.O.C. for supporting under Grant NSC 97-2120-M-260-001.

References

1. Wang, C.-H., Sun, T.-P.: A novel color mixing control system for light-emitting diodes. In: Proc. SPIE, vol. 7482, p. 74820K (2009)
2. Irvakili, A., Joyner, V.: A digitally-controlled, bi-level CMOS LED driver circuit combining PWM dimming and data transmission for visible light networks. In: 2010 IEEE GLOBECOM Workshops (GC Wkshps), December 6-10, pp. 1067–1071 (2010)
3. Muthu, S., Schuurmans, F.J., Pashley, M.D.: Red, green, and blue LED based white light generation: issues and control. In: 37th IAS Annual Meeting Conference Record of the Industry Applications Conference, vol. 1, pp. 327–333 (2002)
4. Ying, S.P., Tang, C.-W., Huang, B.-J.: Characterizing LEDs for Mixture of Colored LED light sources. In: International Conference on Electronic Materials and Packaging, EMAP 2006, December 11-14, pp. 1–5 (2006)
5. Shlayan, N., Venkat, R., Ginobbi, P., Mercier, G.: A Novel RGBW Pixel for LED Displays. In: 19th International Conference on Systems Engineering, ICSENG 2008, August 19-21, pp. 407–411 (2008)
6. Wu, H., Feng, Q., Lv, G.: A novel LCD backlight with mixed white and colorful LEDs. In: AFRICON 2009, September 23-25, pp. 1–5 (2009)

An Embedded Controller for Smart Products

Limin Liu

Institute of Embedded Systems
IT School, Huzhou University
Huzhou, Zhejiang, 313000, China
liulimin@ieee.org

Abstract. Embedded controllers are important cores for most smart products. A controller will manage various functions in an electronic product, such as mobile phone. In order to develop smart products, the embedded controller is asked the more powerful performance. In this paper, a common embedded controller for smart products is discussed.

Keywords: Embedded systems, embedded controller, electronic products, micro controller.

1 Introduction

Smart products are popular more and more, from space robots to mobile phones. Children like smart toys. People love smart mobile phone. Smart cards are used everyday. The smart products are intelligent electronic products. The smart ability is from some electronic parts. The electronic parts are depended on some embedded cores [1]. The core is a controller. There is one embedded controller in one smart product at least [2]. Some products take more controllers, such as a car. Fig.1 shows the relationship between smart products and embedded controller.

Actually all intelligent actions of smart products are produced by their inside controllers [3]. The controllers are composed of embedded cores and interfaces.

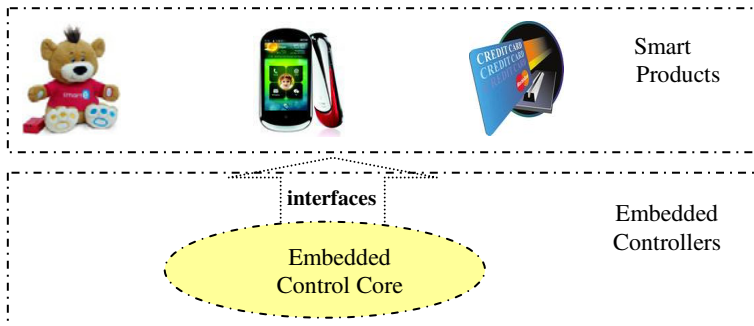


Fig. 1. Smart products and embedded controllers

On the other hand, embedded controllers are combination of software and hardware. Then, how is an embedded controller to act some smart products? The organization and functions of the controller should be known.

2 Organization of Embedded Controllers

Embedded Controllers normally are designed for special smart products or control objects. The word “embedded” means to put inside [4].

A common application structure of some embedded controller is indicated as Fig.2

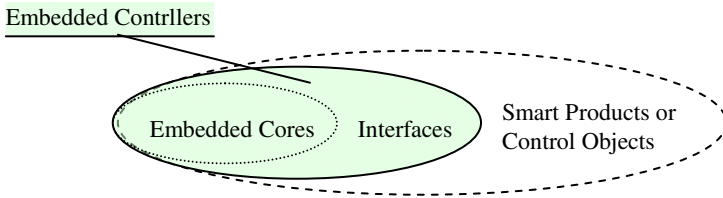


Fig. 2. An application structure of embedded controller

An embedded controller for smart products may consist of an embedded core, interfaces bus and memory [5]. Its organization is shown as Fig.3.

The embedded core is the controller center of the controller, like the brain. It should include one or more CPU units, such as MPU, microprocessor, DSP, digital signal processor and MCU, micro-controller.

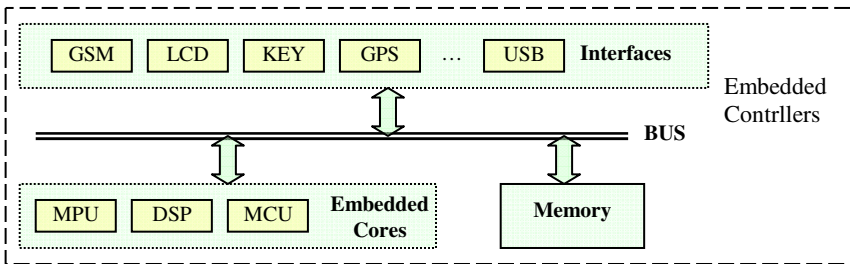


Fig. 3. Organization of an embedded controller

The core sends instructs to manage interfaces, then to drive peripherals. The codes are stored in memory, such as EEPROM or Flash memory. All data communication in the controller may be taken through some internal bus. Normally, there are three buses, control, address and data, in a controller.

With more interfaces, a controller can manage various peripherals. There are many different peripherals as controlled objects, such as mobile phones, LCD displays, keyboards, GPS navigation aid, printers and so on.

To control different objects, the various interfaces are required. For mobile phone, the GSM interface is necessary. The LCD, KEY, GPS and USB interfaces will be employed by LCD displays, keyboards, GPS navigation aid and printers.

3 Development of Embedded Controllers

Most embedded controllers are unique, and are developed specially. It is because that each controlled object has different features and various control requirements.

From hardware, there are two ways to design embedded controller. One is to use fixed logic devices, such as MPU, DSP and MCU units. Another way is to depend on PLD, programmable logic devices, such as FPGA or CPLD chips [6].

3.1 A Software and Hardware for Controller

The development of an embedded controller for smart products is a design procedure with combination of hardware and software [7]. Actually, the design of a controller depends on different operating demands. In order to match the demands, a special hardware and corresponding software are essential. The Fig.4 shows hardware and software for an embedded controller.

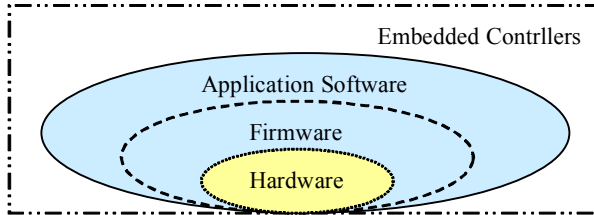


Fig. 4. Software and hardware of an embedded controller

The hardware is wrapped by firmware in a controller. The firmware is firmed software. Normally firmware directly drives hardware. For some large application, application software is programmed. Otherwise, for a common controller, the firmware is only software.

There are two characteristics for firmware. The first one is to face hardware. The other is for codes to be stored in firmed memory. The firmed memory mainly is semiconductor memory, such as EPROM, EEPROM and Flash memory. The magnetic and optical memory, like hard disks, floppy disks and CDs, are not firmed memory.

The firmware is required for embedded controllers. An important reason is user in moving environment for most smart products. The controllers are asked to match various use environments, especially anti-vibration and better reliability [8]. So, the semiconductor memory is only choice to store software of embedded controllers. And the firmware is only software for the similar situation.

In order to increase operation efficiency, C and Assembly are common computer languages. The port definition codes of MCU 8051 are as follows.

```

; Ports
;LCD control signals
    RS      EQU  P2.7  ;
    RW      EQU  P2.6  ;
    EP      EQU  P2.5  ;
;Beep run with '0'
    BEEP    EQU  P2.0  ;
; valves control
    V6_PIN  EQU  P1.5   ; port 1.5-1.0 for valve
control
    V5_PIN  EQU  P1.4
    V4_PIN  EQU  P1.3
    V3_PIN  EQU  P1.2
    V2_PIN  EQU  P1.1
    V1_PIN  EQU  P1.0
    V_ALL_PIN EQU      P1
; work state display
    LED_VOPEN_DISP EQU  P1.6   ; valve open state
display

```

From above program, a software designer must be familiar with hardware for define MCU port pins. The detailed pins represent different peripherals, such as read/write control lines, beep, valves and LED.

3.2 A Design Based on PLD

PLD, programmable logic device, is hardware chip. Its function will be designed by software. It is an ideal device for embedded controller application [9]. An embedded controller based on PLD is shown as Fig.5. In fact, the controller is a SoC, system on a chip [10].

There is in-memory, build in memory, or ex-memory, external memory in an embedded controller based on PLD. Application software in memory would be run through CC. A peripheral can be controlled by some IP or directly managed by CC in a controller. Here, all devices or equipments outside of controller are considered peripherals, such as executors in control systems, office machines connected computers, display and control components for users, communication devices and so on.

In Fig.5, CCs communicate with IPs, or external bus through internal bus. For a common controller, the internal bus is necessary [11]. According to applications, the external bus may be not required for some cases.

An obvious advantage of PLD controllers is secret. Since all functions of hardware are accomplished through software programming, the solution of this special controller is just known by its designer.

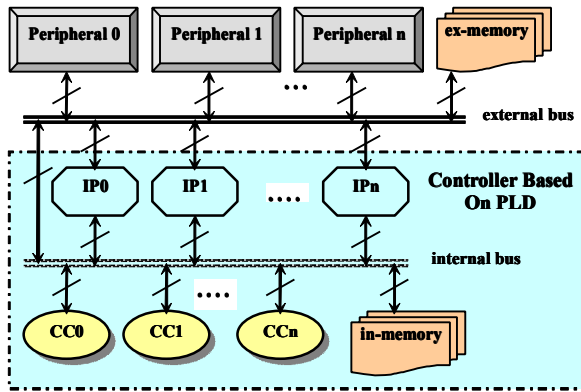


Fig. 5. An embedded controller based on PLD

When to program some PLD devices, the computer languages are hardware description languages, such as VHDL and Verilog. A code frame of VHDL for LCD controller is as follows.

```

Library
Entity lcd_fct is
  Port( );
End lcd_fct;
Architecture struct of lcd_fct is
  Signal
  Constant
Begin
Res: process;
Clk :clk_div;
LCD: process;
  Begin
    If init then
      Initialization;
    Eelsif clr then
      Clear LCD;
    Eelsif addr then
      Write address to LCD RAM;
    Eelsif data then
      Write data to LCD RAM;
    End if;
  End process;
Us: ustimer;
Ms: mstimer;
End struct;

```

4 Conclusions

Embedded controllers are important cores for most smart products. A controller is composed of embedded core, interfaces, bus and memory. Its hardware may be fixed or programmable logic devices.

The development of embedded controller is a combination procedure of software and hardware. Its software mainly is firmware. A controller based on PLD is a competitive solution.

Acknowledgments. This research was supported in part by the National Natural Science Foundation of China under grant 60872057, by Zhejiang Provincial Natural Science Foundation of China under grants R1090244, Y1101237, Y1110944 and Y1100095. We are grateful to NSFC, ZJNSF and Huzhou University.

References

1. Ostua, E., Viejo, J., et al.: Digital Data Processing Peripheral Design for an Embedded Application based on the Microblaze Soft Core. In: Proc. 4th South. Conf. Programmable Logic, San Carlos de Bariloche, Argentina, vol. 3, pp. 197–200 (2008)
2. Sifakis, J.: Embedded systems design - Scientific challenges and work directions. In: Proceedings of DATE 2009, vol. 4, p. 2. IEEE Press, Nice (2009)
3. Kay, D., Chung, S., Mourad, S.: Embedded test control schemes using iBIST for SOCs. *IEEE Transactions on Instrumentation and Measurement* 54(3), 956–964 (2005)
4. Wallner, S.: A configurable system-on-chip architecture for embedded and real-time applications: concepts, design and realization. *Elsevier Journal of Systems Architecture* 51, 350–367 (2005)
5. Liu, L., Luo, X.: The Reconfigurable IP Modules and Design. In: Proc. of EMEIT 2011, Harbin, China, vol. 8, pp. 1324–1327 (2011)
6. Kilts, S.: *Advanced FPGA Design: Architecture, Implementation, and Optimization*. Wiley, New Jersey (2007)
7. Dimond, R.G., Mencer, O., Luk, W.: Combining Instruction Coding and Scheduling to Optimize Energy in System-on-FPGA. In: Proceedings of 14th Annual IEEE Symposium on Field-Programmable Custom Computing Machines, Napa, CA, USA, vol. 4, pp. 175–184 (2006)
8. Liu, L.: A Hardware and Software Cooperative Design of SoC IP. In: Proc. of CCIE 2010, Wuhan, China, vol. 6, pp. 77–80 (2010)
9. Saleh, R., Wilton, S., et al.: System-on-chip: reuse and integration. *Proceedings of the IEEE* 94(6), 1050–1069 (2006)
10. Claasen, T.M.: An industry perspective on current and future state of the art in system-on-chip (SoC) technology. *Proceedings of the IEEE* 94(6), 1121–1137 (2006)
11. Smit, G., Schuler, E., Becker, J., Quevremont, J., Brugger, W.: Overview of the 4S Project. In: Proceedings of 2005 International Symposium on System-on-Chip, Tampere, Finland, vol. 11, pp. 70–73 (2005)

Fuzzy Comprehensive Evaluation on the Quality of Vehicle Equipment Maintenance Based on Membership Degree Transformation Algorithm*

XiaoLing Wei^{1,**}, ZhiWen Chen¹, and ZhenZhong Zhou²

¹ College of Mechanical and Electrical Engineering, Hebei University of Engineering,
Handan 056038, China

² Handan Municipal Wastewater Treatment Co., Ltd.,
Handan 056002, China
wxl.csm@163.com

Abstract. In this paper, we were through digging the information about target classification implicated in the index membership and defined index distinguish right, to eliminate the redundancy index membership which has no effect on the target classification and redundant data of index membership, to pick up the effective value which has effect on the target classification to calculate membership, finally to build a new membership degree transformation algorithm. The new algorithm, is used in quality fuzzy comprehensive evaluation of vehicle maintenance, can make results of evaluation reliable and authentic.

Keywords: The quality of vehicle equipment maintenance, Fuzzy comprehensive evaluation, Membership degree transformation.

0 Introduction

Vehicle and equipment maintenance quality evaluation system is a complex system, it contains many complex uncertainty, the fuzzy evaluation model is a suitable methodology of evaluation. The essence of the fuzzy comprehensive evaluation is membership degree transformation. However, in the conversion process, in which membership degree is transformed from the indicators to the target, the existing conversion methods can not reveal which part of the target classification useful, in the indexing membership degree, result the useless redundancy value of the indexing membership degree for the target classification is used to calculate the membership degree of target. According to the redundancy value of the membership degree conversion process, in this paper, it will build a conversion membership degree model

* Fund: Funded by Natural Science Foundation of Hebei Province (F2010001047).

** Born in Shanxi Province Hejin City, professor, tutor of masters, research in design and theory of mechanical.

without interference of the redundancy value, this model is used for fuzzy comprehensive evaluation of vehicle maintenance quality to improve accuracy and efficiency of quality evaluation of vehicle equipment maintenance.

1 The New Membership Degree Transformation Model Based on M(1,2,3)

1.1 Defining the Index Distinguishing Right

First, from the classification point of view, we behold which index membership has effect on the target Q classification.

- (1) If $\mu_{j1}(Q) = \mu_{j2}(Q) = \dots = \mu_{jp}(Q)$,. If we us real number $\alpha_j(Q)$ said normalized quantitative value of the contribution of index j for target Q classification, then $\alpha_j(Q) = 0$.
- (2) If there is an integer k , at this time, the real number $\alpha_j(Q)$ should be maximum. (3) Similarly, if the closer the values of membership $\mu_{jk}(Q)$ to k , the greater of $\alpha_j(Q)$ •on the contrary, the smaller of $\alpha_j(Q)$.

The above three conditions describe, the real number, that reflects the magnitude of the contribution of index j to target $\alpha_j(Q)$, depends on the degree of centralization and decentralization of the value of membership $\mu_{jk}(Q)$;while, the value of membership $\mu_{jk}(Q)$ can be described in the number by the entropy $H_j(Q)$ of membership:

$$H_j(Q) = - \sum_{k=1}^p \mu_{jk}(Q) \cdot \log \mu_{jk}(Q) \tag{3}$$

$$v_j(Q) = 1 - \frac{1}{\log p} H_j(Q) \tag{4}$$

$$\alpha_j(Q) = v_j(Q) / \sum_{t=1}^m v_t(Q) \quad (j = 1 \sim m) \tag{5}$$

The real number $\alpha_j(Q)$ is defined, from formulas (3), (4)and(5), as the distinguishing right of the index j to the target Q , and the distinguishing right $\alpha_j(Q)$ should meet the condition:

$$0 \leq \alpha_j(Q) \leq 1 \quad \sum_{j=1}^m \alpha_j(Q) = 1 \tag{6}$$

1.2 Determining the Effective Value

1: If $\mu_{jk}(Q)$ ($k = 1 \sim p, j = 1 \sim m$) is the membership of index j of target Q that belongs to the kind C_k , and $\mu_{jk}(Q)$ meets the formula (1); $\alpha_j(Q)$ is the distinguishing right of index j about target Q , we name:

$$\alpha_j(Q) \cdot \mu_{jk}(Q) \quad (k = 1 \sim p) \tag{7}$$

is the effective distinguishing value of the kind k membership of index j , named the kind k effective value for short.

1.3 Defining the Comparable Value

2: If $\alpha_j(Q) \cdot \mu_{jk}(Q)$ is the kind k effective value of index j , $\beta_j(Q)$ the important weight of index j about target Q , we name:

$$\beta_j(Q) \cdot \alpha_j(Q) \cdot \mu_{jk}(Q) \quad (k = 1 \sim p) \tag{8}$$

is comparable effective value of the kind k membership of the index j , named the kind k comparable value for short.

3: If $\beta_j(Q) \cdot \alpha_j(Q) \cdot \mu_{jk}(Q)$ is the kind k comparable value of index j ($j = 1 \sim m$) of target Q , we name:

$$M_k(Q) = \sum_{j=1}^m \beta_j(Q) \cdot \alpha_j(Q) \cdot \mu_{jk}(Q) \quad (k = 1 \sim p) \tag{9}$$

Is the kind k comparable sum of target Q .

4: If $M_k(Q)$ is the kind k comparable sum of target Q , $\mu_k(Q)$ is the membership of Q belongs the kind C_k , then:

$$\mu_k(Q) = M_k(Q) / \sum_{t=1}^p M_t(Q) \quad (k = 1 \sim p) \tag{10}$$

So far, by the above formulas, we get the conversion from each index membership to the target membership; and in this process, the priori knowledge won't be increased or the classification information won't be lost.

2 An Application Example

In this paper, we use the system, the weight and the membership vector of quality evaluation of vehicle equipment maintenance from the literature[1], shown in form 1. In

the form 1, the vector of the bottom index is the membership of the bottom index about 5 comments levels (Best C1, Better C2, Normal C3, Worse C4, Worst C5), we divide the level of quality of vehicle equipment maintenance into 5 level:best, better, normal, worse, worst.

Form 1, The system and fuzzy membership of quality evaluation of vehicle equipment maintenance

Target layer	Level 1 index	Level 2 index	Membership Vector (C ₁ , C ₂ , C ₃ , C ₄ , C ₅)
Quality Comprehen-siv e Evaluation of vehicle equipment maintenanc-e Q	A ₁ : Engine system (0.323)	Power (0.15)	(0.0, 0.8, 0.2, 0.0, 0.0)
		Revolving speed (0.20)	(0.0, 0.0, 0.2, 0.8, 0.0)
		Oil pressure (0.10)	(0.0, 0.8, 0.2, 0.0, 0.0)
		Injection pressure (0.15)	(0.0, 0.0, 0.2, 0.8, 0.0)
		Injection advance angle (0.15)	(0.0, 0.8, 0.2, 0.0, 0.0)
		Engine noise (0.10)	(0.0, 0.0, 0.2, 0.8, 0.0)
		Engine emission (0.15)	(0.2, 0.2, 0.6, 0.2, 0.0)
	A ₂ : Electrical system (0.181)	Regulator (0.35)	(1.0, 0.0, 0.0, 0.0, 0.0)
		starting motor (0.35)	(0.0, 0.8, 0.2, 0.0, 0.0)
		Aid electrical system (0.30)	(0.0, 0.8, 0.2, 0.0, 0.0)
	A ₃ : Transmissio n system (0.114)	Transmission (0.30)	(0.0, 0.2, 0.6, 0.2, 0.0)
		Cluth (0.25)	(0.0, 0.0, 0.2, 0.8, 0.0)
		Transmission shaft (0.25)	(0.0, 0.2, 0.6, 0.2, 0.0)
		Driving axle (0.20)	(0.0, 0.0, 0.0, 0.0, 1.0)
	A ₄ : Steering system (0.068)	Floating steering wheel gap (0.30)	(0.0, 0.8, 0.2, 0.0, 0.0)
		Minimum turning radius (0.20)	(0.0, 0.2, 0.6, 0.2, 0.0)
		Steering torque (0.25)	(0.0, 0.2, 0.6, 0.2, 0.0)
		Maximum Angle of steering (0.25)	(0.0, 0.0, 0.0, 0.0, 1.0)
	A ₅ : Braking system (0.049)	Brake pressure (0.35)	(1.0, 0.0, 0.0, 0.0, 0.0)
		Brake pedal free travel (0.30)	(0.0, 0.8, 0.2, 0.0, 0.0)
		Braking distance (0.35)	(0.0, 0.2, 0.6, 0.2, 0.0)
A ₆ : Driving system (0.163)	Shock absorber (0.30)	(0.0, 0.0, 0.2, 0.8, 0.0)	
	Tires (0.35)	(0.0, 0.2, 0.6, 0.2, 0.0)	
	Driving wandering (0.35)	(0.0, 0.2, 0.6, 0.2, 0.0)	
A ₇ : Performanc-e condition (0.103)	Vehicle appearance (0.35)	(0.0, 0.2, 0.6, 0.2, 0.0)	
	Acceleration performance (0.35)	(0.0, 0.8, 0.2, 0.1, 0.0)	
	Fuel consumption per hundred kilometers (0.30)	(0.0, 0.2, 0.6, 0.2, 0.0)	

(1)According to the evaluation matrix of form 1, for example in electrical system A_2 , the steps of calculating synthesis membership is that:

① The A_2 has 3 bottom indexes $B_{21} \sim B_{23}$ •the index membership matrix is that:

$$U(A_2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.8 & 0.2 & 0 & 0 \\ 0 & 0.8 & 0.2 & 0 & 0 \end{bmatrix}$$

We use the $j(j = 1 \sim 3)$ line of $U(A_2)$ to calculate the distinguishing right of index B_j , the vector of distinguishing right is $\alpha(A_2) = (0.4204, 0.2898, 0.2898)$

② According to the form 1•the importance weight vector of the index $B_{21} \sim B_{24}$ about A_2 is $\beta(A_2) = (0.35, 0.35, 0.30)$

③ We calculate the kind k comparable value of the index B_{2j} , and get the comparable value matrix of A_2 is:

$$N(A_2) = \begin{bmatrix} 0.1471 & 0 & 0 & 0 & 0 \\ 0 & 0.0811 & 0.0203 & 0 & 0 \\ 0 & 0.0696 & 0.0174 & 0 & 0 \end{bmatrix}$$

④ We calculate the kind k comparable sum of the index A_2 , and get the comparable sum vector of A_2 is $M(A_2) = (0.1471, 0.1507, 0.0377, 0, 0)$

⑤ According to the comparable sum vector $M(A_2)$ •we calculate the membership vector $\mu(A_2)$ of A_2 :

$$\mu(A_2) = (0.4385, 0.4492, 0.1123, 0, 0)$$

Similarly, we can get $\mu(A_1) \cdot \mu(A_3) \cdot \mu(A_4) \cdot \mu(A_5) \cdot \mu(A_6)$ and $\mu(A_7)$ •along with $\mu(A_2)$, will get the matrix $U(Q)$ of quality evaluation of vehicle equipment maintenance:

$$U(Q) = \begin{bmatrix} \mu(A_1) \\ \mu(A_2) \\ \mu(A_3) \\ \mu(A_4) \\ \mu(A_5) \\ \mu(A_6) \\ \mu(A_7) \end{bmatrix} = \begin{bmatrix} 0 & 0.3601 & 0.2370 & 0.4029 & 0 & \\ 0.4385 & 0.4492 & 0.1123 & 0 & 0 & \\ 0 & 0.0756 & 0.2838 & 0.3060 & 0.3346 & \\ 0 & 0.3153 & 0.2371 & 0.0575 & 0.3901 & \\ 0.4998 & 0.2772 & 0.1819 & 0.0410 & 0 & \\ 0 & 0.1162 & 0.4324 & 0.4514 & 0 & \\ 0 & 0.4852 & 0.4099 & 0.1049 & 0 & \end{bmatrix}$$

(2) The $U(Q)$ is calculated by the same step as Step•, the membership vector of quality evaluation of vehicle equipment maintenance is:

$$\mu(Q) = (\mu_1(Q), \dots, \mu_7(Q)) = (0.1141, 0.3235, 0.2707, 0.2515, 0.0407)$$

(3) Recognition

Set $\lambda(\lambda > 0.6)$ into confidence level,, calculating

$$K_0 = \min \left\{ k \left| \sum_{t=1}^k \mu_t(Q) \geq \lambda, 1 \leq k \leq 5 \right. \right\}$$

Then, we judge that the Q belongs to the level K_0 , and is Greater than the confidence level of λ . In this example, the Q belongs to level “Normal”, the confidence level is about $70.78\% \cdot 0.1141 + 0.3235 + 0.2702 = 0.7078$.

3 Conclusion

In this paper, we build a membership conversion model which re not interference from redundant data, it can eliminate the redundant data in the membership conversion process, and is used in quality comprehensive evaluation of vehicle equipment maintenance. We can improve the Reliability and security of quality of vehicle equipment maintenance.

References

1. Liu, Z.Y., Zhang, Y., Ma, H.W.: Analysis of Vehicle Maintenance Quality Based on the Fuzzy Synthesis Appraisal. *Journal of Academy of Military Transportation* 10(6), 31–48 (2008)
2. Cao, Q.-K., Ruan, J.-H., Liu, K.-D.: Fuzzy evaluation on the safety of coal gas station based on a new algorithm of membership degree transformation. *Journal of China Coal Society* 35(3), 467–470 (2010)
3. Xie, Q.-H., Zhang, Q.: Model of Fuzzy Comprehensive Evaluation of Engineering Equipment Repair Quality. *Journal of PLA University of Science and Technology* 4(1), 63–66 (2003)

Moving Vehicle Detection Based on Union of Three-Frame Difference

Caiyuan Chen and Xiaoning Zhang

Key Lab of Road and Traffic Engineering at the Ministry of Education
Tongji University Shanghai 200092, China
ccyuanxx@126.com

Abstract. Due to problems of blank holes created by traditional frames difference in moving object detection, this paper develop a method of moving vehicle detection, which is based on three-frame-difference and union set. First, we difference the three grayed images and enhance them. Second, dichotomize images and filter some interference. Third, make the union set operation. Finally, detect the vehicle with morphology process and connectivity analysis. The results of experiment show that this method is characterized by small calculating amount, fast speed, accurate detection and a good robustness.

Keywords: vehicle detection, three-frame difference, images union, morphology.

1 Introduction

In recent years, vehicle detection has become a major concern in the field of computer vision. It takes video image as the main object, captures the video stream image, extracts traffic information, and finally detects vehicles.

Video detection in intelligent transport system (ITS) has been more widely used, which plays an important role in ITS. However, the dynamic change of background image caused by weather, light, and shadow makes it difficult to detect and track vehicles in video. So the segmentation of moving objects, lighting changes, and several vehicles adhesions brought some challenges in video detection.

2 Two-Frame Difference

Two-frame difference [1] is based on relevance of a series of images. The basic idea is to subtract between two adjacent images, then filter the stationary background by the method of threshold and extract moving objects.

At first, calculate the difference between the k^{th} frame image $F_k(x,y)$ and the $k+1^{\text{th}}$ frame image $F_{k+1}(x,y)$, obtain the difference image $D_k(x,y)$:

$$D_k(x,y) = | F_k(x,y) - F_{k+1}(x,y) | \quad (1)$$

Then, set a threshold value T for binary operation on the difference image $D_k(x,y)$ and obtain the binary image $I_k(x,y)$:

$$I_k(x, y) = \begin{cases} 1, & D_k(x, y) > T; \\ 0, & D_k(x, y) \leq T \end{cases} \quad (2)$$

When value of pixel in the difference images exceeds the given threshold value T , the pixel is considered as an object pixel; otherwise as a background pixel. And then, go on to connectivity processing on the binary image $I_k(x,y)$. While the size of connective area exceeds a given threshold value, it is considered as a moving area; otherwise it is considered as a background area.

There are some disadvantages in this method [2]. First, vehicle usually is missed when the speed is extremely low. Secondly, if the inner gray of the object is relatively homogeneous, the overlap part in moving maybe turns to be blank holes. In actual application, some scholars improved the two-frame difference by differencing three adjacent images and then making logical operation [3]. And some proposed to take an image block as a difference unit and so on [4]. As a whole, those methods have optimized the results in moving object detection.

3 Three-Frame Difference

To solve the problems in two-frame difference for low speed vehicles or homogeneous color vehicles, this paper provides an algorithm of union of three-frame difference to accumulate motion trajectory to ensure that the moving vehicles can be detected.

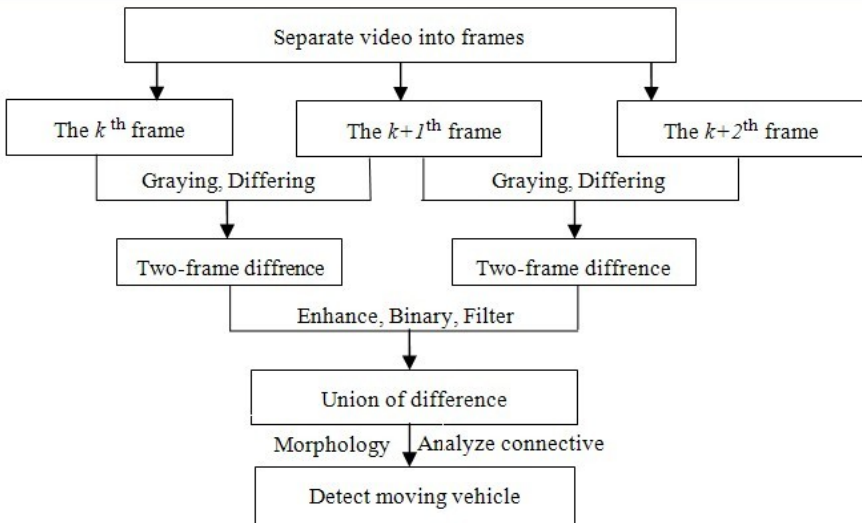


Fig. 1. Flow Chart of Three-Frame Difference

3.1 Image Graying

The first step in video image processing is to separate video into individual frame images. The frame rate of camera usually is 24 or 30 frame/second. Here we adopt 30 frame/second as rate. The original frame images obtained from the video are color images. But variety of color in the color image means abundant information included. It needs large memory capacity to storage as well as takes huge system resources to operate, which goes against the fast processing. So the next step of image processing is to gray color image.

Gray image contains only the brightness information without color information. The brightness that is usually divided into 256 (0-255), where 0 is darkest (black), 255 is the brightest (white). Weighted average method is commonly used in gray image conversion, which is to give R, G, B different weights according to importance or other indicators.

$$\text{Gray}(x,y) = (w_r * R(x,y) + w_g * G(x,y) + w_b * B(x,y)) \quad (3)$$

w_r , w_g , w_b are respectively the weight of R, G, B. The human eye is highly sensitive to green, red followed by and blue minimum. So we take $w_r = 0.2989$, $w_g = 0.5870$, $w_b = 0.1140$ in the weighted average method and thus get the most reasonable gray image.

3.2 Three-Frame Difference

The adjacent frames are different and the difference mainly due to the movement of vehicles. According to the theory, we difference three adjacent frames twice, the first time for the first frame and the second frame and the second time for the second frame and the third frame. That is to say we subtract the k^{th} gray frame $F_k(x,y)$ from the $k+1^{\text{th}}$ frame $F_{k+1}(x,y)$ to get the difference image $D_{k+1,k}(x,y)$ in formula(4). And then, subtract the $k+1^{\text{th}}$ gray frame $F_{k+1}(x,y)$ from the $k+2^{\text{th}}$ frame $F_{k+2}(x,y)$ to get the difference image $D_{k+2,k+1}(x,y)$ in formula(5):

$$D_{k+1,k}(x,y) = | F_{k+1}(x,y) - F_k(x,y) | \quad (4)$$

$$D_{k+2,k+1}(x,y) = | F_{k+2}(x,y) - F_{k+1}(x,y) | \quad (5)$$

3.3 Enhancing Contrast of Difference Images

Due to the effect of light or underexposure, the contrast of image is usually low, which leads to the details of moving object looks indistinct and then makes it difficult to identify and detect vehicles. Therefore, it is necessary to expand the scale of image gray and enhance the contrast, so as to improve the efficient. Here we adopted algorithm of adaptive histogram equilibrium for image enhancement.

Adaptive histogram equilibrium is to set sliding window for every pixel in the same size. Equalize local histogram based on sliding window in order to process the center pixel value of the window. Adaptive histogram equalization can find gray

transformation function in each sub-region, rather than the entire image, which ensure all the details of sub-region have been enhanced in image. So the performance of adaptive histogram equalization makes the details of regions more adequate.



Fig. 2. Three Gray Images from the 136th, 137th and 138th Frame

3.4 Dichotomizing Images

The point of dichotomizing images is to determine the threshold value T . Threshold selection method contains three categories: binary into global, local and dynamic. Here a method of global threshold is adopted for difference image. After differing frames, most of things left in the difference images are moving objects. At the result of experiments, 25 are selected as the threshold value T with adaptability to some extent.

3.5 Filtering Images

There are some inevitable isolated points in the two binary images, so it is necessary to filter those points as to get the image with less interference. Here method of median filter [4] is adopted. Median filter is a nonlinear filter. Its basic idea is to use median gray value of the neighborhood to instead the gray value of the pixel. Specifically speaking, that is to set a sliding window with $N \times N$ points to scan the image, each point of the window is instated by the median value of the window. The pixels in the window are sorted according to the gray value. If $N \times N$ is odd, then takes the middle value; If $N \times N$ is even, and then takes the average of the two middle values.

As a result of various experiments, 3×3 was fixed as the square of median filter was obtained. According to the results, the isolated points are reduced obviously, What's more, the images become smoother with little effect on vehicles, such as Fig.3 (a). That is beneficial for the next steps.

3.6 Uniting the Two Filtering Images

Because the images have been filtered, some information on moving vehicles has been lost. Hence the union of the two images is been considered as to make up some filtered information.

The principle of uniting the two filtering images as follows: as long as there is moving object in one difference image $D_{k+1,k}(x,y)$ or in the other one $D_{k+2,k+1}(x,y)$, then considered that there is moving object in union of the images $Union(x,y)$. Only there is

no moving objects in both difference images, then considered that there is no object in union of the images. The function of computing $Union(x,y)$ is in formula(6). Some necessary and useful vehicle information is made up after uniting, which is helpful for the next detection.

$$Union(x,y) = \begin{cases} 1, D_{k+1,k}(x,y) = 1 \text{ or } D_{k+2,k+1}(x,y) = 1; \\ 0, D_{k+1,k}(x,y) = 0 \text{ and } D_{k+2,k+1}(x,y) = 0 \end{cases} \quad (6)$$

Because we are only interested in the vehicles on the road, there is no need to analyze and process other information. Picking up road image from the whole image can effectively avoid the disturbance which is outside of road. According to actual video image, we pick up the part of road and make sure that contains sufficient information for detection, such as Fig.3 (b). Here we avoid the disturbance of the vehicles on the other road and the shaking trees on the edge of road.

3.7 Applying Morphology

In practical processing, even images have been united, there still exists some blank holes and cracks in the union of two images. That will bring some adverse effect on connectivity analysis. Hence it is necessary to apply morphology [5] on uniting image.

At first, operate open morphology with a disk structure so as to get rid of the isolated pixels. Then operate close morphology with a square structure so as to connect the whole vehicle region without interference among different vehicles. The result is showed in Fig.3 (c).

3.8 Calculating Numbers of Moving Vehicle

Label of connective components is based on 8 connected image blocks [6, 7] in analysis of connectivity. 8 connected image blocks means detecting the connectivity according to horizontal, vertical and four diagonal directions. It is assumed that the noise part is far less than the real moving part. Thus set S as a threshold value, if the size of connective component exceeds S , then it is considered as moving vehicle, if not, considered as background. Finally add up the totally numbers of moving vehicles. Based on the image of finally processing result, we detect the number of moving vehicle is 9 at that moment.

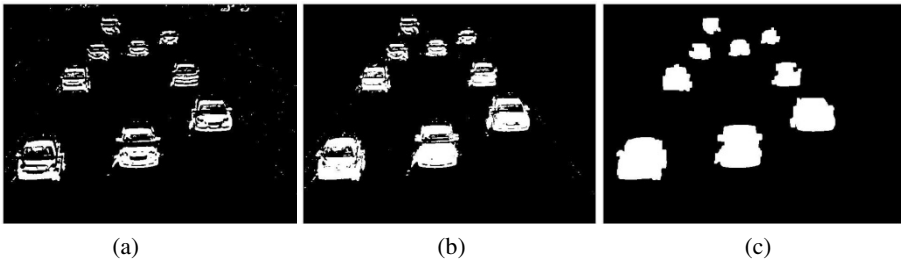


Fig. 3. (a) Images after filtering; (b) union image of the two filtering images; (c) image of final processing result

4 Conclusion

The results of experiments show that the three-frame difference is simple in calculation as well as accurate and fast in image segmentation. It can solve the problems exist in two-frame difference, which is difficult to detect vehicles in slow speed or with homogenous color. Due to calculate the totally numbers of moving vehicles with strong robustness, it can be used to volume detecting. However, we should point out that the edge of moving vehicle may become expansive, which needs to be resolved in further study.

References

1. Liu, W.X.: Vehicle Detection Based on Video Image Processing. Ms D Thesis. Shen Yang University of Technology, Shen Yang (2007)
2. Yin, M., Zhang, H., Meng, H.: An HMM-Based Algorithm for Vehicle Detection in Congested Traffic Situations. In: Proceedings of the 2007 IEEE Intelligent Transportation Systems Conference Seattle, WA, USA, September 30-October 3 (2007)
3. Lv, G., Zhao, S., Zhao, J.: A New Method of Object Detection Based on Three-Frame Difference and Connectivity Analyses. *Liquid Crystals and Displays* 22 (2007)
4. Gan, M., Chen, J., Liu, J., Wang, Y.: Based on Three-Frame Differencing and Edge Information. *Journal of Electronics & Information Technology*, 32 (2010)
5. Xie, F., Zhao, D.: *Visual C++ Digital Image Processing*, pp. 58–437. Publishing House of Electronics Industry, Beijing (2008)
6. Song, J.: Study of Traffic Flow Detection and Statistics Technology video-based. Chang An University (2010)
7. Du, L.-L.: Research on Video-based Vehicle Detection in ITS. Ocean University of China (2009)

Improving the Learning Performance by Providing Intellectual Suggestions

Mohammad Sarem¹, Mostafa Bellafkih², and Mohammad Ramdani¹

¹ Faculty of Sciences and Technique, Mohammadia, Morocco

² National Institute of posts and Telecommunications, Rabat, Morocco
mohsarem@gmail.com

Abstract. In developing a diagnosis system for detecting student's learning problems, providing intellectual suggestions is helpful and powerful tool to improve the learning performance. However, it is difficult and time consuming for teachers to give personalized suggestions to each student, particularly when there are many students in class. To cope with this problem, this study proposes concept-effect relationship model (CER) as a tool to identify the learning problem of students. Based on that, an intellectual learning guidance system has been proposed.

Keywords: Concept map, Fussy set theory, Bayesian network, Concept-effect relationship model, learning guidance.

1 Introduction

With vigorous development of the Internet, many scholars have attempted to adopt computer and communication technologies for research on education to support students during their learning performance on the Internet [11], in pursuit of this goal, many systems have been developed. Notable examples include the computer-assisted tutoring and testing systems [5, 10, and 14] and educational diagnosis and guidance systems [8, 13, and 18], specifically, the agent systems for guiding students through on-line course material [3 and 12] and learning feedback during on-line evaluation [16] which provide learning suggestions for students after testing. In recent years, researchers have proposed different approaches for developing adaptive learning systems based on learning behaviors of students during their interaction with e-learning systems [15]. As known, to improve students' learning performance, the teacher should give them additional suggestions. However, it is time consuming for him to give personalized suggestions to each student, particularly when there are many students in class. Among the existing models, Concept-effect relationship (CER) has been proved to be an effective way of improving the learning performance of student [11]. The CER model can be used for identifying the learning problem of students, furthermore, takes into account the effect of each concept in the learning process. As result of that, it has been used to successfully give personalized suggestions to them for several scientific courses [6, 7]. In this study, we propose an innovation learning strategy based on CER model and fuzzy set theory that can be used for providing intellectual guidance via e-learning environments.

2 Learning Guidance Strategy

2.1 Concept-Effect Model

In the CER model, the diagnosis of student learning problems mainly depends on the prerequisite relationships between the concepts to be learned. During tutoring, students learn new concepts among previously learned concepts, from this standpoint, the effect of learning one concept on the learning of other concepts called “Concept-effect relationships”.

Consider two concepts to be learned, say C_{i+1} and C_{i+2} (see Fig.1). If C_{i+1} is a prerequisite to efficiently performing the more complex and higher level concept C_{i+2} , then a concept-effect relationship " $C_{i+1} \rightarrow C_{i+2}$ " is said exist with effect " $W_{C_{i+1},C_{i+2}}$ ", where the " $W_{C_{i+1},C_{i+2}}$ " is the relevance degree between C_{i+1} and C_{i+2} (see section 2.2) . From this standpoint, if a student fails to answer most of the test items concerning C_{i+2} due to a lack of understanding of the questions posed or because of carelessness, the problem is likely because the student has not thoroughly learned " C_{i+2} " or its prerequisite concepts " C_{i-1} " or " C_{i+1} ". Therefore, teachers could identify the learning problems of students by tracing the concept-effect relationships.

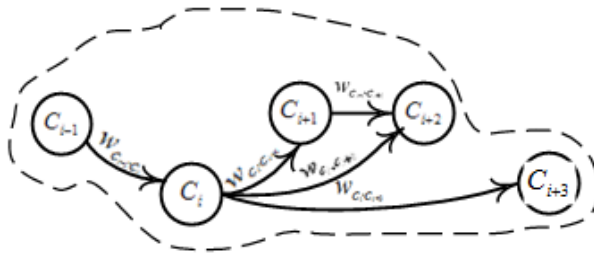


Fig. 1. Illustrative example of concept maps

In the CER model, all of the possible learning paths will be taken into consideration to find the poorly-learning paths. In the illustrative example given in Fig.1, to learn concept " C_{i+2} ", there are two learning paths for the subject unit:

PATH 1: $C_{i-1} \rightarrow C_i \rightarrow C_{i+1} \rightarrow C_{i+2}$

PATH 2: $C_{i-1} \rightarrow C_i \rightarrow C_{i+2}$

A threshold, θ , is used to indicate the acceptable error rate. If the ratio of incorrect answers to the total strength of concept is less than the proposed threshold, the student is said to have learned concept; otherwise, the student is said to have failed to learn concept and thus the concept is added to the To-Be-Enhanced learning path [11]. Among the To-Be-Enhanced learning paths, those with the maximum weight are defined as the critical learning paths that students are asked to restudy it in more detail before learning other concepts. In the next section, we propose a probabilistic approach to calculate the error ratio for each student to answer the items related to each concept.

2.2 Probability Approach for Calculating the Error Ratio

Let the test portfolio of the learners and the conceptual weight relationships into the matrix G and the matrix QC, where matrix “G” given as follows:

$$G = \begin{matrix} & S_1 & S_2 & \dots & S_n \\ \begin{matrix} Q_1 \\ Q_2 \\ \vdots \\ Q_m \end{matrix} & \begin{bmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \dots & g_{mn} \end{bmatrix} & \end{matrix}_{n \times m},$$

where “ g_{S_i, Q_j} ” denotes the score of question “ Q_j ” of the learner “ S_i ”, $g_{S_i, Q_j} = 1$ denotes the student S_i gets the right answer in question Q_j , and $g_{S_i, Q_j} = 0$ denotes the student S_i has a wrong answer in question Q_j , $1 \leq i \leq n$ and $1 \leq j \leq m$, “n” is number of learners and “m” is number of questions. Furthermore, let the questions-concepts matrix shown as follows:

$$QC = \begin{matrix} & C_1 & C_2 & \dots & C_p \\ \begin{matrix} Q_1 \\ Q_2 \\ \vdots \\ Q_m \end{matrix} & \begin{bmatrix} qc_{11} & qc_{12} & \dots & qc_{1p} \\ qc_{21} & qc_{22} & \dots & qc_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ qc_{m1} & qc_{m2} & \dots & qc_{mp} \end{bmatrix} & \end{matrix}_{p \times m},$$

where “ qc_{ij} ” denotes the weighting or degree of relevance for each concept C_i to each question “ Q_j ” integrated from multiple experts and $0 \leq qc_{ij} \leq 1$.

From the probability viewpoint, the QC matrix can be useful to calculate the probability of failure for a student. This probability is calculated as the ratio of “ qc_{ij} ” to “ $\sum_{j=1}^m qc_j$ ”, moreover, if the student fails to answer more than one question, then the (1) can be used to calculate failing rates for each concept.

$$P(C_j) = \frac{\sum_{i=1}^m (1 - g_{ij}) \times qc_{ij}}{\sum_{i=1}^m qc_{ij}}, \tag{1}$$

where “ g_{ij} ”- represents the student’s answers for each of the questions and the test set has m questions. If “ g_{ij} ” is 1 then the student answers the question “ Q_j ” correctly, otherwise “ g_{ij} ” = 0.

2.3 Calculation Relevane Degree between Concepts

Concepts and concepts map play an important role in this study. Researchers argue that the notion of concept can be thought of as an atomic unit of knowledge and this knowledge is created by forming connections among concepts [9]. Moreover, each course unit is made up of a certain number of related concepts, and each concept has its own size and its own special importance or effect in other concepts. In[2], we proposed an approach to construct concept map and calculated the relevane degree

between concepts based on the corresponding relation of concepts and questions (Question-concepts matrix) to explore the degree of student familiarization with the concepts, according to collected data from the students' assessment records where the relevance degree between concepts determine as follows:

$$W(C_i \rightarrow C_j) = \text{Max}(W_{Q_x C_i} \times W_{Q_y C_j} \times \text{Conf}(Q_x \rightarrow Q_y)) \tag{2}$$

,where “ $W(C_i \rightarrow C_j)$ ”denotes the relevance degree of the relationship “ $C_i \rightarrow C_j$ ”converted from the relationship “ $Q_x \rightarrow Q_y$ ”, “ $C_i \rightarrow C_j$ ”denotes a concept appearing in the question “ $Q_x \rightarrow Q_y$ ”, “ $C_i \rightarrow C_j$ ”denotes a concept appearing in the question “ $Q_x \rightarrow Q_y$ ”, ($W_{Q_x C_i}$ denotes the weight of the concept “ $C_i \rightarrow C_j$ ”in the question “ $Q_x \rightarrow Q_y$ ”, $W_{Q_y C_j}$ denotes the weight of the concept “ $C_i \rightarrow C_j$ ”in the question “ $Q_x \rightarrow Q_y$ ” and $\text{Conf}(Q_x \rightarrow Q_y)$ denotes the confidence of the relationship“ $Q_x \rightarrow Q_y$ ” obtained by Apriori Algorithm [1].

Using the proposed approach provides us a visual knowledge representations, which show relationships among concepts as standardized learning order of different concepts, effect between it and can also serve as a tool of detect the learning barriers.

3 Intellectual Learning Guidance System

This study proposes a combination of Bayesian networks (also known as belief networks or Bayes nets for short) and CER model to provide learning guidance, which uses concept map as a tool for representation domain knowledge.

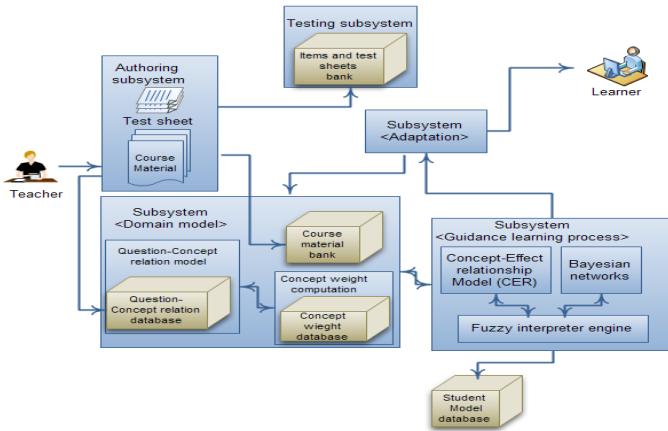


Fig. 2. System architecture

The architecture of system is depicted in Fig.2, which has five major subsystems/functional modules, including the authoring subsystem, testing subsystem, “domain model”, subsystem “Adaptation” and subsystem “guidance learning process”.

The details of each module/subsystem are described as follows:

1. Authoring subsystem: This subsystem is responsible for creating (i.e. teacher/tutor) courses content and setting the weight of each concept to each question.
2. Testing subsystem/module: This module is responsible for providing evidence about what learner knows and his knowledge level/background. Moreover, it used for mining relationships between concepts in domain model.
3. Domain model: This subsystem contains question- concept relation computation module and concept weight computation module.
 - The question-concept relation computation module used to set the relationship between the concepts appeared in test sheets by teacher or domain expert.
 - The concept weight computation module used to compute the weight or the relevance degree of each concepts using formula (2), which represent the Strength of effect is (for more detail see algorithm described in [2]).

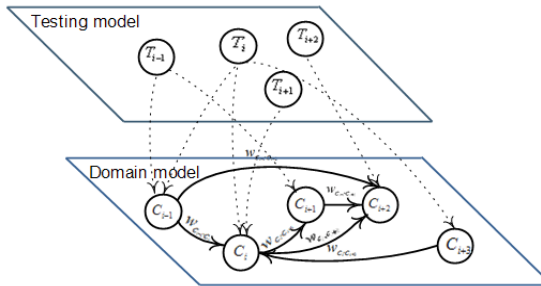


Fig. 3. Relationship between Testing model and Domain model

Both question-concept relation computation module and concept weight computation module collaborate together to generate concept maps that display relationships and concepts nodes.

4. Adaptation subsystem: This module is responsible for representing content and necessary adaptation. Adaptation algorithm in this context belongs to behavioral class model that means it is reactive with current student’s knowledge level [21] and conceptual weight.
5. Subsystem “Guidance learning process”: This module is essential part in the proposed system, where concept-effect relationship model which used to detect learning problem and Bayesian networks to provide forward guidance. Moreover,

to make the learning guidance more understandable to student and tutor/teacher, the fuzzy set theory is proposed. Let the four membership functions are proposed as shown in Fig. 4.

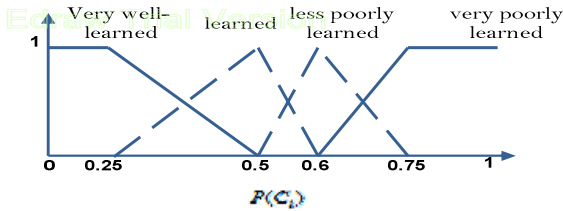


Fig. 4. The given membership functions

According to the demonstration was proposed in section 2.2, it can be seen that the student does not understand the concepts “ C_{i+2} ” and “ C_{i+3} ”, because the $P(C_{i+2}) = 0.92$, and $P(C_{i+3}) = 1$, where student’s learning status is very poorly learned (see Fig.4). For that reason to provide teaching suggestions to student system advice him to restudy the related concepts. Assuming that the teacher has defined the acceptable error ratio to be 0.2, as seen in Fig.1, it can easily determine the related concepts (C_{i-1} , C_i and C_{i+1}), however, concept C_{i+1} learned very well, thus both concepts “ C_{i-1} ” and “ C_i ” are assigned as misunderstanding of concepts “ C_{i+2} ” and “ C_{i+3} ”; moreover, the student should learn “ C_{i-1} ” before learning “ C_i ”, this does not represent that the learner does not understand “ C_{i-1} ” at all. It means that it is very probably that the learner has not fully understood a part it only. Therefore, concept “ C_{i-1} ” will be strongly recommended.

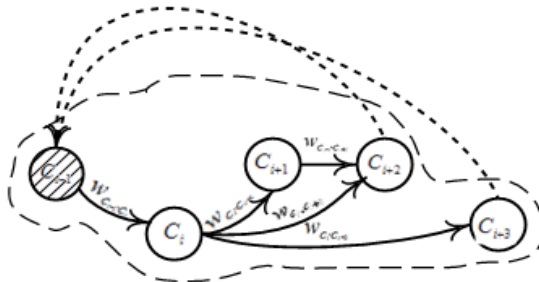


Fig. 5. Learning suggestion based on the concept-effect relationship model

4 Conclusions

In recent years, researchers have proposed different approaches for developing adaptive learning systems based on learning behaviors of students during their

interaction with e-learning systems. To improve the learning achievement of students, it is important to diagnose the learning problem and provide learning suggestions for individual students. In the last decade, several studies have been conducted for diagnosing student learning problems and providing appropriate learning guidance for individual students [20]. Among the existing models, Concept-effect relationship (CER) has been proved to be an effective way of improving the learning performance of student [11]. However, it is difficult and time consuming for teachers to give personalized suggestions to each student, particularly when there are many students in class. In this study an intellectual guidance system has been proposed to identify the learning problem of students. The next step of our study will be testing this approach with actual students and teachers and observing how it affects their learning performance.

References

1. Agrawal, R., Srikant, R.: Fast algorithms forming association rules. In: Proceedings of the 20th International Conference on Very Large Database, Santiago, Chile, pp. 487–499 (1994)
2. AL-Sarem, M., Bellafkih, M., Ramdani, M.: A New Method for Constructing Concept Maps in Adaptive E-Learning Systems. In: Lin, S., Huang, X. (eds.) CESM 2011, Part II. CCIS, vol. 176, pp. 180–185. Springer, Heidelberg (2011)
3. Butz, C.J., Hua, S., Maguire, R.B.: A Web-based Bayesian Intelligent Tutoring System for Computer Programming. In: Evolution of the Web in Artificial Intelligence Environments, pp. 221–242. Springer, Heidelberg (2008)
4. Chen, S.M., Sue, P.J.: A new method to construct concept maps for adaptive learning systems. In: Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao, pp. 2489–2494 (2010)
5. Chiou, C.K., Hwang, G.J., Tseng, J.C.R.: An auto-scoring mechanism for evaluating problem-solving ability in a web-based learning environment. *Computers & Education* 53(2), 261–272 (2009)
6. Chu, H.C., Hwang, G.J., Tseng, J.C.R., Hwang, G.H.: A Computerized Approach to Diagnosing Student Learning Problems in Health Education. *Asian Journal of Health and Information Sciences* 1(1), 43–60 (2006)
7. Günel, K., Aşlıyan, R.: Determining Difficulty of Questions in Intelligent Tutoring Systems. *The Turkish Online Journal of Educational Technology – TOJET* 8(3), 14–21 (2009)
8. Hooper, S.: Cooperative learning and computer-based instruction. *Educational Technology Research & Development* 40(3), 21–38 (1992)
9. Hsu, C., Kuo, R., Chang, M., Heh, J.: Implementing a Problem Solving System for Physics based on Knowledge map and Four Steps Problem Solving Strategies. In: 2nd International Conference on Advanced Learning Technologies (2002)
10. Hung, C.-L., Hung, Y.W.: A Practical Approach for Constructing an Adaptive Tutoring Model Based on Concept Map. In: VECIMS 2009 - International Conference on Virtual Environments, Human-Computer Interfaces and Measurements Systems, Hong Kong, China, May 11-13 (2009)
11. Hwang, G.J.: A conceptual map model for developing intelligent tutoring systems. *Computers & Education* 40(3), 217–235 (2003)

12. Hwang, G.J.: A test-sheet-generating algorithm for multiple assessment requirements. *IEEE Transactions on Education* 46(3), 329–337 (2003)
13. Hwang, G.J., Cheng, H.C., Chu, H.C., Tseng, J.C.R., Hwang, G.H.: Development of a web-based system for diagnosing student learning problems on English tenses. *Journal of Distance Education Technologies* 5(4), 80–98 (2007)
14. Hwang, G.J., Chu, H.C., Yin, P.Y., Lin, J.Y.: An innovative parallel test-sheet composition approach to meet multiple assessment criteria for national tests. *Computers & Education* 51(3), 1058–1072 (2008)
15. Hwang, G.J., Hsiao, C.L., Tseng, J.C.R.: A computer-assisted approach to diagnosing student learning problems in science courses. *Journal of Information Science and Engineering* 19(2), 229–248 (2003)
16. Jong, B.S., Lin, T.W., Wu, Y.L., Chan, T.Y.: Diagnostic and remedial learning strategy based on conceptual graphs. *Journal of Computer Assisted Learning* 20(5), 377–386 (2004)
17. Lee, C.H., Lee, G.G., Leu, Y.H.: Application of automatically constructed concept map of learning to conceptual diagnosis of e-learning. *Expert Systems with Application* 36(2), 1675–1684 (2009)
18. Ozdemir, B., Alpaslan, F.: An intelligent tutoring system for student guidance in Web-based courses. In: *4th International Conference on Knowledge-based Intelligent Engineering System and Allied Technologies*, vol. 2, pp. 835–839 (2000)
19. Panjaburee, P., Hwang, G.J., Triampo, W., Shih, B.Y.: A multi-expert approach for developing testing and diagnostic systems based on the concept-effect model. *Computers & Education* 55, 527–540 (2010)
20. Shabalina, O.A.: Models and methods of modeling learning process. Ph.D thesis, University of Astrakhan (2005)

The Study of Substation Automation Remote Communication for Multi-granularity Access Control System

Lichun Shi and Haimei Xiao

Henan Polytechnic, 450007, Zhengzhou, China
Shilichund2011@163.com

Abstract. This paper based on the demands of access control of substation automation system, analyzed two operation objects of access control including information model and information exchange service of IED, and designed a multi-granularity access control system according to RBAC and granularity of the object and access policy. The design of the system is centralized in logical, and distribute in realization. It provides a centralized control for user's access, the access policy is generated by privilege verify and implemented in the distributed IED. And instance of remote configuration is illustrated to demonstrate that the design can fulfill the needs of access control in substation automation.

Keywords: Substation automation, access control, RBAC, remote configuration.

1 Introduction

The power system is affected by the negative influence of information technology with the widely used of computer and network technology in power system. The hacker and virus in public network are becoming more and more rampant, of course that hacker's activities have been found by electric power system which is at home and abroad. In the substation automation, the current protocols have not set function to ensure the control instruction which is from the authorized initiator, and have no effective measures to prevent the terrorist threat during operating the circuit breaker; what's more, this phenomenon is very likely to happen through the communication line on the forged SCADA instructions. It has been becoming a very urgent task that how to protect the power dispatching control system effectively and its network information safety; and also how to realize the control of access permission of the important information or key operation. At present, the research institutions of domestic and foreign are in favor of the security problem in telecommunications for power system, the entry point of the research mainly focused on the security architecture of power information system and safety assessment problems, but has not always proposed the practical and feasible solution for the substation automation system, data exchange system configuration process exists the problem of access control.

2 Substation Automation System’s Requirement of Access Control

2.1 The Operation Object of Needs with Access Control

International Electric Committee published the IEC1850 standard in 2004, which summarize the substation information transmission that required for communication service, and design the ACSI (Abstract Communication Service Interface, ACSI) which is independent of the network and application layer protocol. ACSI provides the model of specific substation information and information exchange for substation. Figure 1 demonstrates the relationship between content and the model. In the image at left shows the hierarchical relations of model, data attribution form data, the date form logical node, and the logic point form logic device, eventually form the server model. Figure on the right side mainly demonstrates the information of control modules which is based on function, and the control modules mainly has relation with logic point, in which the constant group, generic objects oriented substation event (GOOSE), generic substation status event (GSSE), data of multicast sampled and single-cast sampled etc. all of these 5 kinds of control modules are function with logic point 0 (LINO).

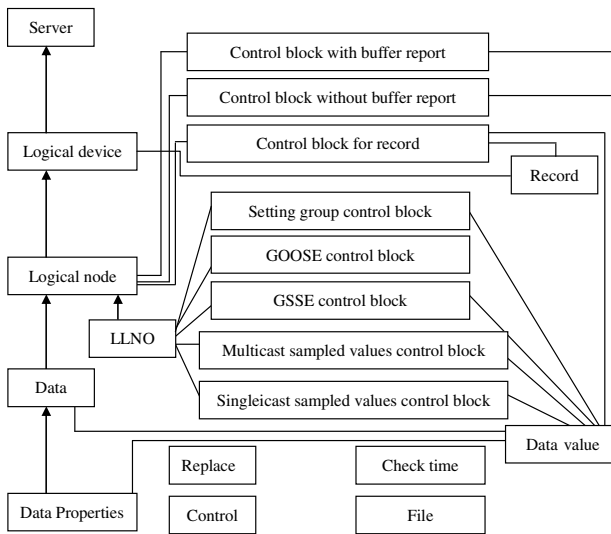


Fig. 1. Information Model and Communication Service Model for Substation

For the substation in the communication between different devices, it is realized by application of relation model in the formal standard, which is consisting of Two-Party-Association (TPAA) and Multicast-Application-Association (MCAA), both of the two means of communication have their corresponding access control problem. According to the IEC61850 standard, application of intelligent electronic device (IED) in substation can be seen as ACSI equipment, which could provide the corresponding information model and information exchange service model and when

the client access the server, according to the client access permission the access control will limit some operation which is related to logic device, logic point, date etc. In view of this, according to the specific user's permission, the access control in substation automation is designed to establish a set of rules and in order to limit the operation ability about the equipment data object and its function of different users.

2.2 The Command Message with Access Control Requirement

In the substation automation system, according to the function and demand the message is divided into 7 categories in the fifth part of IEC61850 standard, in which the seventh types is command message with requirement of access control. This message is used to send control command, which can be issued by the man-machine dialogue interface of local or remote and this situation requires high security. In the IEC6180 standard, according to the function and the logical abstraction the substation automation system can be divided into three layer system structure, which is layer substation, layer interval, layer process, what's more, the Ethernet is used among all the layers.

3 The Application of Access Control Model Based on Role in Substation Automation System

Just as the analysis above, substation automation system has higher demands for access control. Compared with traditional discretionary access control (DAC) and mandatory access control (MAC), the role based access control (RBAC) is a breakthrough in flexibility and control of particle size; on the other hand, it is replacing the traditional access control model gradually. Therefore, we think that RBAC is an effective solution of device access control problem in substation network system. So the following is the description of access control strategy, system function structure, and multi-granularity strategy execution in substation system.

3.1 Access Control Strategy of Substation System

For the access control of substation automation, we could define the authorization policy according to a four group <subject, object, action, condition>.

SUBJECT (sub)

Information of the requester is outputs that authenticated by the user in actually, which contain the identity of the user, but it maybe include substation operators, maintenance personnel, and administrator in the power system. Due to use the role-based access control model, so we use the role name to mark SUBJECT, which is the authorization in RBAC, and the role name is related to a set of access rights.

OBJECT (obj)

The target object resource index, according to the information model of IEC61850, can use the method of object reference which is called LDName/LNName [.Name [...]] to express the operation object or service that corresponding.

ACTION (act)

The operation is allowed to target object. When we operate the substation model, all the operation is up to application requirements, according to the fifth part of IEC61850, we could define the accession of device like this: Great, Delete, View, Set/write, Get/read, Execute etc. Great allows the user to create some specific application objects in LN; Delete allows the user to delete application object in specific LN; View allows the user to get the object that existed and its definition details; Set/write allows the user to set attribution values of the object; Get/read allows the user to get attribution values of the object; Execute allows the user to execute the application service which is permitted.

CONDITION (con)

To specify some constraints, that is to say under what conditions a request will be granted. Boolean expressions of some certain conditions can realize negative authorization, and it can also express the access control policy that more flexible.

According to the four groups, we could define the basic component that in the access control system.

Request of access control

REQUEST (REQ): obj sub act

In which $sub = role \cup group \cup uid$, but obj must identify a unique object, and a attribution exists in the act.

Authorization strategy

POLICY (POL) is four groups: <obj, sub, act, and con >.

If permission which is the attribution of action is grant and the condition of con is true, then the requester who is called sub could execute act for obj; On the other hand, if the condition of con is false, then the requester sub couldn't execute act for obj.

Decision of access control

Access control decision is three groups :< obj, sbj, act>, that is the output of the request evaluation module of authorization. If permission which is the attribution of action is grant, then the corresponding obj will execute act; If permission is deny, then the sbj will be refused to execute act for obj.

3.2 The Function Structure of Multi-granularity Access Control System

In the substation automation system, the date and function that in IED equipment is the access objet, generally the embed system is often used in IED equipment; in addition the available computer resources are very limited. Multi-granularity access control system is composed of the centralized access control service and the strategy execution module which is distributed to specific equipment. We can choose the access control service according to the specific circumstance, and the server device is arranged in the security position to ensure its safety. The function frame fig of accessing control system.

3.3 The Processing Method of Different Granularity Object in Multi-granularity

The access control of substation automation system includes coarse-grained and fine-grained, the coarse-grained is mainly refers to the information exchange service control, and the fine-grained is refers to the data access control of information model. The two controls are achieved by the RBAC model, so they have a lot in common,

such as the role of information sharing. The access controls of information exchange service and information model data are controlled by a unified strategy description, the two parts is the basic requirement for information role. In addition, in the realization of the process, the user / role management (UA in RBAC: user \leftrightarrow role) and role privilege management (PA in RBAC: role \leftrightarrow privilege) and other aspects are the same, because of its own characteristic in information model, the two have many different points in many ways, such as: The control for information exchange service, whose control object is a dynamic service; and the control for data object information model, whose control object is a static data, but the operational semantics for the object are different; The processing for information exchange service only needs to answer yes or no, so understanding the semantic of access resolution is not necessary; but access control of information model data must understand the semantic of access resolution, and then execute the corresponding action according to the access resolution (for example: create a data object or modify and delete the data etc.); In addition, the characteristics of information model in its own has different aspect in deal with exchange information service when decision point generates access resolution, such as managing the jurisdiction access according to the model tree structure. Therefore, the processing for the two kinds of control is separated from each other. In order to realize the mutual independence and simple operation, we must separate the two strategies, and then contact the two strategies through the consistency of role information.

4 The Application of Access Control System in Remote Configuration of Substation IED

With the development of automation technology and network technology, the remote configuration function of IED in the station already has the feasibility; it is mean that we should set the spacer layer such as relaying protection, fault recorder etc. through the network in the scheduling terminal, which make the system maintenance more convenient. At present, some of the intelligent devices in the station are equipped with a similar function, but when considering the safety, the distance distribution function does not put into use. The method of remote configuration IED is transferring file, as shown in figure 2.A. In the case of allows remote configuration, according to the information of the type of equipment, model and database, the dispatching terminal will generate the remote configuration automatically, which include setting the constant, casting or return the soft plate, sampling frequency by fault recorder, current variation value of various parameters; B. The operator can dispatch the remote configuration list in the end of the dispatching, and send to the end of station system after confirming authorization and completing the various settings. C. Anglicizing the configuration file in the end of station, and completing the set. Caution: profile in the station can be analyzed in the main control unit or a communication management machine, and then send the configuration information to IED according to protocol that used by device, shown as the point line in figure 3(A); with the development of communication network and the improvement of performance, it is possible that the configuration file can be sent to the device directly, and then accept analysis, shown as dotted line in figure 3(A).

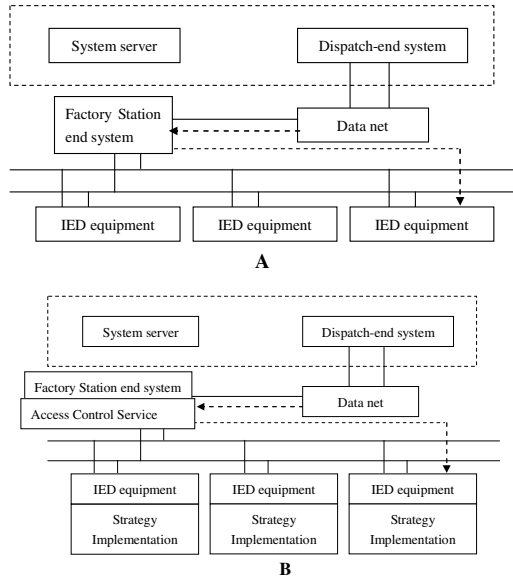


Fig. 2. Process of Remote Configuration

When we use the access control system, the access control server can be set in substation or dispatch center according to the requirement, here we set the access control server in substation, equip each IED device with execution strategy module. As shown in figure 3(B). According to substation communication system and characteristics of the network in the IED61850 standard, we set the definition of substation configuration language SCL, which communicate by model switching, and the five definitions Header, Substation, IED, Lnodetype and Communicate are included in it. In order to control the access of operations on data in the configuration and we extend the language SCL according to the security requirements, in which some description for identity information of the user are added. A. The identity element is included in the message which is send to the substation by substation; B. After station receives the message, the station will send it to access control server and confirm the access privilege; C. Firstly, access control service will identify the operator's identity; secondly evaluate the access request through the strategy management, and then generates the corresponding strategy execution element, finally send it to IED in order to execute; D. IED will analyse the configuration files according to the strategy executing elements, and then execute the corresponding operation. Addition: the above method is under the assumption that the IED equipment is consistent with the 6180 standard in the circumstances, on the contrary, if the method doesn't meet the condition, strategy execution module would be set in a main control unit or a communication management machine.

5 Conclusion

The design of this system is based on the requirement of access control in the substation automation system, and analyzing two operations of object which are the IED equipment information model and information exchange service and then research access strategy according to the size dividing machine, at last design a kind of multi-granularity control system based on the RBAC method. The system is logically centralized, distributed in realization; it can provide centralized control and check permission for the user's access, what's more, it also can generate the corresponding access resolution according to the request, and then distributed it to each IED device with access control in order to perform the control resolution.

References

1. Li, J.-P., Liu, T.-F., Zhang, C.-Q.: Analysis and Research on the Security of Remote MIS Based on ASP+SQL. *Journal of Northern Jiaotong University* (2002-06)
2. Wu, G.: Study on Privilege Management in MIS. *Office Informatization* (2008-12)
3. Jiang, F.-W., Yang, T.-N.: Role-crossed Distributed Authorization Model in Information System. *Journal of Chongqing Institute of Technology* (2007-08)
4. Li, T.: Apply the RBAC strategy to enterprise Information Systems. *Computer Knowledge and Technology* (2008-35)
5. Zou, S., Tong, X., Tong, L., et al.: Design of the General-Purpose Security Access Control Module for a Traction Transformer on-line Monitoring Information System. *Automation of Electric Power Systems* (2003-05)
6. Xiong, G.-H., Li, Z.-K.: Role-Based Access Control in Educational Administration System. *Network & Computer Security* (2007-07)

Parameter Estimation of Sinusoidal Signal under Low SNR

Mingjian Sun*, Xiaowei Zhang, and Fanming Meng

College of Science, Harbin Engineering University,
150001 Harbin, China

{MingjianSun, XiaoweiZhang, FanmingMeng}@hrbeu.edu.cn,
sunyong722@126.com

Abstract. Based on the known frequency, the amplitude and initial phase of sinusoidal signal can be estimated by using the proposed cross-correlation function in this paper. And the problem of signal loss on phase when using autocorrelation function to estimate amplitude is solved. By using the superposition algorithm, the SNR of the observed signal can be improved effectively and the signal energy would not lose, this improves the accuracy of the estimation. Simulation results prove that amplitude and initial phase can be estimated accurately at low SNR by using the cross-correlation function which based on the superposition algorithm.

Keywords: Sinusoidal signal, cross-correlation function, superposition, amplitude, initial phase.

1 Introduction

The parameter estimation of sinusoidal signal has important applications in the area of radar, sonar, electronic warfare, communications, biological and vibration signal processing [1]. The sensor is often mixed with strong noise when it receives signals, so it is important to estimate the sinusoid signal parameters accuracy under low SNR. The estimation methods given in literatures [2-3] are based on FFT algorithm. But in the case of non-equal interval sampling, the high accuracy parameter estimation cannot be obtained by the Fourier transform. Although the maximum likelihood estimation (ML) [4] can make the mean square error to minimum, ML algorithm is complicate, slow and inconvenient to real-time processing, so it is said in literature [5] that ML estimation is seldom applied directly. We can estimate the amplitude effectively for the autocorrelation function is a good noise suppression in statistical [6-7], but the phase information is lost at the same time. In this paper, we use cross-correlation function to estimate the amplitude and initial phase of the sinusoidal signal, and we retain the phase information which is lost when using autocorrelation function. In practice, the above methods have certain requirements of SNR [8-9], they can estimate the amplitude and initial phase with high accuracy in the case of high SNR, but the estimated effect is not ideal in the case of low SNR, so we need to improve the SNR of observed signal.

* Corresponding author.

In this paper, without losing the signal energy, we take advantage of the characteristics of periodic signal and random noise and use the superposition method to improve the SNR of observed signal. Simulation results show that the method can estimate the amplitude and initial phase with high accuracy at low SNR.

2 Cross-Correlation Function

Given two signals $x(t)$ and $y(t)$, assume T is the sampling time of the signals, $y(t + \tau)$ is the time-shift sample of $y(t)$. Therefore, the cross-correlation function of $x(t)$ and $y(t)$ can be shown as follows:

$$R_{xy}(\tau) = \frac{1}{T} \int_0^T x(t)y(t + \tau)dt.$$

Suppose that the observation process is ergodic, then the cross-correlation function of stochastic process can be taken by the cross-correlation function of sample,

$$R_{xy}(\tau) = \frac{1}{N} \sum_{i=1}^N x(i)y(i + \tau), \quad i = 1, 2, \dots, N.$$

With known frequency, the amplitude and initial phase are obtained by the cross-correlation function. Suppose the measured signal as follows

$$x(t_i) = s(t_i) + G(t_i), \quad i = 1, 2, \dots, N, \tag{1}$$

Where $s(t_i) = a \sin(2\pi f_0 t_i + \varphi_0)$ is the source signal, a is the signal amplitude, f_0 is the natural frequency, φ_0 is the initial phase, $G(t_i)$ is the zero mean and σ^2 variance additive white Gaussian noise.

Let $x_e(t_i, \varphi) = \sin(2\pi f_0 t_i + \varphi)$, $i = 1, 2, \dots, N$, $\varphi \in (-\pi, \pi]$ be a sinusoidal signal whose frequency is f_0 and amplitude value is 1. Obviously, $s(t_i) = ax_e(t_i, \varphi_0)$, $i = 1, 2, \dots, N$. $G(t_i)$ is supposed as the standard white Gaussian noise, therefore, the noise $G(i)$ and the signal $x(t_i, 0)$ are independent, then

$$R_{Gx_e(t_i, 0)} = \frac{1}{N} \sum_{i=1}^N G(t_i) \times x_e(t_i, 0) = \frac{1}{N} \sum_{i=1}^N [x(t_i) - s(t_i)] \times x_e(t_i, 0) = 0,$$

that is

$$\sum_{i=1}^N [x(t_i) - ax_e(t_i, \varphi_0)] \times x_e(t_i, 0) = 0. \tag{2}$$

Take $x_e(t_i, 0) = \sin(2\pi f_0 t_i)$ and $ax_e(t_i, \varphi_0) = a_1 \sin(2\pi f_0 t_i) + a_2 \cos(2\pi f_0 t_i)$ into the equation (2), where $a = \sqrt{a_1^2 + a_2^2}$, $\cos(\varphi_0) = \frac{a_1}{\sqrt{a_1^2 + a_2^2}}$, so that

$$\sum_{i=1}^N [x(t_i) - a_1 \sin(2\pi f_0 t_i) - a_2 \cos(2\pi f_0 t_i)] \times \sin(2\pi f_0 t_i) = 0. \tag{3}$$

Let $Y = [x(t_1) \sin(2\pi f_0 t_1) \quad x(t_2) \sin(2\pi f_0 t_2) \quad \dots \quad x(t_N) \sin(2\pi f_0 t_N)]^T$,

$$H = \begin{bmatrix} \sin^2(2\pi f_0 t_1) & \cos(2\pi f_0 t_1) \sin(2\pi f_0 t_1) \\ \sin^2(2\pi f_0 t_2) & \cos(2\pi f_0 t_2) \sin(2\pi f_0 t_2) \\ \vdots & \vdots \\ \sin^2(2\pi f_0 t_N) & \cos(2\pi f_0 t_N) \sin(2\pi f_0 t_N) \end{bmatrix}, A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix},$$

then equation (3) could be written as $Y = HA$, where H^T is the transposed matrix of H , $(\#)^{-1}$ is the matrix inversion symbol, $A = (H^T H)^{-1} (H^T Y)$. We obtain the estimation of amplitude and initial phase by $a = \sqrt{a_1^2 + a_2^2}$ and $\cos(\varphi_0) = \frac{a_1}{\sqrt{a_1^2 + a_2^2}}$.

3 Improve the Estimation Accuracy

In the condition of higher SNR, the amplitude and initial phase of the signal can be estimated with higher accuracy. But at low SNR, the estimated effect is not very satisfactory [8-9]. In order to improve the estimation accuracy, the signal-to-noise ratio could be raised by using the following superposition method.

The observed signal is

$$x(t) = \begin{cases} s(t) + G(t) & , \quad t \in [0, t_T] \\ 0 & , \quad \text{others} \end{cases},$$

where $s(t) = a \sin(2\pi f_0 t + \varphi_0)$, $T = \frac{1}{f_0}$ is the natural period, $t_T = mT$,

$(m = 1, 2, \dots)$, $G(t) = \sigma \text{randn}(\text{size}(s(t)))$ is the zero mean and σ^2 variance additive white Gaussian noise. The observed signal is superposed by translation and truncation operators to improve SNR. For the linearity property of the operators, the translation and truncation operations of $x(t)$ are equivalent to the operations on $s(t)$ and $G(t)$. The following translation and truncation processes are for $s(t)$, the processes of $G(t)$ are the same.

a) Translation operator: If the operator L_b is satisfied the equation $L_b(f(t)) = f(t+b)$, then L_b is defined as the translation operator. When $b < 0$, $f(t)$ is translated towards right for $|b|$. When $b > 0$, $f(t)$ is translated towards left for $|b|$.

b) Truncation operator: If the operator τ_c is satisfied the equation

$$\tau_c(g(t)) = \begin{cases} g(t) & , t \in [0, c] \\ 0 & , \text{others} \end{cases}$$

then τ_c is defined as the truncation operator, where $t \in [0, c] \subseteq [0, t_r], [0, t_r]$ is the definition domain of $g(t)$.

Steps of the superimposed processes are shown as follows:

(I) : Translate the signal $s(t)$ towards left for $\frac{T}{2}$, then

$$s_1(t) = L_{\frac{T}{2}}(s(t)) = \begin{cases} -a \sin(2\pi f_0 t + \varphi_0) & , t \in [-\frac{T}{2}, t_r - \frac{T}{2}] \\ 0 & , \text{others} \end{cases} \tag{4}$$

(II) : Truncate (4) in the interval $[0, t_r - \frac{T}{2}]$, that is,

$$\begin{aligned} s_2(t) = \tau_{(t_r - \frac{T}{2})}(s_1(t)) &= \begin{cases} \tau_{(t_r - \frac{T}{2})}(-a \sin(2\pi f_0 t + \varphi_0)) & , t \in [-\frac{T}{2}, t_r - \frac{T}{2}] \\ 0 & , \text{others} \end{cases} \\ &= \begin{cases} -a \sin(2\pi f_0 t + \varphi_0) & , t \in [0, t_r - \frac{T}{2}] \\ 0 & , \text{others} \end{cases} \end{aligned} \tag{5}$$

(III) : Truncate $s(t)$ in the interval $[0, \frac{T}{2}]$, that is,

$$s_3(t) = \tau_{\frac{T}{2}}(s(t)) = \begin{cases} a \sin(2\pi f_0 t + \varphi_0) & , t \in [0, \frac{T}{2}] \\ 0 & , \text{others} \end{cases} \tag{6}$$

(IV) : Translate $s_3(t)$ towards right for $t_r - \frac{T}{2}$, then

$$\begin{aligned} s_4(t) = L_{-(t_r - \frac{T}{2})}(s_3(t)) &= \begin{cases} L_{-(t_r - \frac{T}{2})}(a \sin(2\pi f_0 t + \varphi_0)) & , t \in [0, \frac{T}{2}] \\ 0 & , \text{others} \end{cases} \\ &= \begin{cases} -a \sin(2\pi f_0 t + \varphi_0) & , t \in [t_r - \frac{T}{2}, t_r] \\ 0 & , \text{others} \end{cases} \end{aligned} \tag{7}$$

(V) : From equation (5) and (6), we get

$$s_5(t) = -(s_2(t) + s_4(t)) = \begin{cases} a \sin(2\pi f_0 t + \varphi_0) & , t \in [0, t_r] \\ 0 & , \text{others} \end{cases}$$

(VI) : Superimpose $s_5(t)$ and $s(t)$, then

$$s_5(t) + s(t) = \begin{cases} 2a \sin(2\pi f_0 t + \varphi_0) & , t \in [0, t_r] \\ 0 & , \text{others} \end{cases}$$

After one superposition, the source signal $s(t)$ will be changed to

$$P_1(t) = s(t) + s_5(t) = 2s(t), \quad t \in [0, t_r].$$

The transformed noise $G_1(t)$ could be obtained by the above method.

After a superposition, the signal $x(t)$ will be noted as

$$Q_1(t) = \begin{cases} P_1(t) + G_1(t) & , t \in [0, t_r] \\ 0 & , \text{others} \end{cases}, \text{ where the amplitude of } P_1(t) \text{ is noted as } a_1 = 2a.$$

Then, the transformed signal $Q_1(t)$ will be superimposed once again and be noted as

$$Q_2(t) = \begin{cases} P_2(t) + G_2(t) & , t \in [0, t_r] \\ 0 & , \text{others} \end{cases}, \text{ where } P_2(t) \text{ is the result of } P_1(t) \text{ after one}$$

superposition. The amplitude of $P_2(t)$ is noted as a_2 and $a_2 = 2a_1 = 2^2 a$.

Through the superposition for n times, we conclude that

$$Q_n(t) = \begin{cases} P_n(t) + G_n(t) & , t \in [0, t_r] \\ 0 & , \text{others} \end{cases}. \tag{8}$$

Where $P_n(t)$ is the result of $P_{n-1}(t)$ after one superposition. The amplitude of $P_n(t)$ is noted as a_n and $a_n = 2a_{n-1} = \dots = 2^{n-1} a_1 = 2^n a$. Then we can get $P_n(t) = a_n \sin(2\pi f_0 t + \varphi_0) = 2^n a \sin(2\pi f_0 t + \varphi_0)$. We can define the transformed signal

$$x'(t) = \frac{Q_n(t)}{2^n} = \frac{P_n(t)}{2^n} + \frac{G_n(t)}{2^n} = s(t) + \frac{G_n(t)}{2^n}. \tag{9}$$

According to equation (9) the transformed signal $x'(t)$ can be gotten by the observed signal $x(t)$. Gaussian noise is not periodic, then the noise cannot be amplified by 2 exponential form, so that $G_n(t) < 2^n G(t)$. According to the equation (1) and (9), it is clear that the transformed source signal $s(t)$ is same as before and the noise of $x'(t)$ is lower than the noise of $x(t)$, that is, the SNR of the signal $x(t)$ can be improved by using the superposition algorithm and the signal energy would not be lost. So that the cross-correlation function based on superposition can estimate the amplitude and the initial phase of sine signal under low SNR more accurately.

4 Experimental Results

The proposed cross-correlation function is simulated by Monte Carlo method.

The observed signal is $x(t_i) = s(t_i) + G(t_i)$, $i = 1, 2, \dots, N$, where $s(t_i)$ is the sinusoidal signal defined by equation (1), $G(t_i)$ is the zero mean and σ^2 variance additive white Gaussian noise. The simulation parameters are shown as follows:

The frequency $f_0 = 10$ Hz, initial phase $\varphi_0 = \frac{\pi}{6}$ rad, amplitude $a = 1$, sampling frequency $f_s = 640$ Hz, sampling points $N_1 = 4096$, and the SNR is defined as

$$SNR = 10 \log \frac{\sum_{i=1}^{N_1} s^2(t_i)}{\sum_{i=1}^{N_1} x^2(t_i)} \text{ dB.}$$

The mean-square error of amplitude and initial phase are defined as:

$$MSE a = \frac{1}{N_2} \sum_{i=1}^{N_2} (a - a_i')^2, \quad MSE \varphi_0 = \frac{1}{N_2} \sum_{i=1}^{N_2} (\varphi_0 - \varphi_{0i}')^2 \text{ rad} \quad (i = 1, 2, \dots, N_2, \text{ where}$$

a_i' 、 φ_{0i}' are the i times estimation of amplitude and initial phase, N_2 is the simulation times.)

(1) We first study the impact of superposition times and sampling frequency on the accuracy of the estimation. Table 1 lists the estimated amplitude a^* and initial phase φ_0^* of the observed signal and the mean-square error obtained by Monte Carlo method for 500 times direct cross-correlation function at $SNR = -18$ dB.

Table 1. The estimation results by using direct cross-correlation function at $SNR = -18$ dB

n		$f_s^* = 320\text{Hz}$	$f_s^* = 640\text{Hz}$	$f_s^* = 1280\text{Hz}$
6	a^*	1.0714	1.0455	1.0147
	φ_0^* rad	0.4953	0.4911	0.5078
	MSE a	2.5524	1.0354	0.1087
	MSE φ_0 rad	0.0949	0.0430	0.0240
8	a^*	1.0398	1.0293	1.0133
	φ_0^* rad	0.4923	0.4944	0.5247
	MSE a	0.7935	0.4287	0.0890
	MSE φ_0 rad	0.0831	0.0527	0.0215
10	a^*	1.0576	1.0221	1.0166
	φ_0^* rad	0.5041	0.5037	0.5155
	MSE a	1.6603	0.2445	0.1375
	MSE φ_0 rad	0.0894	0.0496	0.0225
16	a^*	1.0538	1.0361	1.0128
	φ_0^* rad	0.4972	0.5014	0.5108
	MSE a	1.4447	0.6499	0.0815
	MSE φ_0 rad	0.0981	0.0506	0.0255

The meanings of symbols in Table 1 are shown as follows:

The superposition time is n . f_s^* is the sampling frequency. a^* and φ_0^* rad are the mean estimation of amplitude and initial phase respectively. $MSE a$ and $MSE \varphi_0$ rad are the mean-square error of amplitude and initial phase respectively.

Table 1 shows that the mean-square error of parameters estimation decrease when n and f_s^* increase. The most suitable one is $n = 10$, $f_s^* = 640Hz$. In this way, it can reduce computing time and can also improve estimation accuracy.

(2) We study the impact of SNR on the accuracy of the cross-correlation function. Table 2 lists the parameter estimation of the observed signal and the mean-square error obtained by Monte Carlo method for 500 times direct cross-correlation function at different SNR. The symbol meanings are same as that in Table 1.

Table 2. The estimation results by using direct cross-correlation function

SNR dB	a^*	φ_0^* rad	$MSE a$	$MSE \varphi_0$ rad
-15.1860	1.0082	0.5134	0.0335	0.0117
-17.0749	1.0237	0.5230	0.2804	0.0165
-18.6267	1.0165	0.5110	0.1363	0.0253
-19.9577	1.0135	0.5020	0.0908	0.0399
-21.1116	1.0392	0.5060	0.7669	0.0434
-22.1326	1.0446	0.4856	0.9936	0.0541
-23.0343	1.0526	0.4768	1.3820	0.0727
-24.0149	1.0502	0.4508	1.2610	0.0980
-25.0185	1.0527	0.4435	1.3890	0.1099

Table 2 shows that a high estimation accuracy is obtained by using cross-correlation function at high SNR(SNR>-24dB), the estimation accuracy of cross-correlation function decreases when SNR<-24dB. It shows higher SNR is required when using cross-correlation function.

(3) Table 3 lists the estimated value of the observed signal amplitude and initial phase, the mean-square error and SNR are obtained by Monte Carlo method for 500 times cross-correlation function based on superposition method(here, the signal is superimposed 10 times, that is, in equation (8), $n = 10$) at different SNR. The symbol meanings are same as that in Table 1.

Compared with Table 2, the superimposed cross-correlation function can improve estimation accuracy and reduce mean square error. And it shows the effectiveness of the superposition method. But when SNR<-25dB, the accuracy will decrease. Therefore, the superposition method can improve the SNR of the observed signal at a certain extent.

Table 3. The estimation results by using cross-correlation function based on superposition method

<i>SNR dB</i>	a^*	ϕ_0^* rad	<i>MSE_a</i>	<i>MSEϕ_0 rad</i>
-15.1875	1.0106	0.5170	0.0562	0.0113
-17.0642	1.0059	0.5106	0.0173	0.0149
-18.6305	1.0182	0.5133	0.1662	0.0240
-19.9557	1.0331	0.5059	0.5474	0.0308
-21.1054	1.0372	0.5110	0.6927	0.0413
-22.1257	1.0563	0.4863	1.5869	0.0524
-23.0260	1.0467	0.4826	1.0905	0.0655
-24.3142	1.0578	0.4948	1.6700	0.0869
-25.0378	1.0714	0.4863	2.5456	0.0959

5 Conclusions

In order to estimate amplitude and initial phase of the sinusoidal signal under low SNR, a new algorithm with no energy loss is proposed to improve SNR. Then, amplitude and initial phase are estimated by using cross-correlation function. The accuracy of estimation is improved in a certain range. Meanwhile, the problem that amplitude and initial phase cannot be estimated at the same by using auto-correlation function is solved. It is found that sinusoidal parameters cannot be estimated by using cross-correlation function which based on superposition if the observed signal is mixed with harmonic waves.

References

1. Luo, P.: Statistical Signal Processing. Publishing House of Electronics Industry, Beijing (2009)
2. Li, C., Zhang, X., Zhang, Z., Chang, J.: Phase estimation of sinusoid signal based on DFT and error analysis. *Journal of Beijing University of Aeronautics and Astronautics* 33(5), 580–584 (2007)
3. Zhu, J., Tang, B., Du, Z., Yao, X.: Fast Frequency Estimation Algorithm Using High order Approximations to Kernel Function and Interpolation on Fourier Coefficients. *Journal of Data Acquisition & Processing* 24(6), 797–801 (2009)
4. Rife, D.C., Boorstyn, R.R.: Single Tone Parameter Estimation From Discrete-Time Observations. *IEEE Transactions on Information Theory*, 591–598 (1974)
5. Deng, Z., Liu, Y.: The Starting Point Problem of Sinusoid Frequency Estimation Based on Newton's Method. *Acta Electronica Sinica* 35(1), 104–107 (2007)
6. Li, T.F., Xiong, J., Wang, J.: A New Testing Method for High-frequency Power Wave Trapper Measurement Based on Correlation Theory. In: *Industrial Mechatronics and Automation, ICIMA*, pp. 340–342 (2009)

7. Chen, M., Liu, Z.: The Detection of Weak Sinusoidal Signal by Multi-Layer Auto Correlation. *Noise and Vibration Control* 26(5), 28–30 (2006)
8. Wang, D., Su, Y., Ma, X.: A New Method of Pole Extraction Using the Cross-Correlation-Based Processing. *Acta Electronica Sinica* 33(6), 1015–1018 (2005)
9. Qi, G., Lv, J.: Variance Analysis Oil Sinusoid Frequency Estimators Based on The Argument of The Sample Autocorrelation Function. *Journal of Dalian Maritime University* 33(4), 5–9 (2007)

Contour Algorithm Based on GVF Snake

Fan Zhang^{1,2}, Shengwei Li¹, Gang Zheng³, and Rui Li¹

¹ College of Computer and Information Engineering, Henan University, Kaifeng, Henan 475001, China

² Institute of Image Processing and Pattern Recognition, Henan University, Kaifeng, Henan 475001, China

³ Kaifeng Electric Power Company, Kaifeng 475001, China

Abstract. In order to solve the initial active contour problem of Snake model, Contourlet transform is introduced into the GVF Snake model, which will provides a way to set the initial contour, as a result, will improves the edge detection results of GVF Snake model effectively. Firstly, the contours of the object in images can be obtained based on Contourlet Transform, and this contours will be identified as the initial contour of GVF Snake model. Secondly, then GVF Snake model is used to detect the contour edge of human gait motion. Experimental results show that the proposed method can extract the edge feature accurately and efficiently.

Keywords: Contour Extraction, GVF Snake models, Contourlet Transform, Gait Recognition.

1 Introduction

Feature extraction and description is the key steps for gait recognition. How to extract the motion contour effectively is most important in the gait recognition, and it is the main point of this paper. Active contour model is a powerful tool to solve the contour extraction in gait recognition.

Active contour model, also called snakes, is a framework for delineating an object outline from a possibly noisy 2D image [1]. This framework attempts to minimize an energy associated to the current contour as a sum of an internal and external energy. The external energy is supposed to be minimal when the snake is at the object boundary position. The most straightforward approach consists in giving low values when the regularized gradient around the contour position reaches its peak value. The internal energy is supposed to be minimal when the snake has a shape which is supposed to be relevant considering the shape of the sought object. The most straightforward approach grants high energy to elongated contours (elastic force) and to bended/high curvature contours (rigid force), considering the shape should be as regular and smooth as possible. In two dimensions, the active shape model represents a discrete version of this approach, taking advantage of the point distribution model to restrict the shape range to an explicit domain learned from a training set [2-4].

The Contourlet transform has a fast implementation based on a Laplacian Pyramid decomposition followed by directional filter-banks applied on each band-pass sub-band.

Contourlet transform can be used to captures smooth contours and edges at any orientation. The transform decouples the multi-scale and the directional decompositions. The multi-scale decomposition is handled by a Laplacian pyramid. The directional decomposition is handled by a directional filter bank. In order to solve the initial active contour problem of Snake model, in this paper, Contourlet transform is introduced into the GVF Snake model, which will provides a new way to set the initial contour, as a result, will improves the edge detection results of GVF Snake model effectively.

2 GVF Snake Model

Active contour model (snakes) can move under the influence of internal forces within the curve itself and external forces derived from the image. There are two key difficulties in Snake models, the first is the location of the initial contour setting, and the second is the bad convergence effect to concave boundary regions of image.

Snake models minimize the energy functional to restrain the contour of target object. A traditional snake is a curve $X(s) = (x(s), y(s)) \ s \in [0,1]$, and its minimizing energy function is as follows,

$$E = \int_0^1 \frac{1}{2} (\alpha | X'(s) |^2 + \beta | X''(s) |^2) + E_{ext}(X(s)) ds, \tag{1}$$

where α and β are parameters which control the snake’s tension and rigidity respectively. The first term is the internal force, which controls the curve changes, while the second term E_{ext} is the external force, which pulls the curve to desired features. Different E_{ext} can be constructed in different models.

To analyze the movement of snake model curve from the aspect of force balance, the minimized E of a snake must satisfy the Euler equation:

$$\alpha X''(s) - \beta X''''(s) - \nabla E_{ext} = 0. \tag{2}$$

As the Snake contour is dynamic, the $X(s)$ can be viewed as the function of t and s , then,

$$X_t(s, t) = \alpha X''(s) - \beta X''''(s) - \nabla E_{ext}. \tag{3}$$

Once getting solution of the equation (3), we will find a solution of equation (2).

Snake model is sensitive to the initial position. An initial contour should be as close as to the real contour; otherwise the result is often unsatisfied. Continuous image information is often used to obtain the initial contour of an image when Snake model is used in the image sequences. But sometimes the initial contour is not close to the true profile.

Because of some problems in snake model, the numbers of improved approaches have been proposed. In [5], the GVF (gradient vector flow) field replaces the potential force field in equation (3). Defining a new snake, GVF Snake, is computed as a diffusion of the gradient vectors of a gray-level or binary edge map derived from the image. GVF Snake has the better performance than snake in many fields. GVF Snake

expands the capture region of the curve, and can force the curve into the concave regions.

Assuming $f(x, y)$ is the contour image of a gray-scale image $I(x, y)$, and then ∇f is the vector field. If ∇f is iterative diffused to the edge of image, it will form a gradient vector flow field (GVF), $V(x, y) = (u(x, y), v(x, y))$. Then the minimize the energy function of contour is as follows,

$$\mathcal{E} = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2 \, dx dy . \tag{4}$$

The parameter μ is the control parameters, which is set according to the image noise. The parameter μ provides the tradeoff between the first part and the second part of the energy function. When ∇f is small, the curve is away from the target edge, the partial derivatives dominate the energy function. When ∇f is large, the energy is mainly depends on the second term of the integrand. Energy minimization is achieved by iterative approximation to the target edge, and the following Euler equations must be satisfied:

$$\begin{aligned} \mu \nabla^2 u - (u - f_x)(f_x^2 + f_y^2) &= 0 \\ \mu \nabla^2 v - (v - f_y)(f_x^2 + f_y^2) &= 0 \end{aligned} , \tag{5}$$

where ∇^2 is the Laplacian operator. $f_x^2 + f_y^2$ is the maximum in edge regions, and is zero in homogeneous regions. So this can be controlled edge gradient only in the boundary, and it will avoid the effect of the diffusion weaken. Equation (5) can be solved by view u and v as the function of time t , then,

$$\begin{aligned} u &= \mu \nabla^2 u(x, y, t) - b(x, y)u(x, y, t) + c_1(x, y) \\ v &= \mu \nabla^2 v(x, y, t) - b(x, y)v(x, y, t) + c_2(x, y) \end{aligned} . \tag{6}$$

3 Contourlet Transform Based GVF Snake Edge Detection

In the GVF field, some points have an important influence to the initial contour setting. When these points within the target, the initial contour must contain these points; and when these points outside the target, the initial contour do not contain these points, otherwise it will not converge to correct results. We call these points the critical points. The creation of critical points has the following factors: (1) Critical points are associated with the image gradient ∇f . By minimizing the energy functional, ∇f will be spread out, the impact of noise generated by the local minima will produce the critical points. (2) Critical points are associated with the smoothing coefficient μ . The higher the value is, the greater the smoothing effect is. The critical points will reduce, and GVF performance will reduce. Choosing a large-scale Gaussian smoothing and a smaller value will get good results. (3) Critical points are associated with iterations number of solving equations. When the number of iterations is more than a thousand times, GVF field has changed dramatically. Although we

require the final converged solution, but in the actual image processing applications, an intermediate result is acceptable. Therefore, the initial contour should contain the critical points as much as possible. So we can reduce the need of capture and reduce the amount of computation.

Contourlet offers a much richer sub-band set of different directions and shapes, which helps to capture geometric structures in images much more efficiently [7-9]. The Laplacian pyramid (LP) is first used to capture the point discontinuities, and then followed by a direction filter banks (DFB) to link point discontinuities into linear structures. In particular, contourlets have elongate supports at various scales, directions, and aspect ratios. The contourlets satisfy anisotropy principle and can capture intrinsic geometric structure information of images and achieve better a sparser expression image than discrete wavelet transform (DWT).

In order to solve the initial active contour problem of Snake model, Contourlet transform is introduced into the GVF Snake model, which will provides a way to set the initial contour, as a result, will improves the edge detection results of GVF Snake model effectively. Firstly, the contours of the object in images can be obtained based on Contourlet Transform, and this contours will be identified as the initial contour of GVF Snake model. Secondly, then GVF Snake model is used to detect the contour edge of human gait motion.

The proposed algorithm is described as follows:

(1) Applying Contourlet transform in the original image. The Contourlet transform coefficients, $C_{j,k}^{(l)}[n]$ ($0 \leq l \leq J, 0 \leq j \leq J, 0 \leq k \leq 2^l$), and approximate sub-band coefficients $C_0[n]$, can be obtained in different scale and different directions. Where 2^j denotes the scale of LP, l denotes the DFB decomposition levels, k denotes the direction serial number.

(2) Let the approximate sub-band coefficient $C_0[n] = 0$, and let remains unchanged $C_{j,k}^{(l)}[n]$.

(3) Processing $C_{j,k}^{(l)}[n]$ using a threshold.

(4) Detecting the maximum point of $C_{j,k}^{(l)}[n]$, and other non-maximum points are assigned to 0. Then the matrix $M_{j,k}^{(l)}[n]$, the modulus maximum matrix of Contourlet transform coefficients $C_{j,k}^{(l)}[n]$ is obtained.

(5) Applying the Contourlet inverse transform to $M_{j,k}^{(l)}[n]$ and $C_{j,k}^{(l)}[n]$, getting a single pixel wide edge of object in images.

(6) Using the edge that obtained by Contourlet transform as the initial contour of the GVF Snake model. Applying GVF snake model control the movement of contour until get the target profile.

(7) Decomposing the each level of LP decomposition of low-pass images. Setting the upper contours as the initial goal of edge contours, searching the target edge contours continually until the ultimate goal edge is detected. Then the iteration is terminated.

4 Experimental Results

In this paper, the edge contours of human motion gait are used as the experimental objects. Contourlet transform is introduced into the GVF Snake model, which will provide a way to set the initial contour to the edge extraction. Experimental implementations are based on the MATLAB 7.0. Experimental images come from Biometrics and Security Research Center of gait recognition database [7]. This algorithm is based on the GVF Snake model and Contourlet Transform processing, GVF Snake model $\mu = 0.2$, the number of iterations is 120 times.

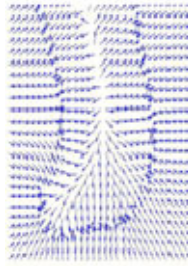


Fig. 1. The foot force field.

Fig. 1 shows the foot force field of gait contour. Some of the characteristics of GVF Snake model can be found according to the balance perspective of force field. When the initial contour is not setting in false zero force field, we can get very good results. The Calculating of the number of iterations does not affect the GVF force field. GVF force field not only keeps the diffusivity and the smoothness of Laplacian force field, but also maintains the characteristics that the gradient force field locates in the real contours and points to the real contours.

Fig. 2 is the comparison of extracted body contour using GVF Snake model and Contourlet transform based GVF Snake model respectively. The parameters of both two models are $\mu = 0.2$, and the number of iteration is 100 times. Fig. 2a is the final result of contour using GVF Snake model. When we increase the number of iterations continually, and it mains increase the computing time, the edge contour still cannot deeper recessed area of body, the search is terminated.

Fig. 2b is the final result of contour using the Contourlet transform based GVF Snake model. Body crotch area usually is difficult to track. GVF Snake model can not get a good result. Contourlet transform based GVF Snake model uses the edge which is obtained by Contourlet transform as the initial contour of the GVF Snake model. Applying GVF snake model, we can control the movement of contour until getting the target profile. Guaranteed point boundary gradient vector flow field and boundary gradient strength is better preserved, well into the recessed area, extracting the external continuous contour of the body without affect by the influence of various complexity factors of body, and gaining a better body contour.

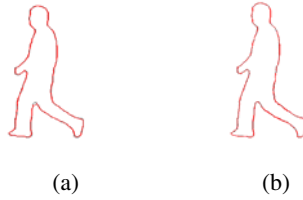


Fig. 2. The comparison of extracted body contour using GVF Snake model (a) and Contourlet transform based GVF Snake model (b) respectively.

5 Conclusions

The moving target detection is the key step in gait recognition. In order to solve the initial active contour problem of Snake model, the paper introduced the Contourlet transform into the GVF Snake model, which will provides a way to set the initial contour, as a result, will improves the edge detection results of GVF Snake model effectively. Based on Contourlet transform, the contours of object are identified as the initial contour in GVF Snake model. The human motion gait images are used as experimental objects. Experimental results show that the method can accurately and efficiently extract the edge feature.

Acknowledgments. This research was supported by the National Natural Science Foundation of China grants No. 60873039, the Medical Science and Technology Research Project of Henan Province, grants No. 2011020114.

References

1. Kass, M., Witkin, A., Terzopoulos, D.: Snake: Active contour models. *International Journal of Computer Vision* 1(4), 321–331 (1987)
2. Zosso, D., Bresson, X., Thiran, J.: Geodesic Active Fields - A Geometric Framework for Image Registration. *IEEE Transactions on Image Processing* 20(5), 1300–1312 (2011)
3. Song, W., Keller, J., Haithcoat, T., Davis, C.: Automated Geospatial Conflation of Vector Road Maps to High Resolution Imagery. *IEEE Transactions on Image Processing* 18(2), 388–400 (2010)
4. Niu, X.: A semi-automatic framework for highway extraction and vehicle detection based on a geometric deformable model. *Photogrammetry & Remote Sensing* 61, 170–186 (2006)
5. Xu, C., Prince, J.L.: Snake, shapes, and gradient vector flow. *IEEE Transactions on Image Processing* 7(3), 359–369 (1998)
6. Do, M., Vattedi, M.: The Contourlet Transform: An Efficient Directional Multi-resolution Image Representation. *IEEE Transactions on Image Processing* 14(12), 2091–2106 (2005)
7. Center for Biometrics and Security Research, Appendix: Springer-Author Discount, <http://www.cbsr.ia.ac.cn/>

The Design of Partial Electric and Control System for 500kV Substation

Haimei Xiao and Lichun Shi

Henan Polytechnic, 450007, Zhengzhou, China
xiaohaimeid2011@163.com

Abstract. The good design of the electric section in substation and its stable move may effect the electrical system of substation directly, this paper aim at the design of an electric primary section in substation and its supervisory system for the 500kv substation. based on the principle of connecting line of the electric primary section in substation and the function of the realization, and then by comparing the design of the double bus bars of double with bypass line subsection and line of 3/2 circuit breaker. we fixed the main wiring scheme. The design based on Programmable Logic Controller and supervisory control and data acquisition is an electric primary section of monitoring system.and the monitoring and control of the state and parameter in electric primary section in substation are realized by the monitoring system, which ensure the security of system, improve the level of automation and intelligent.

Keywords: Substation, electric primary section, supervisory control system, PLC.

1 Introduction

With the high-speed development of China economy, the demand of energy continues to expand, transmission and transformation engineering voltage rises the growing of grid structure become more and more complexity, and the grid real-time information transmission quantity increase multiple, the reliability requirements for the operation of power grid are also getting higher and higher. The reliability of design to operation normal in substation mainly achieved by two aspects: One is the rational design of electric primary section. The electrical main wiring design is the first part in the electric primary section, which may influence the substation itself and the reliability, flexibility and economy of the power system, therefore, it is very important to determine the main wiring. The other is the design of supervisory control. Real time monitoring of substation equipment and various types of data can help us find and remove the fault of system, and maintain the security and reliability for the system effectively. Therefore, take the 500kv substation for example to research and design the main wiring and its supervisory control for electric primary section in the substation. Improving the application of all the above is very crucial for realizing the design of electrical system.

2 The Design of Electric Primary Section in Substation

Take the 500kV hub substation for example to illustrate the design of electric primary section in 500kV substation. 500kV super voltage substation has a big capacity, high voltage, high power transmission line loop, and all of these are the hub of Substation. The hub substation contains 500kv into four lines, the length are 200km and 300km, which is belong to four independent power supply; 110kV feeder line 8 times, from the load side were: 40km, 50km, 60km, 70km. Four main installation transformer substations, the system take the infinite into consideration.

2.1 The Design of Main Wiring

The design of electric primary section in substation is the main wiring. Main wiring is composed of power and distribution circuit of high voltage electrical equipment by connecting lines; it reflects the various electrical equipments, connection mode and the mutual relations constitute the main body of electrical substation. Main wiring is an important part in power system; it influences the security, stability, flexibility and the arrangement of power distribution device, protective relaying, automatic device and the choice of the control method. Therefore, it must be designed to meet the reliable work, flexible operation and convenient maintenance and operation, and with the possibility of economic and development requirements. The following is the comparison of the two design scheme.

First, wiring by double bus bars of double with bypass line subsection scheme.

Wiring use double bus bars of double with bypass line subsection scheme equivalent to add the circuit breaker into the bus section, although it has the advantages of double bus bar with bypass, but it costs a lot of the investment, occupy more space of the device, so it must have the following conditions:

- (1) When the outgoing and ingoing line equipment is 12 to 16, we must set the relay in the same bus bar;
- (2) When the outgoing and ingoing line equipment is above 17, we must set the relay in two bus bars.

Second, wiring by 3/2 circuit breaker.

Wiring 3/2 circuit breaker sends two loops in every two circuit breaker, generally used in 500kV power grid bus bar connection. The following is the advantage of this scheme:

- (1) It is very flexible to operate, under normal circumstance the two bus bars and all circuit breakers will be on work, and multi ring power supply;
- (2) It is very convenient to repair, when there is something wrong with a group of bus, it is no need to switch the loop, for example there's no need to switch the circuit breaker when any of them stop to work;
- (3) Reliable operation, and each loop has two sets of circuit breakers to supply power, that is to say when one circuit breaker stops to work, any loop will continue with power.

The disadvantage of line of 3/2 circuit breaker is that it will use many equipments, especially the circuit breaker and current transformer, which will cost a lot of money and very complex to connect the line.

Compared with the two design of the main wiring on the reliability and economy, 3/2 circuit breaker will has less device in low power, when there's something wrong with the bus bar, there will be no device low power, and it is very easy to operation the isolate switch, and flexible operation, high reliability. In terms of investment, when the number of elements is 6 to 10, the design is very cheap compared whit the double bus bars of double with bypass line. When the original is more than 10, 3 / 2 circuit breaker connection investment will increase, but the investment of reliable system is also far less then the major accident. Therefore, this design adopts the method 2.

2.2 The Selection of Main Transformer Capacity

The selection of main transformer capacity is very crucial, if the small capacity can't meet the requirement of load growth, on the contrary, it will loss many power and can't meet the requirement of reducing the power. So, the capacity of the substation transformer is considered according to the general substation built after 5~10 years planning considerations, and should be in accordance with the load, so that the load will achieve 60% to 70% of the maximum load S_{max} . The main transformer capacity of the substation is 370MVA, and the model of transformer is SFPZ-370000/500.

2.3 The Selection of Electric Equipment

The first thing is to choose the material, when choose the bus bar, for example, we generally choose steel core, then select the section and model according to the current density. According to the minimum cross-sectional area method we check bus thermal stability and the corona condition check bus critical corona voltage, ultimately to determining bus types. Here we select the steel core aluminum stranded conductor which type is LGJQT-1400 mm². On the basis of the places installed and structure type of circuit breaker, we choice the less oil circuit breaker in the outdoor, we choose like this: $U_e \geq 500kV$; $I_e \geq 897.22A$; $I_{ekd} \geq I_{dt} = I_x = 2.202kA$. Selecting LW12-500 less oil circuit breaker according to the "manual design of electric power engineering", which in table 1.

Tabel 1. The main technical parameters of Outdoor-type oil-minimum breaker for LW12-500

Model	Rated voltage	Rated Current	Rated cut-off current	limit flux Peak	RMS	Inherent Separating brake time	5 s Stable current
LW12-500	500kV	4000A	63kA	160 kA	125 kA	20ms	63kA

Check: rated voltage $U_e = 500kV = UN$; rated current $I_e \geq 4000A$.

1) Rated breaking current calibration:

500kv bus bar short-circuit current of sub-state $I_x(3) = 12.64kA$

The rated breaking current of LW12-500 circuit breaker is $63\text{kA} > 12.64\text{kA}$, meet the requirements.

2) The check of dynamic stability:

500kv bus short circuit impulse current: $I_{gh}=32.232\text{kA}$; the passing current limits LW12-500 circuit breaker: $I_{et}= 125\text{kA}$, $I_{sk} < I_{et}$, meet the requirements.

3) The check of stability thermal:

If the protection time of backup is 1.9s, according to the selected circuit breaker parameters, the inherent opening time of circuit is 0.02s, the time of cease-fire is $t=0.03\text{s}$.

Devices short-circuit duration: $t_1=1.9\text{s}+0.03\text{s}+0.02\text{s}=1.95\text{s}$. The thermal effect of short circuit: $Q_k500= I_x(3) \cdot t_1=(12.64\text{kA})^2 \cdot 1.95\text{s}=31.55\text{kA}^2 \cdot \text{s}$; Allowable heat effect: $Q_k500=(63\text{kA})^2 \times 5\text{s}=19845\text{kA}^2 \cdot \text{s}$. Therefore, the circuit breaker meets the heat stability. The method of breaker selection of the rest of circuit breaker circuit is similar to 500kV circuit breaker. The selection of 110kV side isolator choice

1) In order to protect of electrical equipment and bus maintenance, we choose the isolation switch with a grounding knife switch.

2) The isolation switch is mounted in the outdoor, so we choose the outdoor type.

3) The rated voltage of this circuit is 110kV, therefore, the rated voltage of isolation switch is more then 110kV, and the rated current of isolating switch is much more then the circuit breaker 897.22A.

4) Therefore, we choose the type GW4-110/1000, which is the single grounding isolating switch, the main parameters shown in table 2.

Table 2. The main parameters of Single-grounded high-voltage isolation switch for GW4-110/1000

Model	Rated voltage	Rated current	Peak current limit by	4s Thermal current
GW4-110/1000	110kV	1000A	80kA	25kA

3 Design of Supervisory Control

The design of supervisory control for electric primary section in substation plays an important role in safe and stable operation of power system, reduce the cost of substation. The supervisory control in substation must with information technology and communication technology assistance, real-time detection of electrical substation is a part of equipment state and operation data, after get the relevant information, we must carry out the necessary analysis for relevant information, and submitted to the monitoring terminal by the type of audio notification and intuitive form, at last provide the direct evidence for management and the operating personnel.

The overall structure of the system is shown in figure 1. The design of the system is divided in 3 layers: the first layer is located in the upper of the machine, and control computer and configuration software made up of the substation, design for double redundant, which is located in the control center; the second layer is PLC. We select the series of products from Siemens, design for soft redundancy, mounted in the field

of control room. The third layer is a field device, including circuit breaker, isolating switch equipment state supervisory and the voltage, current and other parameters supervisory. We connect the substation supervisory computer to Ethernet through a common card, using TCP/IP protocol to complete the communication between these. PLC is responsible for supervisory equipment state feedback data; acquire parameter information, according to the host computer instructions and its control program of control equipment. Computer with configuration software can exchange data through the PLC, and keep watch on the equipment status at the real-time, at then concurrent control commands to PLC. At the same time, the system uses the redundancy structure. Configuration software support control redundancy equipment, redundancy network, redundancy computer and other redundancy mode. There are two PC computers in the control system. One is the master station, and the other is slave station. Normally only master and PLC will complete communication, data from the slave station through the main station for backup and synchronization, if the master fails, the slave will take over the master's work, when the claim is restored, it will be back automatically or manually. The configuration software provides an effective support structure for redundancy of the PLC, and communicate with the PLC main system normally, when the main system fails, PLC control tasks from the main system switch to the backup system, configuration software will be based on redundancy switch logo disconnect the communication with the main system, and then communicate with the backup control system to ensure the accuracy of switching time.

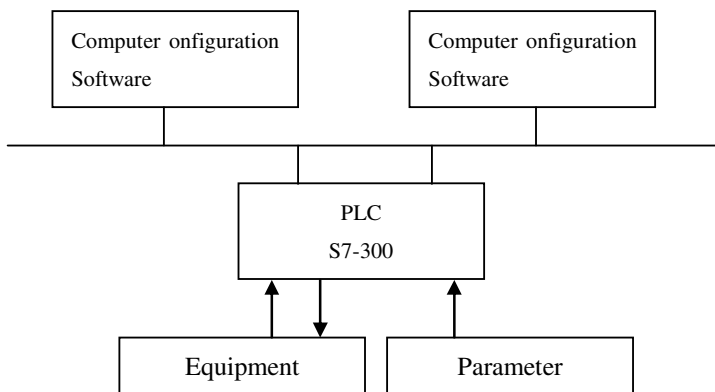


Fig. 1. Supervisory system structure

Reliability and stability of PLC is suitable for real-time control, easy to program, and can be used for instruction or ladder programming. Therefore, we use PLC to supervise electric primary section. In the choice of the PLC model and configuration, we must analyze the control object and the I/O points statistics, besides we should set aside reserve margin about 20%—30%, the system has expansion capabilities. After analysis the system control task, we choose the S7-300 series of product to meet the requirements of control system, S7-300 uses a modular structure, fast, and wide variety of modules can be combined to achieve control. Module selection for the PLC

system: CPU module—cpu315-2DP, which integrate a PROFINET interface, support TCP/IP protocol, and can be more easily and communication between other devices equipped with Ethernet. Digital input modules-SM321(DI32×DC24V); digital output module-M322(DO32×DC24V/0.5A); analog input module-SM331(DO32×DC24V/0.5A); power modules-PS307 10A; interface module-IM360/361.

Configuration software is some of the data acquisition and process control, special software, which is in the automatic control system monitoring layer-level software platform and development environment, flexible configuration to provide users with fast building automatic control system monitoring functions, general level of software tools. In order to supervise the equipment and parameters of electric primary section, we use the “king view” to realize the man-machine interface. The design steps: (1) Designed graphical interface, the establishment of PLC equipment; (2) Creating real-time database, so that test data points will be corresponded to the PLC data. Which include PLC substation supervisory data and command flag that written to the PLC configuration software; (3) Producing graphical interface for displaying work state and alarm state, and operation button; (4) Animation connection. The establishment of corresponding relationship between graphic and data will ensure its appearance changing with the data changing, and can be used for audible alarm. (5) Action script. When control the device, we should press the action button trigger mouse event, execute script, and mark the corresponding flag, etc.

There is a special driver to realize the communication between the configuration software and S7-300PLC, so the user need not know the difference between the communication protocol, and the PLC needs no programming and network settings, you can achieve the communication between those. The main problem that this system faced is how to design the PLC for soft redundancy when the work of PLC switched to spare CPU from the master CPU. In order to ensure the data switching, we must switch the configuration into the correct communication object. In order to complete this function, we should ensure the reasonable of the equipment in the configuration software design of graphical interface after provide the whole equipment. As the symbol of the redundancy, we design a data point in real-time database and the data type is “redundant switching sign”, under normal circumstances, the processing is 16 numbers and the point constant is 0, when the constant is 1, the configuration will switch to another IP address, and the data point constant will be cleared automatically.

4 Conclusions

The stability and reliability of electric primary section of the main wiring in substation affects the safety of the whole substation system directly. This paper design an electric primary section in 500kVsubstation, what's more, it meets the requirements besides reliable, economic, flexible. The design and the method for the main selection design is reference and guide for other voltage grade, and the design in this paper has the value of a wide range of application in theory and practice.

The goal of the design, which is the electric primary section in substation, are testing the application status of equipment and predicting the hidden accidents. On the other hand, it can help find the cause of the accident quickly, so we can minimize the loss. Commissioning and operation show that the design and configuration which is based on PLC and electric primary section in substation supervisory control system of configuration software, to achieve the status of equipment operation and monitoring, effectively improve the system's level of automation and increasing the automation level of the system. Also, the supervisory control software have so many functions, friendly interface, what's more, it's very easy to use.

References

1. Wang, M.-X., Zhang, Q.: Analysis on Influence of Ground Electrode Current in HVDC on Ac Power Network. *Power System Technology* 03 (2005)
2. Yuan, Q.-Y.: Present State and Application Prospect of Ultra HVDC Transmission in China. *Power System Technology* 14 (2005)
3. Su, H.-T., Qi, X., Wu, Y.: Study on Market Demand of UHVDC Power Transmission in China. *Power System Technology* 24 (2005)
4. Wu, G.-F., Lu, J.-Y., Shao, F.-Y.: Research on electromagnetic environment of the next voltage level of transmission system in China. *Electric Power* 06 (2005)
5. Shao, F.: Power Frequency Electromagnetic Introduction on DC Transmission Line Caused by a Parallel Proximate AC Transmission Line. *Power System Technology* 12 (1998)
6. Zhu, Y.-Y., Jiang, W.-P., Zeng, Z.-H., et al.: Studying on Measures of Restraining DC Current Through Transformer Neutrals. In: *Proceedings of the CSEE*, vol. 13 (2005)

Research on the Secure Access of E-Commerce

Longjun Zhang

Dept. of Communication Eng., Eng. University of C.A.P.F., 710086 Xi'an, China
longjun_zhang@sohu.com

Abstract. Currently, the security problem of E-commerce is more and more prominent. The security problem is one of the key problems of E-commerce and the security of E-commerce is becoming a bottleneck for its further applications. Based on the analysis on the security problem of E-commerce, this paper presents a secure access scheme of E-commerce. The security of the access scheme has the merits of high security, low computation complexity and is easy to realize. The access scheme has the certain theories meaning and realistic meaning for the further expansion and application to E-commerce.

Keywords: Secure, access, e-commerce.

1 Introduction

With the increasing development of social and economic activities, E-commerce has been stepping into practical applications gradually. The development potential of E-commerce is very captivating, the security problem of E-commerce, however, is more and more outstanding [1]. E-commerce is one of the most important trends of Internet application, and one of the important business models in the modern international business activities. E-commerce has formed a significant part of our economic life. The security is a key issue which assures healthy development of E-commerce, and still is a bottleneck problem in E-commerce application currently. A secure and efficient access scheme becomes a problem that needs to be resolved urgently and plays a significant role in the development of E-commerce[2]. At present, there are many different security resolutions for E-commerce[3]~[5]. However, they are still one way or another some of the shortcomings. Accordingly, we need to design a new secure access scheme for E-commerce.

2 Analysis on the Secure Problem of E-Commerce

2.1 Defect Analysis of E-Commerce Security

Two aspects of the E-Commerce security must be considered.

i) Protect the transferable message;

ii) Avoid the active attack of adversary against the network system, such as modification or fabrication of the information. E-commerce activities are based on Internet. The insecurity of Internet is mainly due to the secure bug of itself. The

secure bug of Internet itself includes the insecurity of Internet operations and Internet protocols. The opening of Internet indicates that operations are based on the open protocols and that anonymous remote access makes all kinds of attacks more convenient. On Internet, anonymous remote access is the main access mode. At such instance, the security of system itself and the security of the information transmitted by system are all facing serious threat. Unlawful user may attack system through two modes: passive mode and active mode. At the passive attack mode, the unlawful user does not interfere with the information flux between the system and the lawful users. He just attempts to pick-up the useful information via interception, etc. With the cryptograph, the unlawful user cannot obtain the plaintext. However, he can obtain the helpful information by analyzing the flux of correlative data. On Internet, unlawful users often take the active attack mode. He can arbitrarily intercept and modify the data at network layer, even delete original data and replace with imitative data. For pretending his identity as a lawful user, the unlawful user intercepts and operates the information that system sends to a lawful user. These two kinds of attacks often occur on Internet.

2.2 The Key of E-Commerce Secure Access

Commonly, the access security of E-commerce can be ensured efficiently by using the cryptosystem technology such as authentication. The authentication is a key issue which helps E-commerce system effectively avoid to attacks. At E-commerce, the most serious secure bug is unlawful access system. That is, the unlawful user accesses the E-commerce system by cheating or breaking the password of lawful users. To avoid the imitation, authentication is used for validating that the user is lawful when the both sides in network build link or transmit information. The practical authentication system also needs to avoid the sender and the receiver cheating each other. The secure authentication plays an important role in the access of E-commerce. The authentication service is an indispensable key of E-commerce access, the authentication assure the authenticity of that an entity claimed.

3 Construction the Secure Access Scheme of E-Commerce

The cryptosystem technology is currently the base and support of E-commerce security mechanism. Generally, the secure access scheme of the E-commerce is based on public key infrastructure (PKI). The secure public key cryptogram algorithm, digital signature algorithm and authentication scheme are the base and important guarantee of the secure mechanism in an E-commerce system. Some access schemes for E-commerce are based on the conventional cryptosystem, for instance RSA, but some leak has been reported in certain literature. Some possible troubles may exist due to the weak security of the encrypt algorithms or digital signature applied. There is some secure threat to the RSA public key cryptosystem. Notwithstanding that the most popular public key technology adopts RSA cryptosystem. The advanced cryptosystem-elliptic curve cryptosystem (ECC) is acknowledged worldwide

gradually and becomes an interesting field of cryptosystem. The research and application of ECC in network security and E-commerce security have obtained a great success. Since popular RSA is not able to satisfy the requirements of E-commerce security, the application of ECC in E-commerce is imperative.

The ECC can be used for encrypt/decrypt of information, constructing digital signature and designing the secure authentication protocol. Since the ECC is more secure than RSA, it may replace the RSA to become commonly used cryptosystem. The constitutor of the famous SET (Secure Electronic Transactions) protocol have taken ECC as the next generation default asymmetric cryptosystem algorithm and some international standardization organizations have taken ECC as new information secure standard.

The secure access scheme of E-commerce in this paper is composed of authentication center AC, it is the trusted third party, server S and client C. Client C communicates with server S by Internet. The server S is mainly used to provide service for E-commerce users to access E-commerce system resource and acts as proxy, through which users request authentication to AC. It is the key device to access E-commerce system. The authentication center AC collects and manages the related information of E-commerce users, accesses to the server, authenticates the E-commerce user who sends access request to access E-commerce system and provides different access authority in accordance with different levels of users. The structure of secure authentication system is client-server, with sever S as client end and the authentication center AC as server end. Such structure provides perfect distribution resolution for secure authentication of E-commerce. Authentication system can put the key information of all E-commerce users and AC in a single and centralized database rather than distributing them in various equipments. It is much easier to insure security if the authentication server operates on a single secure operating system. In the authentication system of E-commerce, authentication and access accreditation of E-commerce users are separated from the practical access. Moreover, important information is saved in a centralized database. In this way, the security and flexibility is improved effectively.

This secure access scheme of E-commerce demand that every actor in the E-commerce system can communicate by Internet. In order to avoid the important plaintext being intercepted by unlawful users, the secure access scheme of E-commerce is based on ECC, which can encrypt/decrypt and digital signature by ECC. With this scheme, AC can implement ECC encrypt/decrypt and signature operation, which can be used to authenticate the identity of and the integrity of information from the opposite transaction party, and provides evidence for the transaction behaviors. S needs to authenticate the identity of C and authenticate the integrity of information sent by C with authentication system. Meanwhile, C needs to authenticate the identity of S and authenticate the integrity of information sent by S with authentication system as well. So the secure authentication scheme can provide mutual authentication and undeniable evidence for transaction behaviors of the participants.

4 Design on the Secure Access Scheme of E-Commerce

4.1 The Secure Access Scheme of E-Commerce

The access process of E-commerce is described as follows:

i) System initialization:

Step 1. Select a public base point $P \in E(F_q)$ of E , its order is $l = \text{ord}(P)$. E is a non-supersingular elliptic curve.

Step 2. Client C selects a random integer $d_C \in \{0, l\}$ as secret key, then computes $e_C = d_C P$ and takes e_C as its public key. The client C has elliptic curve digital signature algorithm Sig_C .

Step 3. E-commerce system server S selects a random integer $d_S \in \{0, l\}$ as secret key, then computes $e_S = d_S P$ and takes e_S as its public key. Access server S has elliptic curve digital signature algorithm Sig_S .

Step 4. E-commerce authentication center AC selects a random integer $d_A \in \{0, l\}$ as secret key, then computes $e_A = d_A P$ and takes e_A as its public key. AC has the elliptic curve digital signature algorithm Sig_A .

ii) Authentication

a) Register

The user must register in system authentication center at first. The exact process is as follows:

Step 1. Firstly, C generates the identity V , which consists of his name and password, then selects a random integer r_1 and computes $M = e_A(r_1 V \bmod l)$.

Step 2. C sends the information R that request qualifies access E-commerce system to AC and signs with ECC digital signature algorithm, then computes $C_R = Sig_C(R)$ and sends (C_R, M) to AC .

Step 3. AC verifies the signature to (C_R, M) in order to detect the information sent by C or the intrusion.

Step 4. AC selects a random integer $k \in \{1, \dots, l-1\}$, then computes $Q = kP$ and sends to C by signing with elliptic curve digital signature algorithm;

Step 5. After verifying the signature of AC , C selects another random integer $r_2 \in \{1, \dots, l-1\}$, and computes: $Q' = r_2 Q = (r_x, r_y)$ and $r = (r_1 r_x) \bmod l$, then sends r to AC after doing digital signature.

Step 6. AC obtains $(r_1 V \bmod l)$ by decrypting M and generates an only identity ID for C , and then computes: $C_{ID} = S_A(ID)$, $y = k^{-1}(r_1 V d_A + r) \bmod l$ and $A_{ID} = e_C(C_{ID} | y)$. Then AC sends A_{ID} to C .

Step 7. After receiving A_{ID} , C decrypts it and obtains C_{ID} and y , then computes: $M' = (r_1^{-1} r_2^{-1} y) \bmod l$.

b) Access E-commerce system

When users want access E-commerce system, that is to say client C sends request access information to sever S , E-commerce system firstly asks C to input lawful identity. C sends his identity, which is obtained from system authentication center AC , to access sever S after encrypting via the public key of S . The access information coming from C includes authentication symbol (V, M') . Sever S picks up

authentication information, then encrypts the request information of C by the public key of AC and signs with the ECC digital signature algorithm, namely server S computes $S_R = \text{Sig}_S(R)$. Subsequently it puts the digital signature and the cryptosystem of request information in the request data package and sends to AC for authentication.

Authentication center AC verifies if S_R is the digital signature of server S by the public key of S, that is by computing $e_S(S_R)$ to make sure that the information is sent by S assuredness. To ensure further security, AC also sends the Access-Challenge response to S. The response information value of Access-Challenge package and the request information value of Access-Request package must be same in order to indicate that they are a pair of request and response. AC authenticates the request information of C that is sent by S after decrypting it with its secret key. Otherwise, it shows S to be an unlawful user. AC will send Access-Reject response to S and refuse the operating request that come from S.

AC verifies the request information of C sent by S to verdict the identity of C is lawful or not, i.e., to compute $M'Q' = Ve_{A+r_x}P$. If the equation is true, it will show that C is a lawful user and allow it accessing system. Otherwise, C is an unlawful user and is refused to access system. AC authenticates digital signature on the result with the public key of S and sends it to S.

After receiving the response package from AC, S decrypts the response package with its secret key and authenticates on it with the public key of AC to confirm the response package is sent by AC or not. If S receives the Access-Reject package, it shows the access request of C is refused. On the contrary, if S receives the Access-Accept package, it shows that the access request of C is accepted.

4.2 Analysis on the Security of the Scheme

The security of access scheme is the important base of E-commerce system, and it is very important to make sure that it can run securely. Although there are very few messages transmitted in access process, each message is designed elaborately. There are complicated interaction and restriction within these messages. It is very likely to make a mistake in design of access scheme, which may result in huge hidden trouble to security. So it is essential to analyze authentication protocol in order to estimate that if it reaches the expected secure goal.

In the secure access scheme which presented by this paper, during the access authentication process, when the data package is transferred among AC, S and C, the ECC digital signature algorithm is adopted, so the identity of information sender and information integrity will be ensured, and it can avoid hacker cheating, modifying or fabricating information. If a hacker wants to cheat or destroy the integrity of information, he has to break the ECC digital signature algorithm. But it is impossible to realize up to now. Therefore this scheme can guarantee the security of E-commerce. To avoid unlawful interception of important plaintext, the important information, such as password etc, must be encrypted by ECC when the data package is transferred among AC, S and C.

The security of this secure access scheme is ensured by the security of ECC, and the security of ECC is based on the difficulty of the elliptic curve discrete logarithm problem. Because there has not been effective solution of elliptic curve discrete logarithm problem yet, the secure access scheme of E-commerce in this paper is

secure. After lawful users dispose the request information and identity by use of ECC encrypt/decrypt algorithm and ECC digital signature algorithm, the unlawful user cannot fabricate the request information and ID of lawful user. Since he cannot resolve the elliptic curve discrete logarithm problem when he does not know the secret key of lawful user, the E-commerce system does not accept his access request. Similarly, because of the difficulty of elliptic curve discrete logarithm problem, the unlawful user cannot obtain the content of the communication information, and cannot do flux-analysis, retransmission, modification or fabrication etc. on the information that he intercepted.

5 Conclusion

This paper introduced the next generation public key cryptosystem ECC into the access scheme of E-commerce, and designed a secure access scheme based on secure authentication to improve the security of E-commerce system. The main consideration in the secure access scheme of E-commerce design is security, transparency to users, easy management and extensibility. The secure access scheme of E-commerce in this paper stops the secure bug frequently occurred in the secure access problem of E-commerce. It is secure and easy to operate and realize. In conclusion, this scheme definitely has theoretic and practical significance to E-commerce.

Acknowledgement. The work is supported by the Basic Research Foundation of Engineering University of C. A. P. F (No.WJY201111).

References

1. Zorkadis, V., Karras, D.A.: Security modeling of electronic commerce infrastructures. Information Systems for Enhanced Public Safety and Security, IEEE/AFCEA (5), 340–344 (2000)
2. Wu, Z.-G.: Security and Atomicity in Electronic Commerce: Model, Protocol and Verification. Journal of Software 3(12), 329–333 (2001)
3. He, W.H., Wu, T.C.: Security of the Tseng integrated schemes for user authentication and access control. IEE Proceedings: Computers and Digital Techniques 147(5), 365–368 (2000)
4. Hui, L.: Server-independent password authentication method for access-controlled Web pages. In: Prof. of IEEE Int. Conf. on Global Telecommunications (GLOBECOM 2000), vol. 1, pp. 361–364, 61–63 (2000)
5. Shor Molly, H.: Remote-access engineering educational laboratories: who, what, when, where, why, and how? In: Proc. of the IEEE American Control Con., pp. 2949–2950 (2000)

A Secure Authentication Scheme of Wireless Mesh Network

Longjun Zhang

Dept. of Communication Eng., Eng. University of C.A.P.F., 710086 Xi'an, China
longjun_zhang@sohu.com

Abstract. This paper addresses the security of WMN which receives little attention thus far. First analyze the unique security requirements of WMN and an authentication based on threshold method. Then propose a new authentication scheme using several authenticate servers. The result shows the computation for the client is comparatively small. It achieves mutual authentication between client, AP and AS and prevents the server-compromise effectively.

Keywords: Authentication, compromise-resilient, wireless mesh network.

1 Introduction

The use of wireless mesh network (WMN) technology to provide internet connectivity is becoming a popular choice for wireless Internet service providers because the technology provides fast, easy and inexpensive network deployment. WMNs represent a unique marriage of the ubiquitous coverage of wide-area cellular networks with the ease and the speed of local-area Wi-Fi networks. However, without a solid security solution, WMNs with open medium won't be able to succeed [1], [2]. To overcome the disadvantage, authenticate server group with several servers is adopted in other paper [3].

Currently, there is no standard method to authenticate a network point in WMN, the security architecture for WMN is still an open issue. Recently, there are lots of papers proposing schemes to authenticate the client, Fu et al. [4] present a new authentication scheme based on a combination of techniques, such as zone-based hierarchical topology structure, virtual certification authority (CA), off-line CA, identity-based cryptosystem and multi-signature. Santhanam et al. [5] propose a network layer authentication mechanism based on Merkle Tree. All these paper just addresses the authentication of the client, without considering the whole security architecture and the users' roaming between different domains. Zhang et al. [6] propose the ARSA architecture which can acquire lost of attack-resiliency. In their architecture, users acquire a universal pass from a third-party broker whereby to realize seamless roaming across WMN domains administrated by different operators; the network model contains brokers, domain operators and users, while the trust model assumes that brokers are fully trustable by both clients and operators, but a client and an operator usually do not completely trust on each other.

2 Related Works

In this part, we briefly analyze one authentication scheme based on threshold [4], which also uses Authenticate Server Group (ASG), similar to our scheme.

Step1: Mobile User (MU) receives certificates of each server from CA, also one symmetrical key K and its shares K_i . Then MU encrypts K_i using public key of each server. This cost n times public key encryption. t or more that t server in ASG recover K and send K to AS-1 (communicating server to MU directly).

Step2: MU constructs the information $R=(R_1||R_2||\dots||R_t)$. The information is first encrypted with K_i , then the ciphertext will be encrypted with the public key of t servers and sent to t servers. Through respective share and secret key, the t servers can recover each R_i . Further, t servers jointly recover R and send to AS-1.

Step3: MU encrypts R with K and public key of AS-1 successively. Then send the result to AS-1.

Step4: AS-1 decrypt the ciphertext with K and its secret key. The access request is permitted if $R'=R$, otherwise denied.

In this scheme, MU must execute public key encryption for $n+t$ times. When the n value is large, it will cost much computation for the resource limited client. Access Point (AP) is assumed to be safe and always trustable. In practice, AP is exposed outside and wireless connected. So it is easy to be captured and the secret information could be acquired or replaced by attackers. Our scheme is proved to be more efficient and practical since it needs less computation for mesh client besides assuming AP could be attacked. It achieves mutual authentication among MU, ASG and AP.

We denote by MU and MP the mobile user and mesh router respectively. ASG is the authenticate server group. $E_k(\)$ mean the symmetrical encryption under the key k .

3 Preliminaries

3.1 Authentication Issues of WMN

Mutual authentication: In WMN, mesh routers are wireless connected and exposed in public area. Hence the router is easy to be captured, removed, replicated and so on. It is important to achieve authentication among client, router and authentication server. However, it needs to access rapidly and safely.

Multi-domain roaming authentication: Unlike conventional wireless network, the future wireless mesh network comprised of multi domains which cover small area. Each area is admined by one independent operator, reauthentication occurs when client move or roam between domains. The conventional settlement for roaming to set up Service Level Agreement does not adapt to WMN.

1) Efficiency

The client in WMN differs depending on their types. It could be normal PC, notebook, laptop or some powerful terminal. It may be PDA, WiFi cell phone or some other terminal with limited resource. The authentication delay and computation cost are crucial for these clients.

2) Hybrid network

WMN is a hybrid network. Via the backbone network, it can join Internet, WLAN, WSN and other network. In different network, the cryptosystem is various. The authentication of clients from different cryptosystem is complicated.

3.2 Security Acquirements of WMN

We assume the wireless backbone, the wired APs and any other WMN operator equipments, as the infrastructure, which is comparatively static. We also use the term “mesh domain” to indicate a subnet comprising a set of mesh routers and their covered mesh clients. One mesh domain is administered by one operator. The security of WMN includes two types as below [6]:

1) Infrastructure security

The security of signalling and data traffic transmitted over the infrastructure.

2) Network access security

The communication security between in client and mesh router. It may also involve the communication security among mesh clients served by the same mesh router, if the route between a client and a router is in multiple hops.

With respect to network access security, we limit ourselves to study the following specific requirements:

1) Router–client authentication and key establishment

A mesh router should authenticate a requesting client from its own domain or other roaming users. The client should also authenticate the router because the router is exposed and easy to be personated. After authenticated, mesh router and client should establish a shared key to secure the traffic between them.

2) Client–client authentication and key establishment

Client should communicate with the legitimate one. This is required authentication between clients. If needed, the communicating clients may need to establish a shared key to secure their information. The network consists of two entities: the mesh clients and the operator. Client buys network access service from the operators like conventional network without changing existent mode.

Large scale WMN consists of domains which are composed of meshes. One mesh includes one mesh router and its communicating clients. Each domain is administered by an operator. The operator uses PKI to secure the infrastructure. The mesh router has much more powerful computation and communication capacities than regular mesh clients. The downlink from the router is one hop, while the uplink to the router from clients may be one hop or multiple hops.

4 Authentication and Key Agreement

4.1 System Initialization

Let TA be the authority in WMN. The system parameter initialization is as follows:

1) TA choose (F_q, E, P, l, h) , F_q is one finite field, q is a large prime. E is an elliptic curve over F_q , P is the base point in E . The order of p is $l \geq 160$.

2) TA randomly pick $k_{TA} \in F_q$ as its secret key and computes $P_{TA} = k_{TA}P$ as its public key. $H(\cdot)$ is one hash function collision-resilient.

3) TA randomly generates a secret polynomial function $t-1$ degree over F_q as follows: $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_{t-1}x^{t-1}$, and generates secret shadows $d_i = f(ID_i)$ ($i=1, 2, \dots, n$) and corresponding public keys.

4) TA choose randomly $s \in F_q$ as the secret key of ASG, $a_0 = s$. Correspondingly, the public key of ASG is $Y_{ASG} = sP$. Let ID_i be the exclusive identity for the servers in ASG, $ID_i \neq ID_j$.

5) TA sends the secret shadows to respective AS_i . The public key of AS_i is $Y_i = d_iP \in E(F_q)$. AS_i ($i=1, 2, \dots, n$) randomly choose two large primes p_i, q_i and keep p_i, q_i, d_i as secret information.

6) MP pick two random large primes p_a and q_a , $p_a \equiv q_a \equiv 3 \pmod{4}$. It calculates $n_a = p_a q_a$, p_a and q_a is the secret key of MP, n_a is the public key.

7) MP constructs $M = ID_a || n_a$, ID_a is the identity of MP. Then it signs M with the scheme in paper [7]. MP randomly pick a integer w ($l < w < l-1$), calculates $wP = (x, y)$, $r = x \pmod{l}$, $e = H(M)$, $\sigma = t + erk_{TA} \pmod{l}$. Then the signature is (M, r, σ) .

4.2 Registration of MU

1) MU chooses a random number r_{MU} , submits its ID and r_{MU} to TA in secure channel.

2) TA checks whether ID is legitimate. After checking the validity, each server AS_i does as follows:

AS_i computes the value of V_i , $V_i = ID \pmod{p_i}$, $V_i = r_{MU} \pmod{q_i}$. By using CRT [8], each server AS_i can obtain V_i by $V_i = r_{MU} b_i q_i + ID b_2 p_i \pmod{n_i}$, where $n_i = p_i q_i$, $i=1, 2, \dots, n$.

3) Each server only keep the value of V_i and put them in a table. ID and random number r_{MU} could not be obtained by attacker because he can't get p_i and q_i from the V_i table, MU obtains public key P_{TA} from TA.

4.3 Authentication

1) $MP \rightarrow MU$

MP periodically broadcast the beacon including the system parameters ($P_{TA}, n_a, ID_a, M, r, \sigma$). After receiving the beacon, MU checks whether P_{TA} is identical from the one he got from TA. Then he verifies the signature as follows:

$$E = H(M), u = er \pmod{l}, \sigma P - uP_{TA} = (x', y'), r' = x' \pmod{l}$$

If $r' = r$, the signature is valid. The system parameters can be accepted and go to 2) or abort the authentication.

2) $MU \rightarrow MP$

MU randomly picks two random numbers: k_{s1} and m . MU calculates $c = (ID || r_{MU} || k_{s1} || m || T)^2 \pmod{n_a}$, where k_{s1} is part of the session key, m is the message to be signed by t servers when access request is permitted, T is the timestamp. MU submits c to MP.

3) $MP \rightarrow MU$

MP decrypts c with its own secret key p_a and q_a , check the validity of ID and T. If ID and T is validate, MP send $(ID || r_{MU} || m)$ to ASG in a pre-established secure channel.

4) $AS_1, AS_2, \dots, AS_t \rightarrow MP$

More than t servers in ASG calculate V_i from MP's ID, r_{MU} according to 2). Then they search the V_i table to match the corresponding V_i . If found, the t servers AS_i respectively sign the message m :

Step1: Let the t servers be AS_1, AS_2, \dots, AS_t . They pick up their own random integers $k_i, i=1, 2, \dots, t$, and $k_i \neq k_j, k_i = k_i P$. Then they send respective k_i to MP.

Step2: MP calculates $k = \sum_{i=1}^t k_i = (x_1, y_1), r_1 = x_1 \bmod l$. MP submit r_1 to AS_1, AS_2, \dots, AS_t .

Step3: AS_1, AS_2, \dots, AS_t do as: $e' = H(m), \sigma_i = k_i + e' r_1 d_i C_i \bmod l, C_i = \prod_{j=1, j \neq i}^t (-ID_j / ID_i - ID_j) \bmod l$. AS_1, AS_2, \dots, AS_t send the subsignature $\sigma_i (i=1, 2, \dots, t)$ to MP.

5) $MP \rightarrow MU$:

MP derives the signature $\sigma' = \sum_{i=1}^t \sigma_i, i=1, 2, \dots, t. e' = H(m), \sigma' P - e' r_1 Y_{ASG} = (x_1', y_1'), r_1' = x_1' \bmod l$. If $r_1 = r_1'$, MP consider the access request is permitted by ASG. MP pick a random number k_{s2} as part of the session key. MP allows the access request and send MU message: $d = Ek_{s1}(ID || k_{s2})$. MU decrypt message d using the key k_{s1} . The final session key is $ks = H(ID || k_{s1} || k_{s2})$. The authentication is finished.

4.3 Security and Efficiency Analysis

1) Security Analysis

The scheme is able to resist the replay attack. Because the access request include the timestamp T . Attackers can't impersonate the legal users without acquiring the timestamp. This scheme is to resist the compromise event. Assuming less than t servers is compromised or captured. They tend to fabricate the signature in order to make MP accept it. We prove that attacker couldn't succeed in fabricating even attacker collect $t-1$ subsignatures. Supposing MP receives $t-1$ correct $k_i, i=1, 2, \dots, t$ and one fake k_i' . $k' = \sum_{i=1}^{t-1} k_i + k_i' = (x_a, y_a), r_a = x_a \bmod l$. The fake subsignature $\sigma_i' = k_i' + e' r_a d_i' C_i' \bmod l$, where d_i' and C_i' is the corresponding secret shadow and coefficient. When MP verifies the signature, $\sigma' = \sum_{i=1}^t \sigma_i + \sigma_i' \bmod l$, according to the equation: $\sigma' P - e' r_a Y_{ASG} = k' + e' r_a (\sum_{i=1}^t d_i C_i + d_i' C_i') P - Y_{ASG} \neq k'$. So the fabricating fails.

2) Efficiency Analysis

In our scheme, MU verifies the improved ECDSA signature to achieve authentication with MP. The improved ECDSA cost much less computation than the original one. Table 1 shows the computation cost for MU, MP and each AS_i in the authentication process. We denote T_{H0} for the time hash function takes. Similarly, T_x indicate multiplication in modulo. T_+ indicate addition in modulo. T_{E+} indicate addition over ECC. T_{Ex} indicate multiplication over ECC. T_{EXP} indicate exponentiation.

Table 1. Computation Cost

MU	$T_{H0} + 2T_x + 2T_{E+} + T_{Ex}$
MP	$T_{EXP} + T_{H0} + 2T_x + (t-1)T_{E+} + 4T_{Ex} + (t-1)T_+$
AS_i	$T_{H0} + 2T_x + T_{Ex} + 2T_+$

5 Conclusion

In this paper, we describe some authentication issues and security requirements of WMNs. Then we analyzed an authentication scheme based on threshold. Finally we present our new authentication scheme. After that, we present a brief security and efficiency analysis and end with conclusions. Through analysis, it proves that our scheme is efficient to resist the compromise.

Acknowledgement. The work is supported by the Basic Research Foundation of Engineering University of C. A. P. F (No.WJY201111).

References

1. Akyildiz, I., Wang, X., Wang, W.: Wireless mesh networks: A survey. *Comput. Netw.* 47(4), 445–487 (2005)
2. Ben Salem, N., Hubaux, J.P.: Securing Wireless Mesh Networks. *IEEE Wireless Communications* 13(2), 50–55 (2006)
3. Yang, Y., Gu, Y., Tan, X., et al.: A New Wireless Mesh Network Authentication Scheme Based on Threshold Method. In: *Proceeding of The 9th International Conference for Young Computer Scientists* (2008)
4. Fu, Y., He, J., Wang, R., Li, G.: Mutual Authentication in Wireless Mesh Networks. In: *Proceedings of ICC 2008* (2008)
5. Santhanam, L., Xie, B., Agrawal, D.P.: Secure and Efficient Authentication in Wireless Mesh Networks using Merkle Trees. In: *33rd IEEE Conference on Local Computer Networks, LCN 2008* (2008)
6. Zhang, Y., Fang, Y.: ARSA: An Attack-Resilient Security Architecture for Multihop Wireless Mesh Networks. *IEEE Journal on Selected Areas in Communications* 24(10) (October 2006)
7. Yang, J., Dai, Z., Yang, D., et al.: One ECC signature scheme and protocol based on identity. *Software Transaction* 11(10), 1303–1306 (2000)
8. Mao, W.: *Modern Cryptography: Theory and Practice*. Prentice Hall PTR, Upper Saddle River (2003)

Research on Capacity Allocation in a Supply Chain System Based on TOC

Kaijun Leng¹ and Yuxia Wang²

¹ School of Logistics and Engineering Management, Hubei University of Economics, 430205, Wuhan, China

² Department of Information and Engineering, Wuhan University of Technology Huaxia College, 430223, Wuhan, China
Kaijunlen@yahoo.com.cn

Abstract. Enterprise optimization can rapidly strip significant “bottom line” costs out of global operations, giving companies a real competitive edge. The benefits of managing supply chain networks by integrating operational, design and financial decisions have been acknowledged by the industrial and academic community. In this paper, we consider a supply chain in which one supplier sells to multiple retailers and suppose that the sum of retailer orders exceeds the supplier’s fixed capacity. To balance supply and demand, the supplier must employ an allocation mechanism, an algorithm for converting an infeasible set of orders into a feasible set of capacity assignments, we proposed a optimal allocation mechanism based on Theory of Constraints(TOC) in face of meeting peak demand in certain period for the whole system, and Genetic algorithm(GA) has been selected in solving the optimal model in our work. Furthermore, a numerical example has been implemented to demonstrate the efficiency of our method.

Keywords: Supply chain management, capacity allocation, Theory of constraints.

1 Introduction

Generally speaking, a supply chain system includes all the interactions among suppliers, manufacturers, distributors, and customers[1]. The integration and synchronization of information and material flows of manufacturing sites in the context of supply chain (SC) has become more practical and has attracted the attention of both industry practitioners and academic researchers. SCM generally deals with strategic and operational problems. By contrast, most of the supply chain literature uses stochastic models to better understand inventory control issues at a strategic level. Deterministic supply chain scheduling models are an interesting emerging area for this type of research. Unfortunately there exist relatively few papers that deal specifically with scheduling problems in the context of supply chain operations[2].

Consider a supply chain in which one supplier sells to multiple retailers and suppose that the sum of retailer orders exceeds the supplier’s fixed capacity. To balance supply and demand, the supplier must employ an allocation mechanism, an algorithm for converting an infeasible set of orders into a feasible set of capacity assignments.

In this paper, we investigate a two echelon supply chain system consist of multi-manufacturers and multi-distributors, we proposed a optimal production allocation mechanism based on Theory of Constraints (TOC) in face of meeting peak demand in certain period for the whole system., The TOC-based strategy is now being implemented by a growing number of companies[3][4]. The performance reported by the implemented companies includes reduction of inventory level, lead-time and transportation costs and increasing forecast accuracy and customer service levels[5]-[11]. As the programming model we set up for the system is a NP-hard problem, Genetic algorithm (GA) has been selected in solving the optimal model in our work.

During the last decade, there has been a growing interest using genetic algorithms (GA) to solve a variety of single and multi-objective problems in production and operations management that are combinatorial and NP-hard [12]-[15]. This approach leads to a hybrid evolutionary algorithm in which the GA constitutes the core of the search strategy, while multiple heuristic rules called in specific circumstances contribute to reconstruct a feasible solution that satisfies all the constraints and objectives.

Furthermore, a numerical example with practical data has been implemented to demonstrate the efficiency of our method.

2 Description and Formulation of the Problem

The model we investigate in this paper is a two-echelon production-distribution supply chain system consists of M manufacturers and N retailers. Normally, we assume that $M > N$.

In the system, all manufacturers are belonged to the same enterprise which can not only produce products with its own brand but also capable of dealing with the ODM (Original Design Manufacture) order for other companies, so each of the these manufacturer may have the choice to either focus on the internal business or other manufacturing outsourcing services, and each manufacturer pursuit the maximization of it's own performance, however, still they have to follow the direct instruction from the headquarters of the company, the limitation of the manufacturers capacity will not be considered in the model. We assume that different manufacturer produce certain product of different type, hence, we note $i = 1, 2, \dots, t, \dots, r, \dots, w, \dots, M$ for the manufacturers which produce three types of products which are A, B and C separately, and W_i stand for the maximum capacity of the manufacturer.

Meanwhile, the manufacturing enterprise distributes its product via certain distributors. In regular sales seasons, the distributors order the product from the manufacturers, we note them as $D_j, j = 1, 2, \dots, N$, and the production schedule are made based on the sales information provided by the distributors, and it's obviously that the accuracy of information are far more different from each other among those distributors for the profit consideration of their own. The inaccuracy information would lead to overstock for the manufacturers and heavily jeopardize the profitability of whole supply chain system. However, in particular sales period such as Christmas or spring festival, shortage of products would appear inevitably, even with overtime working strategy for the manufacturers, it is still hardly satisfy the peak demand in the period due to the constraints of funds available F and production capacity, although the distributor believe that the more inventories they have the higher profit they would

achieve. Hence the main issue in this period is how to optimal the production allocation to maximize the profit of the entire supply chain system.

Though the amount of the order is often considered as a main factor in priority decision, the information accuracy remains a leading concern for the manufacturer. Thus, we proposed an allocation mechanism based on TOC which use the information accuracy level as a buffer to schedule the production allocation in the system as shown in Fig1.

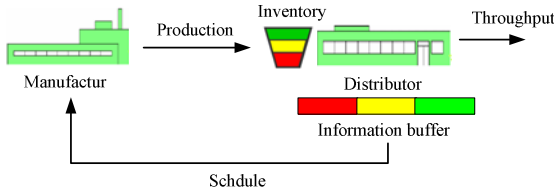


Fig. 1. Allocation mechanism based on TOC

In contrast with the amount-oriented schedule, TOC-based mechanism emphasis the information accuracy of each distributor in daily operation. With different information accuracy level U_j , the higher value of the index the more product those distributor could have. By doing so, it will not only be a schedule mechanism but also an incentive for the whole supply chain system to be more cooperative.

The objective function of our optimal model is to maximize the profit of the supply chain system:

$$\max \sum_{j=1}^N U_j \left[1 - \prod_{i=1}^M (1 - p_{ij})^{x_{ij}} \right] \tag{1}$$

In which p_{ij} stands for the possibility of the demand be satisfied when one batch of overtime production work be allocated to a certain manufacturer for a certain distributor.

According to the TOC-based mechanism, it is necessary to satisfy minimum requirement of each distributor:

$$\sum_{i=1}^M x_{ij} \geq U_j \cdot \sigma \cdot \left(\sum_{i=1}^M W_i \right) \tag{2}$$

σ is the flexible coefficient that the production capacity been put into use for the basic needs of the distributors, and x_{ij} represent the number of batches of overtime work be allocated to manufacturer i for distributor j . While the amount-oriented schedule focus on the amount of distributors need:

$$\sum_{i=1}^M x_{ij} \geq \text{Max}\{D_j, D_{j-1}, \dots, D_1\} \tag{3}$$

However, the schedule process has been limited to certain constraints, first of all is the funds that could be put into production, which often be used in material procurement, labourage etc.:

$$\sum_{j=1}^N \sum_{i=1}^M c_{ij} x_{ij} \leq F \beta_{\theta}, \theta \in [A, B, C] \tag{4}$$

c_{ij} is the overtime production cost of each batch that manufacturer i be allocated to distributor j ; β_{θ} stands for the proportion of each kind of product that satisfy the distributors basic needs. And of course the schedule process will also be limited by the capacity constraint:

$$\sum_{j=1}^N x_{ij} \leq W_i \tag{5}$$

3 A GA Based Solution Procedure

Model (1) is a typical target allocation mathematical problem which has been researched by many scholars. And various methods have been proposed in various literatures [16]-[18]. We choose an improved genetic algorithm in our research. For the problem under study, binary system is selected in chromosome coding, meanwhile the proper chromosome structure has been selected, each chromosomal is a $M \times N$ matrix, and each gene x_{ij} represent the number of batches of overtime work be allocated to manufacturer i for distributor j .in the schedule process.

Different from the traditional GA, we set a brand new race quality examination and rectify procedure in the GA program, which consists of three modules, as shown in Fig2.

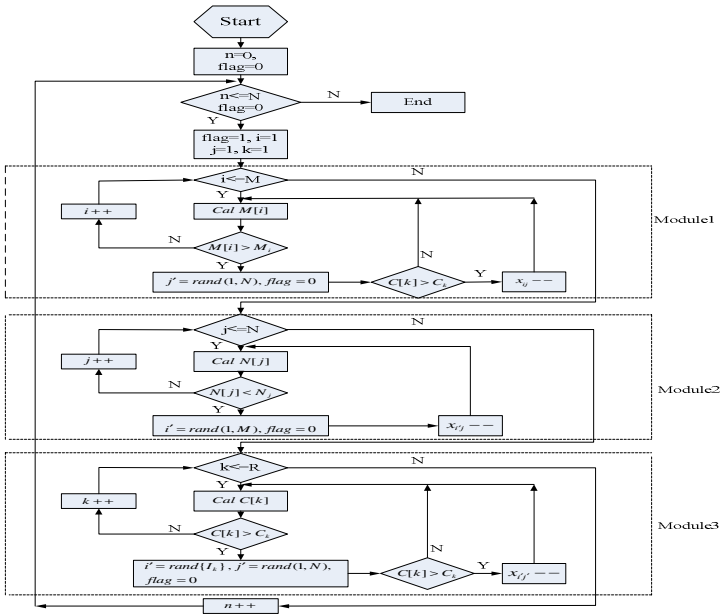


Fig. 2. Structure of the race quality examination and rectify procedure

Then following the traditional GA procedure, we shall have the optimal production allocation schedule of the supply chain system.

4 Numerical Example

In this section we use practical data as numerical example to demonstrate the efficiency of our proposed model. The data were provided by a traditional manufacturing SME (small and medium enterprise) located in Henan province, China. The company owns seven manufacturing plants located in different places and each plants pursuit its own profitability although they have to follow the direction of the headquarters of the company.

The company distributes its products 4 main distributors, and by the year 2009, it owns 15 plants which produce 3 main goods, 7 produce A, 2 for B, and 6 for C. When face the peak demand sales period, the production funds in hand is $F = 7500000$ yuan, and the proportion for each kind of product is $\beta_A = 55\%$, $\beta_B = 15\%$, $\beta_C = 25\%$ separately. The possibility of demand satisfaction of each production allocation process is given in Table1:

Table 1. Possibility of satisfaction for the peak demand

	A1	A2	A3	A4	A5	A6	A7	B1	B1
S1	0.31	0.35	0.19	0.07	0.13	0.24	0.09	0.39	0.24
S2	0.15	0.24	0.04	0.26	0.19	0.20	0.21	0.17	0.31
S3	0.25	0.12	0.03	0.19	0.25	0.22	0.18	0.47	0.14
S4	0.29	0.07	0.19	0.13	0.05	0.19	0.28	0.09	0.43
	C1	C2	C3	C4	C5	C6			
S1	0.1	0.12	0.32	0.32	0.23	0.29			
S2	0.23	0.11	0.23	0.1	0.12	0.07			
S3	0.13	0.23	0.19	0.28	0.13	0.15			
S4	0.35	0.09	0.21	0.02	0.03	0.32			

And other production cost and information accuracy level are shown in Table2:

Table 2. Production cost and information level

	S1	S2	S3	S4	Max Capacity
A1	0.140	0.120	0.140	0.150	8
A2	0.110	0.130	0.150	0.100	7
A3	0.100	0.043	0.100	0.085	9
A4	0.130	0.130	0.130	0.125	5
A5	0.130	0.135	0.140	0.140	6
A6	0.150	0.130	0.150	0.110	8
A7	0.130	0.130	0.130	0.130	3
B1	0.015	0.010	0.010	0.010	10
B2	0.025	0.020	0.020	0.027	15
C1	0.050	0.050	0.050	0.065	12

Table 2. (continued)

C2	0.100	0.100	0.100	0.100	8
C3	0.060	0.060	0.060	0.060	4
C4	0.096	0.096	0.096	0.096	4
C5	0.020	0.020	0.020	0.020	4
C6	0.080	0.080	0.080	0.080	4
Distributor demands	20	30	25	15	
Information level	30	40	10	20	
Buffer level	18	25	10	16	

The final solutions with different mechanism are shown in Table 3, and the variations of the solution for TOC mechanism are shown in Fig3

Table 3.Optimal schedule based on TOC mechanism

	A1	A2	A3	A4	A5	A6	A7	
S1	5	7	0	0	0	0	0	
S2	0	0	0	5	0	2	3	
S3	2	0	0	0	4	0	0	
S4	0	0	6	0	0	0	0	
	B1	B1	C1	C2	C3	C4	C5	C6
S1	0	0	0	0	2	4	4	0
S2	0	7	9	0	2	0	0	0
S3	10	0	0	2	0	0	0	0
S4	0	8	3	0	0	0	0	4
Optimal fitness value	3.999070				Throughput			99.97504

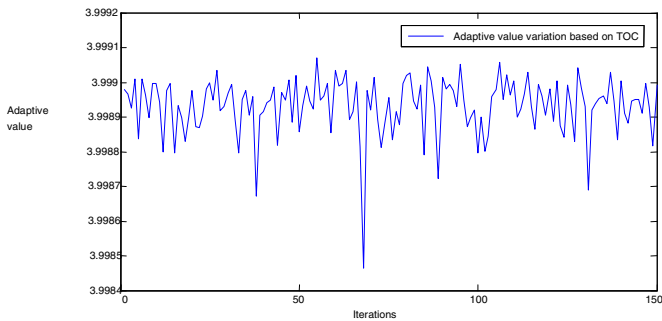


Fig. 3. Variation of the solution based on TOC mechanism

5 Conclusion

In this paper, we investigate a two echelon supply chain system consist of multi-manufacturers and multi-distributors, we proposed a optimal production allocation mechanism based on Theory of Constraints(TOC) in face of meeting peak demand in

certain period for the whole system. And through a numerical example, it has shown the efficiency of our proposed method in contrast with the traditional amount-oriented mechanism.

Acknowledgments. This paper is supported by Educational Commission of Hubei Province(serial number: Q20112201)and Research Center of Hubei Logistics development(serial number: 2011A12).

References

1. Chase, R.B.: Production and Operations Management: Manufacturing and Services. Irwin/McGrawhill (1998)
2. Moon, C., Kim, J., Hur, S.: Integrated process planning and scheduling with minimizing total tardiness in multi-plants supply chain. *Computers and Industrial Engineering* 43, 331–349
3. Simatupang, T.M., Wright, A.C., Sridharan, N.: Applying the theory of constraints to supply chain collaboration. *Supply Chain Management: An International Journal* 9(1), 57–70 (2004)
4. Smith, D.A.: Linking the supply chain using the theory of constraints logistical applications and a new understanding of the role of inventory/buffer management. In: 2001 Constraints Management Technical Conference Proceedings, San Antonio, Texas, USA, pp. 64–67 (2001)
5. Naso, D., Surico, M., Turchiano, B.: Genetic algorithms for supply-chain scheduling: A case study in the distribution of ready-mixed concrete. *European Journal of Operational Research* 177, 2069–2099 (2007)
6. Torabi, S., Ghomi, S., Karimi, B.: A hybrid genetic algorithm for the finite horizon economic lot and delivery scheduling in supply chains. *European Journal of Operational Research* 173, 173–189 (2006)
7. Lee, Y., Jeong, C., Moon, C.: Advanced planning and scheduling with outsourcing in manufacturing supply chain. *Computers and Industrial Engineering* 43, 351–374 (2002)
8. Belvedere, V., Grando, A.: Implementing a pull system in batch/mix process industry through theory of constraints: A case-study. *Human Systems Management* 24(1), 3–12 (2005)
9. Bhattacharya, A., Vasant, P.: Soft-sensing of level of satisfaction in TOC product-mix decision heuristic using robust fuzzy-LP. *European Journal of Operation Research* 177(1), 55–70 (2007)
10. Blackstone, J.H.: Theory of constraints – A status report. *International Journal of Production Research* 39(6), 1053–1080 (2001)
11. Cole, H., Jacob, D.: Introduction to TOC supply chain. AGI Institute (2002)
12. Holt, J.R.: TOC in supply chain management. In: 1999 Constraints Management Symposium Proceedings, Phoenix, AZ, USA, pp. 85–87 (1999)
13. Kendall, G.I.: Viable vision. Heliopolis Culture Group/SAGA Culture Publishing Co. (2006)
14. Luebbe, R., Finch, B.: Theory of constraints and linear programming: A comparison. *International Journal of Production Research* 30(6), 1471–1478 (1992)

15. Patnode, N.H.: Providing responsive logistics support: Applying lean thinking to logistics. In: Constraints Management Symposium, Phoenix, AZ, USA, p. 93 (1999)
16. Ahuja, R.K., Kumar, A., Krishna, J., Orlin, J.B.: Exact and heuristic methods for the weapon target assignment problem. MIT Sloan School of Management Working Paper 4464-03, 1–20 (2003)
17. Lee, Z., Lee, C., Su, S.: An immunity based ant colony optimization algorithm for solving weapon target assignment problem. *Applied Soft Computing* 2, 39–47 (2002)
18. Winston, W.L.: *Operations Research: Application and Algorithms*, 4th edn. Brooks/Cole, Belmont (2004)

Research of Data Acquisition Controller for DO Based on FPGA

WeiSheng Zhong¹ and YaPing Wang²

¹ Nan Chang University, Nan Chang330029, China

² East China Institute of Technology, Nan Chang330013, China
ws0791jx@163.com, ypwang@ecit.cn

Abstract. Real-time sampling control of A/D converter TLC5510 is realized with EP13T144C8—FPGA device of cyclone, which improved the real-time property of data acquisition for DO. Frequency division of system clock and control of A/D sampling are implemented respectively by VHDL language. High-speed storage and sending of transform data are achieved in combination design of FIFO memory. Simulation results under QuartusII development platform are presented at last. The real-time investigation and control of DO could be realized to raise the water quality of sewage treatment.

Keywords: FPGA, TLC5510, VHDL, DO Acquisition.

1 Introduction

Data acquisition is an important section in industrial control system. In sewage Treatment Control System of SBR, DO is a very important parameter, influencing the water quality to a great extent, which require the real-time of DO data acquisition consequently. Sampling process of high-speed A/D converter equipment TLC5510 was controlled by FPGA with its high frequency and real-time characters. Afterwards, the converted data was stored in FIFO high-speed memory inside the FPGA quickly, which provide the possibility of follow-up treatment for DO sampling data. The system was designed with the method of top-down in VHDL language[1].

2 Chip Description of TLC5510

TLC5510, produced by TI with the CMOS craft, is an 8-bit half-flash architecture A/D converter with 20MSPS sampling rate. It's implemented extensively in the regions of high-speed data convert, digital TV, medical graph, video conference and QAM demodulator. There is a standard voltage divider inside the TLC5510, which could acquire 2V full scale reference from 5V power source. At the same time, it could simplify the peripheral circuit design with the inner sample and hold circuit.

The main ports are as follows:

CLK: Clock input;

ANALOG IN: Analog input;

OE: Output enable; When OE=low, data is enabled. When OE=high, D1-D8 is in high-impedance state.

D1-D8: Digital data out. D1=LSB, D8=MSB;

AGND: Analog ground;

DGND: Digital ground;REFB: Reference voltage in bottom;

REFT: Reference voltage in top;

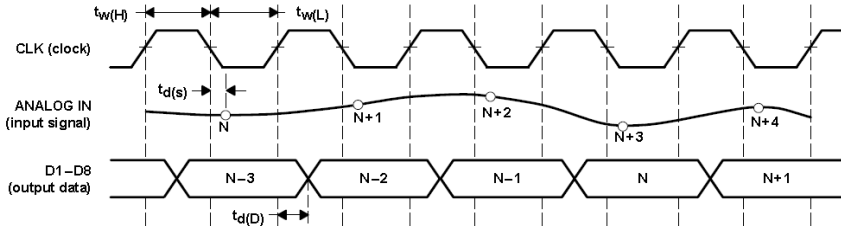


Fig. 1. Time scheduling of TLC5510

3 Design of Frequency Divider Module

The system operating clock was standard signal of 50MHz, inducted outside, but the maximum sampling frequency of TLC5510 is 20MHz. So the system clock was divided by the 2.5 half-integer frequency divider. The divider, whose frequency coefficient is 2.5, consists of XOR, modulo-three counter and a divider with the frequency coefficient 2. The VHDL description was simulated and shown as follows, in which the inclk is the system clock, outclk is the clock resulted by divider:

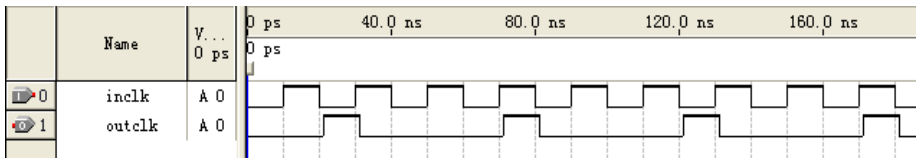


Fig. 2. Simulation diagram of the divider

4 Logic Implementation of Data Acquisition Control Based on FPGA

A state machine was designed in this paper to accomplish the logic control of data acquisition according to features of TLC5510. The main process of the design is to program VHDL code. The logic control of TLC5510 clock signal, data cache and output of converted data were realized respectively in VHDL language. Then the compilation and simulation was completed on the development platform QuartusII. The whole design was downloaded to FPGA chip at last and the simulation diagram and relevant VHDL codes are as follows[2-3]:

```

library IEEE;
use IEEE.STD_LOGIC_1164.ALL;
entity TLC5510 is
port (RST:in STD_LOGIC;
      CLK: in STD_LOGIC;
      D: in STD_LOGIC_VECTOR(7 DOWNT0 0);
      OE: out STD_LOGIC;
      Q: out STD_LOGIC_VECTOR(7 DOWNT0 0);
      LOCK0: OUT STD_LOGIC;
      ADCLK: out STD_LOGIC);
end entity TLC5510;
architecture BEHAV of TLC5510 is
  type STATES is (st0,st1,st2,st3);
  signal CURRENT_STATE: STATES;
  signal NEXT_STATE: STATES;
  signal REGL: STD_LOGIC_VECTOR(7 DOWNT0 0);
  signal LOCK: STD_LOGIC;
begin
  LOCK0<=LOCK;
  ADC: process (CURRENT_STATE)
    begin
      case CURRENT_STATE is
        when st0 =>ADCLK<='0';OE<='1';LOCK<='0';
          NEXT_STATE<=st1;
        when st1 =>ADCLK<='1';OE<='1';LOCK<='0';
          NEXT_STATE<=st2;
        when st2 =>ADCLK<='0';OE<='0';LOCK<='0';
          NEXT_STATE<=st3;
        when st3 =>ADCLK<='0';OE<='0';LOCK<='1';
          NEXT_STATE<=st0;
        when others =>NEXT_STATE<=st0;
      end case;
    end process ADC;
  REG:process(CLK,RST)
  begin
    if RST='1' then CURRENT_STATE<=st0;
    else if (CLK'event and CLK='1') then CURRENT_STATE<=NEXT_STATE;
    end if;
  end if;
  end process REG;
  LATCH:process(LOCK)
  begin
    if LOCK='1' and LOCK'event then REGL<=D;
  end if;
  end process LATCH;
  Q<=REGL;
end BEHAV;

```

In the codes above, RST, CLK are reset signal and system clock signal; D[7...0] is the converted digital signal which was linked to output of A/D converter; OE is the output enable signal of TLC5510; Q[7...0] is the output data; LOCK0 is the latch signal of output data; ADCLK was linked to the converting clock CLK of TLC5510.

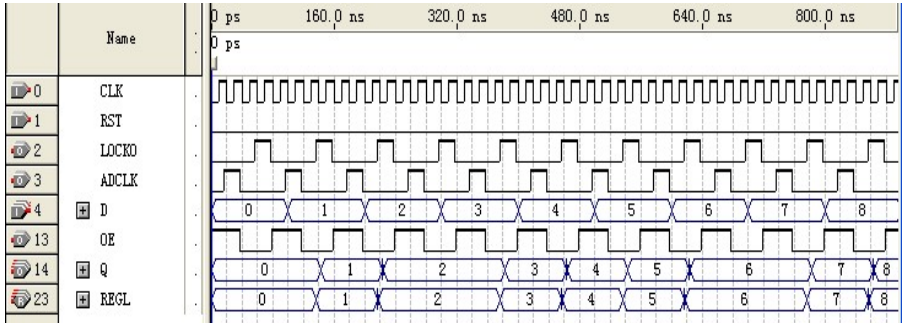


Fig. 3. Simulation diagram of data acquisition

5 Design and Simulation of FIFO Module

FIFO is a module library with settable storage parameters, which was always used for caching in high-speed digital system. FIFO is adopted as cache body in many high-speed systems nowadays. The storing velocity is highly improved because the time read-in and read-out of FIFO is only one clock cycle, in which the adding one operation of address is unnecessary[4-5].

A high-speed FIFO was customized on development platform QuartusII in this paper with the macro block LPM_FIFO in LPM. Relevant parameters were adjusted according to requirement and a FIFO module which met the own needs was generated consequently. After that, the FIFO module could be invoked in the top design of the system by using the component instantiation statements in VHDL language. Component declaration and instantiation statements are as follows:

```

component fifo2
    port(data:in STD_LOGIC_VECTOR(7 downto 0);
          wrreq,rdreq,clock,aclr:in std_logic;
          full,empty:out std_logic;
          q:out std_logic_vector(7 downto 0));
end component;
A:fifo2
port map(DATA,wr,rd,CLK,RST,FULL,EMPTY,Q);

```

Diagram 4 is the function simulated picture.

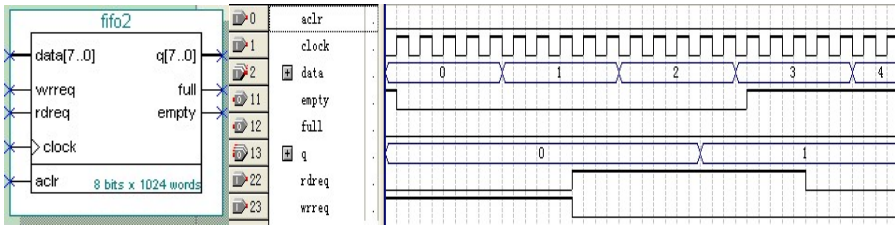


Fig. 4. FIFO module and the function simulated diagram

6 Function Simulation of the System

The system was functional simulated on QuartusII. The simulation diagram is showed in figure 5. The result of simulation indicated that control of high-speed data acquisition and storage were realized successfully in this paper by the data acquisition controller[6].

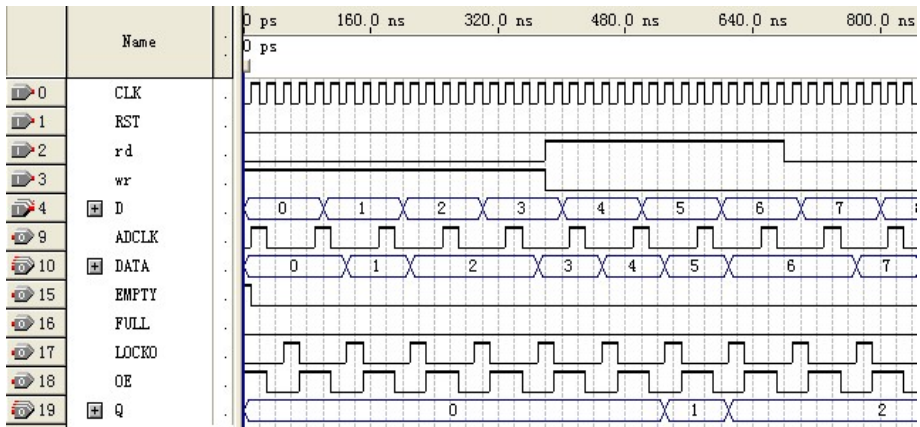


Fig. 5. Simulation diagram of the system

7 Conclusion

In consideration of the requirements on real-time sampling of DO data acquisition system in SBR sewage treatment, the real-time sampling control of high-speed A/D converter TLC5510 is realized in hardware description language VHDL on the cyclone FPGA equipment EP13T144C8. The high-speed sampling performance of TLC5510 was brought into full play and the real-time of DO data acquisition was highly improved. Combined with follow-up signal processing circuit, the real-time investigation and control of DO could be realized to raise the water quality of SBR sewage treatment.

References

1. Yang, P., Xu, Z.: Design of high-speed ADC08200 controller based on FPGA. *Application of Electronic Technique* 35(6), 65–68 (2009)
2. Li, H., Yuan, S.: Design of FPGA/CPLD based on QuartusII. Publishing House of Electronics Industry (2006)
3. Bhasker, J.: A VHDL premier. China Machine Press (2006)
4. Zhang, P., Li, Z.: Development and application of CPLD/FPGA based on VHDL. National Defence Industrial Press (2009)
5. Wang, Q., Han, J., Dong, F.: Design for high-speed data acquisition control logic based on AD7366/AD7367 and FPGA. *Electronic Instrumentation Customer* 16(1), 50–52 (2009)
6. Xiao, Z., Cheng, M.: Design of data acquisition and analysis system based on FPGA. *Application of Integrated Circuits* 35(3), 49–51 (2009)

Research on the Simplification of BOM Maintenance under the Condition of ERP

Jia Zhang and Kaichao Yu *

Faculty of Mechanical and Electronic Engineering,
Kunming University of Science and Technology, Kunming 650093, China
{Jia Zhang, Kaichao Yu, skip_zj}@163.com

Abstract. BOM is the main carrier of information transmission and the core of basic data in the ERP system and is closely related to production and operating activities of all departments in the enterprises. How to improve the quality of the data of BOM and reduce the workload of the BOM maintenance, not only in relation to the efficiency of the system and the use effect, but also directly impact on the successful implementation of ERP. This paper describes the common structures of BOM and the importance and complexity of BOM maintenance, and proposes a more practical and simplified method in BOM maintenance, and provides an application example.

Keywords: ERP, BOM, Simplification.

1 Introduction

Bill of material, as the core basic data of ERP system, is the bond of information from each functional department of the enterprise, guides and controls the production activities of the whole enterprise, and plays an important role in enterprise management and the operating efficiency of ERP system. Today, in the condition of customization-dominated and small batch production, the enterprises have been developing new product to meet soft demands of customers. Therefore, it is very important that we make the rules to modify BOM data reasonably and chose efficient methods of the transformation between different types of BOM.

2 Analyze the Common Method for the Construction of BOM

BOM is a technical document, which also known as the product structure tree or product structure table, and is used to reflect the structure relationship of the product and its parts, components, and raw materials, and their assembly relationship between each other. At the same time, BOM can also be said: the product manuals, maintenance guidance, the list of spare parts and other related documents. [1]

* Corresponding author.

In ERP system, according to the function of BOM, generally BOM is classified: Engineering BOM (EBOM), Process Planning BOM (PPBOM), Manufacturing BOM (MBOM), Buying BOM (BBOM), Sales BOM (SBOM) and so on.

2.1 The Methods for the Construction of BOM

a. Analyze the structure of single-level BOM

The single-level BOM adopts a structure which contains a father part and a direct children part, which reflects corresponding relationship of children part and father part.^[2] For a children part, its father part likely is a product, can also be a component. There have no duplicate data records and smaller data redundancy because that each hierarchy is defined only once in the single-level BOM, which could greatly reduce data storage space.

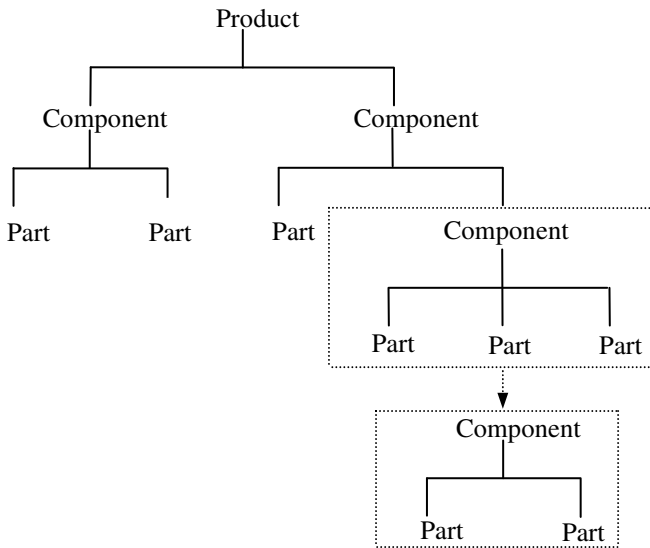


Fig. 1. Structure tree of product A₁

The single-level BOM can reflect structure tree of product clearly. If the regression method is adopted, we could begin with any product or component to find all children parts which are used to make the product or component, similarly, we can trace its father part by a children part quickly. In the single-level BOM, if the structure of component G is changed (as shown in figure 1), all product BOM using the component G would be changed, which reduces the BOM maintenance workload for employee. The structure of component G is changed because of human error, if not promptly corrected; it would cause inestimable loss, which put forward higher requirements to the maintenance of BOM basic data.

b. Analyze the structure of multi-level BOM

The multi-level BOM adopts a structure which contains a father and children parts, which detailed reflect the structure relationship of product and its components and parts, as well as detailed reflect the structure relationship of component and its parts. To a component or part, if it exists in the different components, which needs exclusive record correspond to? [3] The feature of structure of multi-level BOM is that the structures of product have not influence each other and the data records of product have no cross. Therefore, it has higher decomposition rate, and the data is easier to maintain. [2] However, in multi-level BOM, the structure of component or part is defined repeatedly, which leads to make larger data redundancy and increase to workload of basic data entry in the early days of system operation.

3 Simplify the BOM Maintenance

3.1 The Complexity of BOM Management

In mechanical manufacture field, there are too many BOM types, and each one contains the number of materials is truly staggering. In the fierce market competition, changeful customer requirements and small batch production make the number of BOM increase sharply. If the enterprise do not use the proper and efficient methods of BOM maintenance, and which would lead to low efficiency inevitably and even make the serious influence on the process of making products. In ERP system, to fulfill the traceability of products, there are two exclusive BOM correspond to the two products, even if the two products have difference in the specification of a part and component, which makes BOM data and its management become more complicated. [1]

As we all know, for the enterprise in mechanical manufacture field, the number of data of one product BOM is truly staggering, let alone a series of products. Manual entry these data, this means increasing the workload, making mistakes easily, which would go against the operation of ERP system in later stage. Therefore, for the BOM data, it can be imported into system by computer program in the early days of ERP system operation. If there have bad data, it would be modified and debugged in the process of ERP system operation. Take the system data security into account, after the system start to run formally, a little new data should be inputted manually, in order to avoid data importing program failures or human error lead to the changes and loss of historical data.

In ERP system, if a certain type of product has no orders at the present stage, in order to achieve product traceability, facilitate maintenance and repair; the BOM of this product would be stored in the system or external storage as historical data. It wouldn't be deleted but to deal with it by making it become invalid. [1] When enterprise needs to produce the same or similar types of products, just re-enable the product BOM files.

3.2 The Method of Simplifying BOM Maintenance

In ERP system, the introduction of phantom can simplify BOM maintenance, better display processing sequence, facilitate the maintenance of process route, reduce data redundancy in the system and improve the operation of the ERP system efficiency.

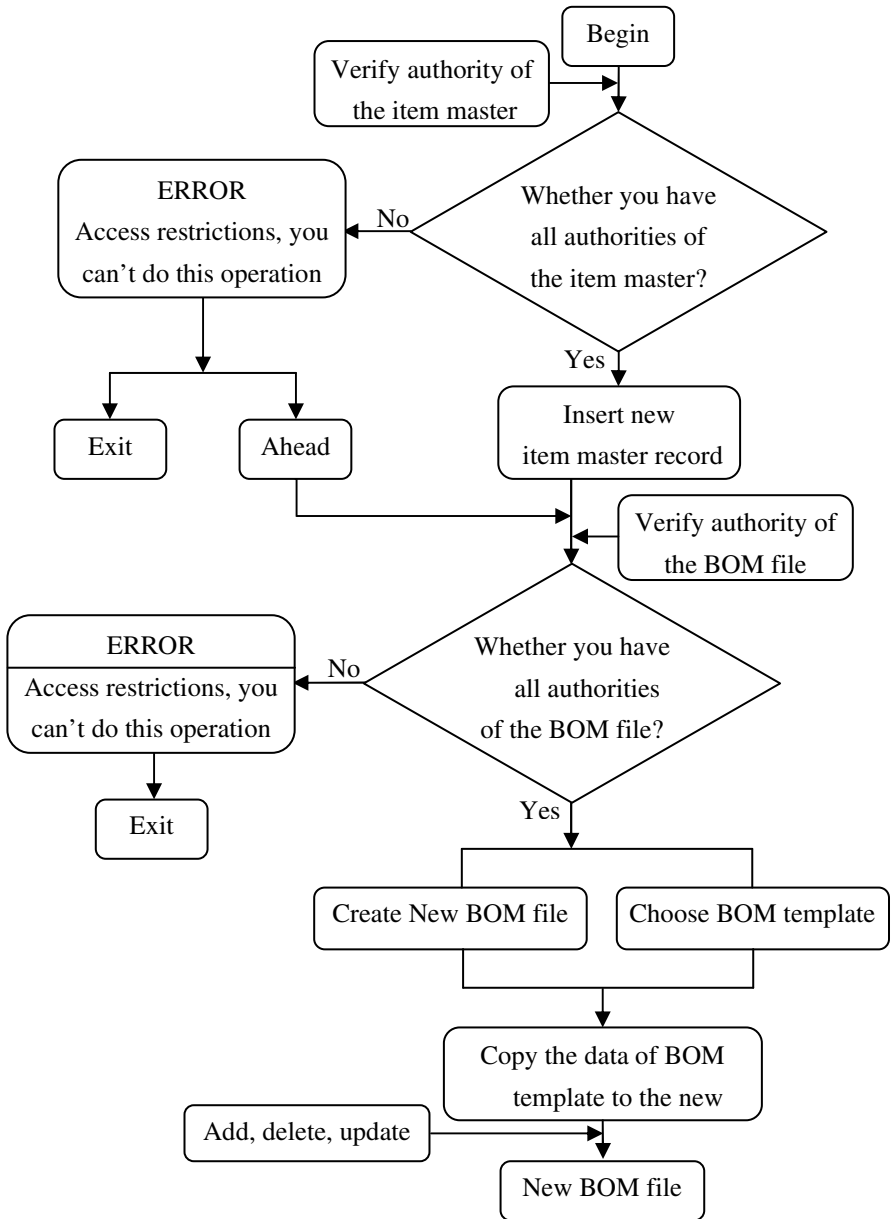


Fig. 3. BOM transformation Schematic

As a transition part in product structure, phantom shows a kind of item which does not exist in and does not appear in drawings and machining process. It can be used to represent a group, a series of, a kind of attribute or a kind of specification materials. Phantom belongs to virtual item and does not exist in inventory. [4] According to the product process

requirements, there must be some temporary components in BOM. For these components with good generality, the enterprise should establish BOM files for it. Especially many BOM files have a lot of the same components, the advantage of which will be more obvious. If the enterprise production tasks are not very urgent, phantom could be produced in advance to prepare for the production of the final product.

The data has some similarities between different type of BOM and the same type of BOM, and the new BOM can extract materials from these BOM. In accordance with order requirements, adding, modifying, deleting materials would generate a new BOM file. [1] As shown in figure 2.

3.3 Distribute BOM Permissions

BOM as an important basic data in production and operating activities of the enterprise, and its accuracy, consistency, Security relate to the survival of the whole enterprise. Therefore, for the BOM data maintenance work, enterprise must choose professional staff and strictly limit the permissions of BOM maintenance staff to ensure the validity and security of BOM data.

To distribute BOM permissions, system administrator can distribute permissions to operators by whom in charge of business contents and set authorization groups by the type of BOM to strictly control the change of the BOM data. According to the requirements of the Company to manufacture products, BOM data will be modified after the competent departments of the layer upon layer for examination and approval. And the BOM data before modifying must be saved to facilitate future verification.

4 Conclusions

ERP is the important mean to the informatization of enterprises. Simplifying the maintenance of BOM which is the core of ERP will promote better implementation of the ERP Project. The methods of simplifying BOM proposed in this paper not only will help the enterprise reduce the BOM maintenance workload, ensure the consistency and reliability of BOM data and improve the operation of the system efficiency, but also will provide reference for the development of BOM module in ERP software.

References

1. Zhang, J., Yu, K.: BOM in the ERP Operation and its Maintenance. *Chinese Manufacturing Information Engineering of China* 5, 23–25 (2011)
2. Hu, S., Tao, S., Yang, M.: The Expression of BOM in ERP System. *Chinese Manufacturing Information Engineering of China* 34, 102–104 (2005)
3. Liu, Y., Yu, M., Zhang, B., Wang, C., Chen, R.X., Su, Y.: Research on Structure of BOM in Enterprise Resource Planning Systems. *Computer Integrated Manufacturing Systems* 4, 309–313 (2003)
4. Wang, X., Yang, Y., Li, C.: *ERP Enterprise Management and Case*. Tsing Hua University Press, Beijing (2007)

Simulation and Analysis of the Common-Emitter Circuit in Analog Electronics Teaching

Xiao-Hong Wu^{1,*}, Bin Wu², Jun Sun¹, and Zhen-Yu Wang¹

¹ School of Electrical and Information Engineering, Jiangsu University,
Zhenjiang 212013, P.R. China
wxh419@ujs.edu.cn

² Department of Information Engineering, ChuZhou Vocational Technology College,
ChuZhou 239000, P.R. China

Abstract. Common-emitter (CE) circuits are the basic and important amplifiers in analog electronics. Theoretical analysis of the CE circuit is carried out to calculate the Q-point, the voltage gain, the input impedance and the output impedance. The CE circuit is simulated to make the analysis of DC operating point analysis, DC sweep analysis, AC analysis and amplitude of output voltage. The simulation results are consistent with the theoretical calculation. It is concluded that a virtual simulation on electronic circuits with Multisim 10 can be easily built and test without the restriction of hardware devices. Multisim 10 provides students with a good experimental platform, and makes them convenient to do experiments, demonstrate and analyze circuits in the process of analog electronics teaching.

Keywords: Analog electronics, Common-emitter circuits, Quiescent point, Electronics design automation, AC analysis.

1 Introduction

Electronics design automation (EDA) is a powerful technology in the field of modern Electronics technology. It provides a bran-new design of electrocircuit system for electronic designer. It not only has strong design capabilities, but also has the testing, analysis and management capabilities. According to the EDA service objects, EDA software is classified into four kinds: the circuit design and analysis, such as Pspice and Multisim 10; the digital circuit design, such as Quartus II; the radio frequency circuit design, such as advanced design system (ADS); the printed circuit board (PCB) design, such as Protel. Multisim 10 produced by American National Instruments is a simulation software for electronic circuits [1]. It contains many kinds of components, which can be chosen to use in experiments, in the component database. At the same time, the new components can be designed to expand the component database. It provides all kinds of virtual instruments that include universal instruments (multimeter, function generator, wattmeter and oscilloscope) and special instruments (bode plotter, word generator, logic analyzer, distortion analyzer and

* Corresponding author.

spectrum analyzer etc) [2,3]. The powerful function of circuit analysis, such as DC operating point analysis, AC analysis, transient analysis, Fourier analysis and noise analysis etc, can help designer to analyze the performance of circuits. Multisim 10 is used to design, test and demonstrate various electronic circuits including analog circuits, digital circuits, interface circuits and RF circuits etc. The virtual simulation laboratory can be established with Multisim 10 on the computers [4]. In this laboratory, the designer can design circuits and conduct the experiments at the same time; the parameters of circuits can be easily tested and analyzed; the designer can directly print the experimental data, test parameters, curve and circuit figures.

Common-emitter (CE) circuits play an important role in analog electronics [5,6]. Quiescent point (abbreviated Q-point) will affect the output waveform and the voltage gain of CE circuits. We calculate the Q-point theoretically and test the Q-point with Multisim 10. The value of bias resistor is adjusted to change the position of Q-point. If the Q-point is an improper operating point, bipolar junction transistor (BJT) will operate in the cutoff or saturation region. On the other hand, if the Q-point is a proper operating point, BJT will operate in linear region and can amplify the signal without distortion. The temperature also influences the Q-point [7]. With temperature sweep analysis of Multisim 10, the temperature in the CE circuit is changed to see the influence of temperature on the Q-point. So Multisim 10 provides students with an excellent experimental platform for simulating the performances of various electronic circuits. It makes students master the content of analog electronics easily and it inspires the innovative abilities of students.

2 Theoretical Analysis of CE Circuit

The voltage-divider bias amplifier is shown in Fig. 1. It is a common-emitter circuit used to stabilize the Q-point. The bias circuit is a voltage-divided circuit composed of Rb1 and Rb2. At the emitter of transistor Q1, Re is connected to stabilize the Q-point. At the frequency of operation, the reactance of the capacitor C3 is so small compared to Re that it is treated as a short circuit across Re. When the input signal u_i is put at the input side, the output side will obtain a phase-reversed and amplitude-amplified output signal u_o . That is to say, the input voltage is amplified.

When the values of the DC currents across the Rb1 and Rb2 far exceeds that of base current I_B of transistor. The Q-point is calculated as follows:

$$U_B \approx \frac{Rb2}{Rb2+Rb1} V_{cc} = \frac{10k}{10k+30k} 12 = 3V \quad (1)$$

In above equation, U_B is the base DC potential of transistor, and the rheostat Rb2 is set to 10k .

$$I_E = \frac{U_B - U_{BE}}{R_e} = \frac{3-0.7}{1k} = 2.3mA \quad (2)$$

$$U_{CE} \approx V_{cc} - I_C(R_c + R_e) = 5.1V \quad (3)$$

The voltage gain is

$$A_u = -\frac{\beta R_c}{r_{be}} = -139 \tag{4}$$

In equation (4), the current amplification coefficient β of transistor Q1 is 100 and the base-emitter ac resistance $r_{be} = 1.44k \Omega$.

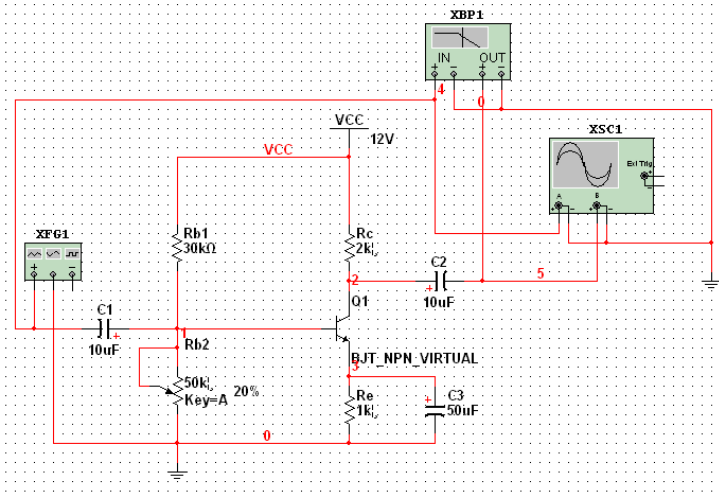


Fig. 1. Voltage-divider bias amplifier

3 Simulation of the Common-Emitter Circuit with Multisim 10

The parameters of Multisim 10 should be set properly before we run it to simulate circuits. In the common-emitter circuit of Fig. 1, the value of the rheostat Rb2 influences the Q-point, and it can be adjusted to an appropriate value to obtain a suitable Q-point. To compare the simulation results with the theoretical calculation, we also set the value of the rheostat Rb2 to 10k . The input signal can be obtained from function generator that can generate three kinds of waves, i.e. sine wave, triangle wave and square wave. Double click on the icon of XFG1 in Fig. 1, and the panel figure will appear with the parameters of function generator. The panel includes two areas: one is “waveforms” and the other is “signal options”. There are three waveforms to choose in waveforms area, and four parameters, such as frequency, duty cycle, amplitude and offset, to be set in signal options area.

3.1 DC Operating Point

When the distortion of the output waveform does not happen, we select Options → Sheet Properties → Show All, and the number of all nodes will show in

Fig. 1. Now, we perform DC analysis simulation with multisim 10 to observe the node voltages with the selection of Simulate → Analyses → DC Operating Point Analysis → Output. Then the variables for analysis are selected in the output page, and after the mouse clicks on button “Simulate” the DC analysis result can be shown in Fig. 2 automatically. $V[1] = 2.84741V$, $V[1]$ is U_B in equation (1), so $V[1]$ equals to U_B approximately. $V[1] - V[3]$ is U_{BE} in equation (2), and $V[1] - V[3] = 0.79258$, so $V[1] - V[3]$ equals to U_{BE} approximately. $V[2] - V[3]$ is U_{CE} in equation (4), and $V[2] - V[3] = 5.87619V$, $U_{CE} = 5.1V$, the error between theoretical value and simulation value is $5.87619V - 5.1V = 0.77619V$ which is small.

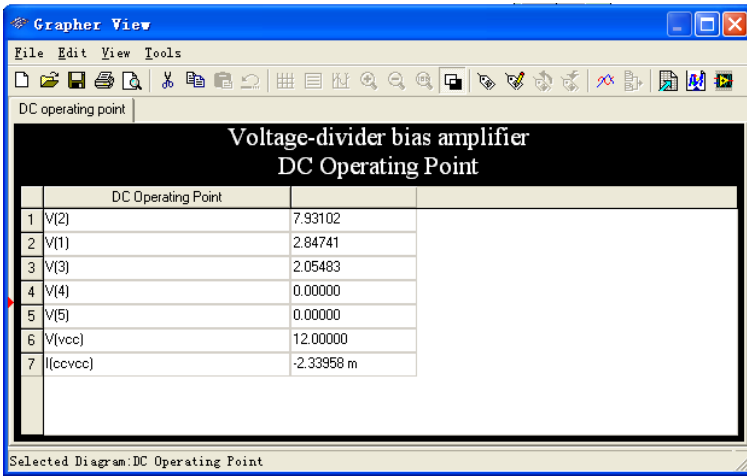


Fig. 2. DC operating point

3.2 AC Analysis

AC analysis of Multisim 10 deals with the frequency response of amplifier. The basic step is Simulate → Analyses → AC analysis, and the dialog box of AC analysis pops up. After we set the parameters well, we press the button “simulate” and the simulation result will be shown in Fig. 3. The result of AC analysis includes magnitude-frequency characteristic and phase- frequency characteristic.

3.3 Amplitude of Output Voltage

We set the amplitude of input sine wave as 10 mV in function generator. The CE circuit of Fig. 1 is simulated, and the amplitude of output wave can be obtained by moving the pointer. In this simulation, the amplitude of output voltage is 1.365V in Fig. 4.

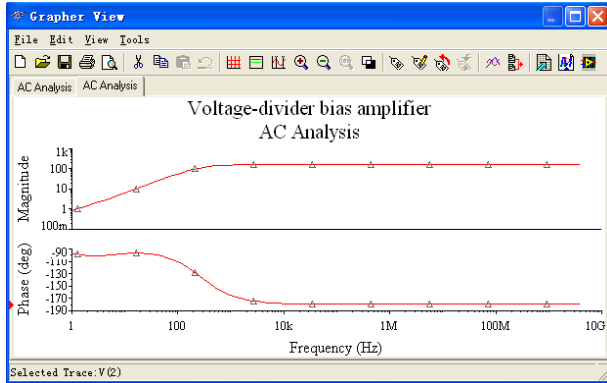


Fig. 3. AC analysis

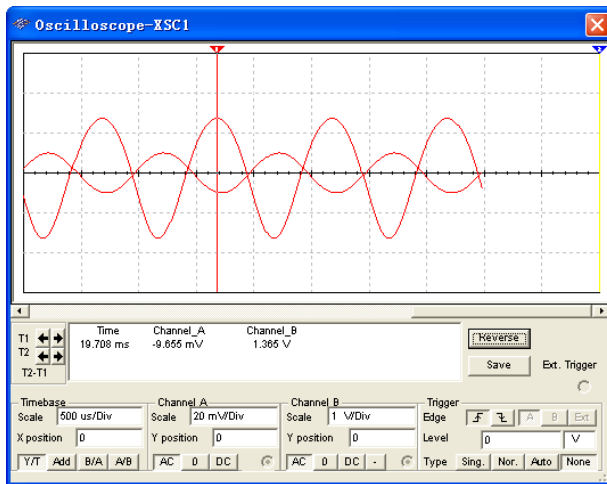


Fig. 4. The amplitude of output voltage

4 Conclusions

In addition to the above simulations on CE circuit, there are still other simulations including noise analysis [8,9], distortion analysis and parameter sweep analysis etc. In this paper, we calculate and analysis the Q-point, the voltage gain, the input impedance and the output impedance theoretically. Then we simulate the CE circuit to carry on the analysis of static state and dynamic state. The simulation results are consistent with the theoretical calculation. We can conclude that a virtual simulation on electronic circuits with Multisim 10 can be easily built and test without the restriction of hardware devices. Multisim 10 provides students with a good experimental platform, and makes them convenient to do experiments, demonstrate

and analyzing circuits in the process of teaching. Using Multisim software to simulate circuits can stimulate the students' imagination and creativity, and cultivate their creative thinking and creative ability [10].

Acknowledgments. The authors would like to thank Jiangsu higher education association "twelfth five-year plan" higher education science project (No. KT2011526) and China education society of electronics "twelfth five-year plan" higher education science project for supporting this research.

References

1. Liao, L.W., Liu, J., Wang, S.: Research of Multisim in the Experiment Teaching. In: International Conference on Computer Science and Software Engineering, CSSE 2008, pp. 515–517 (2008)
2. Huang, Z.: Electronic Circuits simulation and analysis by Computer Based on NI Multisim. Publishing House of Electronics Industry, Beijing (2008) (in Chinese)
3. Lao, W., Lao, J.: Analysis, Design and Simulation of Electronic Circuits. Press of Tsinghua University, Beijing (2007) (in Chinese)
4. Bian, H., Dai, K., Jiang, L.: Application Multisim to Virtual Laboratory for Experiment Teaching. In: International Conference on Computational Intelligence and Software Engineering (CISE), pp. 1–4 (2010)
5. Ji, R.D.: The Simulation and Analysis of BJT Common-Emitter circuit Based on Multisim. Science & Technology Information 30, 211 (2009) (in Chinese)
6. Wang, X.L., Wang, W.L., Zhang, X.D., et al.: Study of Negative Feedback Amplifier Circuit Based on Multisim. In: The 2nd International Conference on Computer Engineering and Technology (ICCET), vol. V6, pp. 85–87 (2010)
7. Boylestad, R.L., Nashelsky, L.: Electronic Devices and Circuit Theory, 9th edn. Pearson Education, Inc., Prentice Hall (2006)
8. Gray, P.R., Meyer, R.G.: Analysis and Design of Analog Integrated circuit, 3rd edn. John Wiley & Sons, New York (1993)
9. Wang, X.L., Zhang, J.Y., Wang, W.L., et al.: Noise Analysis for Electronic Circuit Using Multisim. In: The 2nd IEEE International Conference on Information Management and Engineering (ICIME), pp. 324–326 (2010)
10. Liu, G.D., Wang, S.J.: Study-Based Teaching Practice for Electronics Technology Foundation using Multisim. Journal of EEE 32(5), 60–62 (2010) (in Chinese)

Research on the Evaluation of the Third Party Reverse Logistics Suppliers

QingKui Cao* and JingHua Zhang

College of Economics and Management, Hebei University of Engineering,
056038, Handan, China

Abstract. Reverse logistics has become a developing tendency, so the selection of the reverse logistics supplier becomes a real problem for the enterprise. From the view of customers this paper selects six aspects to built up the third party reverse logistics supplier evaluation index system. Membership conversion algorithm based on M (1, 2, 3) is introduced, and redundancy of target classification from index membership is cleared by the information entropy index distinguish right, and then the RMS values is extracted so as to realize the transformation of the membership degree step by step. Finally, the evaluation of the comprehensive strength overall goal of the third party reverse logistics provider is realized. An empirical study is used to prove that the whole evaluation index system and the algorithm are operable and practical.

Keywords: The third party reverse logistics, Index system, Membership conversion algorithm, Evaluation for the reverse logistics supplier.

1 Introduction

Today, with competition among enterprises getting more and more intense, many enterprises have paid enough attention to reverse logistics. Based on American Reverse Logistics Executive Committee's report, reverse logistics management directly affects customer satisfaction and enterprise profits. So, reverse logistics has become a powerful weapon in enterprises competition. Considering the complexity and uncertainty of reverse logistics, many enterprises are now more likely to outsource the reverse logistics to third-party reverse logistics providers. Thus, small-and-medium-sized enterprises of weak economic basis, they can achieve cost savings through outsourcing and improved operational efficiency. Meanwhile, large-scale enterprises can concentrate on the research and development of core technology and product and improve the enterprise's core competitiveness. According to a survey conducted by The University of Tennessee, Exel and EMST&YOUNG, many business owners said that their logistics cost was reduced by 1.18 percent, goods turnover has been reduced from 7.1 d to 3.9 d average level, stocks have dropped 8.2 percent. Reverse logistics outsourcing will become a development trend. Consequently outsourcing providers' selection has become a practical problem need to be solved. Although domestic scholars made a number of

* School of Hebei University of Engineering, engaged in the research of logistics engineering and supply chain management.

studies on third-party reverse logistics service providers selection, the literature in this research are few. Adding to most scholars applied the traditional AHP method and fuzzy evaluation method to evaluate, but traditional fuzzy comprehensive evaluation method generally get evaluation set by linear weighted average. Generally processing is evaluated using the weighted average set of linear. It is clear that we haven't found a more precise evaluation method, so there is a great potential for further research in this area. therefore , this paper intended to design a set of evaluation index system from a comprehensive perspective ,and to evaluate the third party reverse logistics supplier based on M (1, 2, 3) membership conversion algorithm from the view of customers so as to provide a decision-making tool for customers to select service outsourcing enterprises.

2 Index System Construction of the Third Party Reverse Logistics Supplier

The reverse logistics is a complicated process, including the return items, materials alternative, items reuse, waste processing, maintenance and remanufacturing and so on. So, holding on these principles such as systematic, feasibility, flexibility, combining qualitative and quantitative, this paper selects six aspects such as market competitiveness, service ability, coordination ability, information level, safety and compatibility to built up the third party reverse logistics supplier evaluation index system as shown in table 1.

Table 1. Evaluation index system of the third party reverse logistics service provider

Overall target	First-class index <i>B</i>	Second-class index <i>C</i>
the comprehensive strength of the third Party logistics supplier <i>A</i>	market competitiveness <i>B</i> ₁ (0.20)	Market share rate <i>C</i> ₁₁ (0.20)
		Capital operation capacity <i>C</i> ₁₂ (0.30)
		Technical innovation ability <i>C</i> ₁₃ (0.25)
		Customer satisfaction <i>C</i> ₁₄ (0.15)
		Industry experience <i>C</i> ₁₅ (0.10)
		Service coverage area <i>C</i> ₂₁ (0.20)
	Service ability <i>B</i> ₂ (0.25)	Transportation capacity <i>C</i> ₂₂ (0.10)
		Service price <i>C</i> ₂₃ (0.10)
		Advancement of business processing equipment <i>C</i> ₂₄ (0.20)
	Coordination ability <i>B</i> ₃ (0.15)	Response to service requirements <i>C</i> ₂₅ (0.20)
		Accuracy of processing time <i>C</i> ₂₆ (0.20)
		Communication level with client <i>C</i> ₃₁ (0.30)
		Coordination ability for resources <i>C</i> ₃₂ (0.40)
		The level of business management <i>C</i> ₃₃ (0.30)

Table 1. (continued)

Information level B_4 (0.20)	Degree of the information sharing C_{41} (0.25)
	Functions and standardization of information system C_{42} (0.45)
Safety B_5 (0.10)	Degree of information monitoring C_{43} (0.30)
	Financial stability C_{51} (0.20)
	Safety of information data C_{52} (0.35)
	Safety of plant C_{53} (0.25)
Compatibility B_6 (0.10)	Safety of customer cargo C_{54} (0.20)
	Compatibility of customer culture C_{61} (0.20)
	Compatibility of customer strategy C_{62} (0.20)
	Sharing of benefits and risks C_{63} (0.30)
	Compatibility of production technology C_{64} (0.30)

3 Conversion Algorithms Based on M (1, 2, 3) Membership Grade

3.1 The Algorithm Processor

There are n kind of indexes with influence target state A in a known multi-index system, and each index is divided into t grades which can be described by C_1, C_2, \dots, C_t , and C_k grade is better than C_{k+1} grade.

The divisional right $\alpha_i(A)$ expresses the contribution size of index i affecting target A classification. If the membership grade $\mu_{ik}(A)$ belonging to C_k grade of index i and the importance weight $w_i(A)$ of i affecting target A are given out, then the quantitative description $\alpha_i(A)$ of the data mining technology based on entropy is:

$$\alpha_i(A) = V_i(A) / \sum_{p=1}^n V_p(A), (i = 1, 2, \dots, n) \tag{1}$$

If the membership grade $\mu_{ik}(A)$ of k kind of index i and he divisional right $\alpha_i(A)$ of index i about target A are given out, and then k kinds of effective value of index i are:

$$\alpha_i(A) \times \mu_{ik}(A) \quad k = 1, \dots, t. \tag{2}$$

Eliminating the "unit importance" with some degree of difference between different indexes needs to convert effective value to Valid values. If a k kind effective value $\alpha_i(A) \times \mu_{ik}(A)$ of index i and the importance weight $w_i(A)$ of index i about target A are given out, the k kind valid values of index i is:

$$w_i(A) \times \alpha_i(A) \times \mu_{ik}(A) \quad k = 1, \dots, t. \tag{3}$$

And then get the comparable summation:

$$M_k(A) = \sum_{i=1}^n w_i(A) \times \alpha_i(A) \times \mu_{ik}(A) \quad k = 1, \dots, t. \tag{4}$$

If a k kind comparable summation $M_k(A)$ of target A is given out, and then the C_k kind membership grade of target A is:

$$\mu_k(A) = M_k(A) / \sum_{p=1}^t M_p(A) \quad k = 1, \dots, t. \tag{5}$$

So far, the conversion from all the index membership to the target membership can be realized through the above process, and this process is called M (1, 2, 3) ---- "an effective, two comparable, three synthesis "mode.

3.2 Recognition

As the evaluation rating is orderly and C_k grade is better than C_{k+1} grade, the method of confidence declaration should be used. Let λ be a confidence, and the criterion for it is:

$$C_0 = \min \left\{ c \left| \sum_{t=1}^c \mu_t(A) \geq \lambda, 1 \leq k \leq 4 \right. \right\}. \tag{6}$$

It shows that target A belongs to the C_0 grade, and its confidence is no less than λ .

4 Application Examples

Taking a manufacturing company looking for reverse logistics partners as an example. The evaluation index system is shown with reference to the table 1, and the importance weights of all indexes are got based on expert evaluation and combined with AHP analytic method. The data in the round brackets shown by the table 1 is the corresponding weights in each index. Through rallying expert experience, the evaluation set is divided into four ranks, such as very satisfied (D_1), satisfied (D_2), qualified (D_3) and unsatisfactory (D_4), and the membership vector is got by the expert scoring, data have shown in table 2.

Table 2. The membership vector $(D_1 D_2 D_3 D_4)$

$C_{11}(0.25 \ 0.40 \ 0.30 \ 0.05)$	$C_{25}(0.10 \ 0.30 \ 0.30 \ 0.30)$	$C_{52}(0.20 \ 0.40 \ 0.40 \ 0.00)$
$C_{12}(0.30 \ 0.35 \ 0.25 \ 0.10)$	$C_{26}(0.15 \ 0.40 \ 0.25 \ 0.20)$	$C_{53}(0.30 \ 0.30 \ 0.30 \ 0.10)$
$C_{13}(0.20 \ 0.30 \ 0.40 \ 0.10)$	$C_{31}(0.10 \ 0.30 \ 0.30 \ 0.30)$	$C_{54}(0.20 \ 0.40 \ 0.30 \ 0.10)$
$C_{14}(0.30 \ 0.30 \ 0.25 \ 0.15)$	$C_{32}(0.20 \ 0.30 \ 0.40 \ 0.10)$	$C_{61}(0.40 \ 0.40 \ 0.10 \ 0.10)$
$C_{15}(0.30 \ 0.50 \ 0.20 \ 0.00)$	$C_{33}(0.20 \ 0.30 \ 0.30 \ 0.20)$	$C_{62}(0.30 \ 0.30 \ 0.20 \ 0.20)$
$C_{21}(0.25 \ 0.30 \ 0.25 \ 0.20)$	$C_{41}(0.25 \ 0.30 \ 0.35 \ 0.10)$	$C_{63}(0.40 \ 0.30 \ 0.20 \ 0.10)$
$C_{22}(0.40 \ 0.30 \ 0.15 \ 0.15)$	$C_{42}(0.40 \ 0.30 \ 0.20 \ 0.10)$	$C_{64}(0.35 \ 0.40 \ 0.15 \ 0.10)$
$C_{23}(0.30 \ 0.50 \ 0.20 \ 0.00)$	$C_{43}(0.30 \ 0.35 \ 0.30 \ 0.05)$	
$C_{24}(0.30 \ 0.50 \ 0.10 \ 0.10)$	$C_{51}(0.30 \ 0.30 \ 0.25 \ 0.15)$	

(1) Now taking the service ability as an example, and using conversion algorithms based on M (1, 2, 3) membership grade to calculate as follows: :

By the formula (1) to (3) the valid values matrix of B_2 are got :

$$N(B_2) = \begin{bmatrix} 0.000 \ 6 & 0.000 \ 7 & 0.000 \ 6 & 0.000 \ 5 \\ 0.004 \ 4 & 0.003 \ 3 & 0.001 \ 6 & 0.001 \ 6 \\ 0.013 \ 2 & 0.021 \ 9 & 0.008 \ 8 & 0.000 \ 0 \\ 0.016 \ 1 & 0.026 \ 8 & 0.005 \ 4 & 0.005 \ 4 \\ 0.001 \ 8 & 0.005 \ 3 & 0.005 \ 3 & 0.005 \ 3 \\ 0.002 \ 5 & 0.006 \ 6 & 0.004 \ 1 & 0.003 \ 3 \end{bmatrix}$$

Then the k kind comparable summation matrix of B_2 is:

$$M(B_2) = (0.038 \ 5 \ 0.064 \ 7 \ 0.025 \ 8 \ 0.016 \ 1)$$

The membership grade matrix of index B_2 is got by $M(B_2)$:

$$\mu(B_2) = (0.265 \ 2 \ 0.445 \ 6 \ 0.178 \ 1 \ 0.111 \ 1)$$

(2) Similarly the membership vector of $\mu(B_1)$, $\mu(B_3)$, $\mu(B_4)$, $\mu(B_5)$, $\mu(B_6)$ is got respectively, the evaluation matrix $U(A)$ of comprehensive strength of the third party reverse logistics enterprise for the customer enterprise in this case is formulated together with $\mu(B_2)$:

$$U(A) = \begin{bmatrix} \mu(B_1) \\ \mu(B_2) \\ \mu(B_3) \\ \mu(B_4) \\ \mu(B_5) \\ \mu(B_6) \end{bmatrix} = \begin{bmatrix} 0.265 \ 4 & 0.393 \ 6 & 0.280 \ 9 & 0.060 \ 1 \\ 0.265 \ 2 & 0.445 \ 6 & 0.178 \ 1 & 0.111 \ 1 \\ 0.169 \ 1 & 0.300 \ 0 & 0.360 \ 5 & 0.170 \ 4 \\ 0.333 \ 7 & 0.319 \ 6 & 0.266 \ 3 & 0.080 \ 4 \\ 0.215 \ 2 & 0.384 \ 3 & 0.369 \ 6 & 0.030 \ 4 \\ 0.378 \ 7 & 0.368 \ 9 & 0.148 \ 9 & 0.103 \ 5 \end{bmatrix}$$

The membership vector of comprehensive strength of the third party reverses logistics enterprise is got by the membership transformation combined with importance weights of first-class indexes in the table 1:

$$\mu(A) = (\mu_1(A) \dots \mu_4(A)) = (0.273 \ 6 \ 0.383 \ 9 \ 0.259 \ 3 \ 0.083 \ 3)$$

(3) Recognition

Taking $\lambda > 0.65$ here, as $0.273 \ 6 + 0.383 \ 9 = 0.657 \ 5 > 0.65$, the evaluation of the comprehensive strength of the third party reverse logistics enterprise can be considered as satisfaction level with confidence to 65.75% by the customers. Customers can choose to cooperate with this service enterprise.

5 Conclusion

Conversion algorithms based on M (1, 2, 3) membership grade can eliminate the interference of redundantly numerical value of indexes and make the evaluation result more accurate so that more close to actual level. The proposed method is applied in the evaluation of comprehensive ability of the third party reverse logistics service provider, which can comprehensively reflect the service ability and further penetrate the strengths and weaknesses of each section of their sub goal for the reference of policy-makers.

Acknowledgement. Fund project: Chinese National Natural Science Foundation (60075013), Hebei province National Natural Science Foundation (F2005000482).

References

1. Chen, Y., Tang, C.: The analysis of the effective management strategy of negative logistics. *Soft Science* 18(1), 30 (2004)
2. He, L., Zhou, G.: Analysis on 3PL reverse logistics. *Logistics Technology* (7), 16–17 (2005)
3. Stock, J.R.: *Reverse logistics*, vol. 31. Council of logistics Management, Oak Brook, IL (1992)
4. Cao, Q., Ruan, J., Liu, K.: Fuzzy evaluation on mining investment decision based on membership degree transformation algorithm. *Journal of Hebei University of Engineering (Natural Science Edition)* 27(1), 93–94 (2010)

Simulation Analysis of Traffic Capacity for Ticket Gates of Metro Station

Xin Han, Xie He, and Beihua Cong

Shanghai Institute of Disaster Prevention and Relief, Tongji University,
Shanghai 200092, China
hanxin@tongji.edu.cn, tjlet@163.com, bhcong@tongji.edu.cn

Abstract. The traffic capacity of metro station is very important for safe and efficient passenger evacuation. The proper design of ticket gates contribute to rapid evacuation of metro station. Through field test and simulation analysis, this paper gives out the appropriate obstacle value of ticket gates for software of Building Exodus v4.0 and analyzes the traffic capacity of the ticket gates. The corresponding results are helpful to evaluate the safe margin of crowd evacuation for metro station.

Keywords: ticket gates, side gates, evacuation time, metro station.

1 Introduction

Metro stations are general public places with high density population. The safe evacuation of passengers is the main target of security issue for operations management of metro stations. There are several primal problems to be considered during the design and operation phase of the metro station. The emphasis should be focused on reasonable evacuation routes, accurate estimation of evacuation time, comprehensive evacuation strategy of passengers [1], etc. Taking from the previous research result and practical engineering experience [2], this paper carries out simulation analysis of traffic capacity for metro station, including the effect of the ticket gates, the side gates and the emergency evacuation exits in the emergency operation process.

2 Evacuation Model of Metro Station

For the study of the traffic capacity for large scale metro station including complex functions, the evacuation model of this paper selects the hall layer of a metro station under construction.

The construction area of the evacuation model is about 4375m². It has 4 common exits with width of 6m and 1 emergency exit with width of 3.5m. The 4 common exits are at the corner of the hall layer and the emergency exit is between the middle of the exit 1 and exit 2. The hall layer has 8 gates of ticket machine. The width of every gate routes is 0.5m. The hall layer has 8 ticket gates and side gates separately, they locate at the hall layer symmetrically, as shown in Fig.1 and Fig. 2.

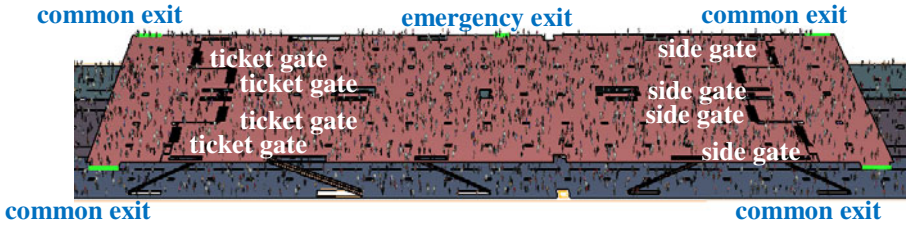


Fig. 1. Evacuation model of metro station

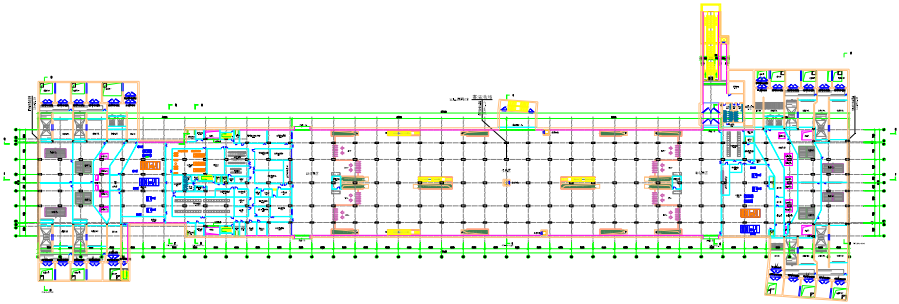


Fig. 2. The hall layer diagram of metro station

3 Assume of the Evacuation Simulation

The walking speed of passengers ranges from 0.97m/s to 1.63m/s.

The boundary layer width of all the exits is 0.3m.

The obstacle value is assumed to 1. According to the user’s guide of the software of Building Exodus v4.0, the default value of the arc between the two nodes is assumed to 0 when the scenario is open area. The default is assumed to 1 when the scenarios are minor ups and downs and the value is assumed to 7 when the scenario is seat jumping [3]. The obstacle value of the gates of the ticket machine is without

Table 1. The passenger assume of the evacuation simulation

Sex	Age	Percent (%)
male	19-29	8
	30-50	30
	51-80	22
female	19-29	7
	30-50	20
	51-80	13

reference, so a field test is conducted for the accuracy of the simulation. The field data is that the evacuation time through the gates is from 2.5s to 3.0s. Thus, the final result is that the obstacle value of the gates should be assumed to 1 according to the comparison of the field data and the simulation data.

4 Simulation of Evacuation Scenarios

Table 2. The evacuation time of different scenarios

The width of emergency exit	The width of the side gates (m)	Evacuation time (s)
The emergency exit is closed	0	171.3
	0.5	244.5
	1.0	185.8
	1.5	182.8
	2.0	177.9
The width of the emergency exit is 3.5m	0	219.9
	0.5	218.8
	1.0	216.0
	1.5	217.2
	2.0	217.9
The width of the emergency exit is 4.0m	0	198.5
	0.5	198.4
	1.0	196.1
	1.5	195.6
	2.0	194.6
The width of the emergency exit is 4.5m	0	179.6
	0.5	179.8
	1.0	178.6
	1.5	179.8
	2.0	181.8
The width of the emergency exit is 5.0m	0	167.8
	0.5	165.9
	1.0	166.7
	1.5	164.0
	2.0	163.7
The width of the emergency exit is 5.5m	0	157.3
	0.5	158.3
	1.0	156.9
	1.5	156.2
	2.0	156.8
The width of the emergency exit is 6.0m	0	146.7
	0.5	165.1
	1.0	146.4
	1.5	147.2
	2.0	147.7



Fig. 3. Ticket gates



Fig. 4. Side gate

5 Analysis of Different Evacuation Scenarios

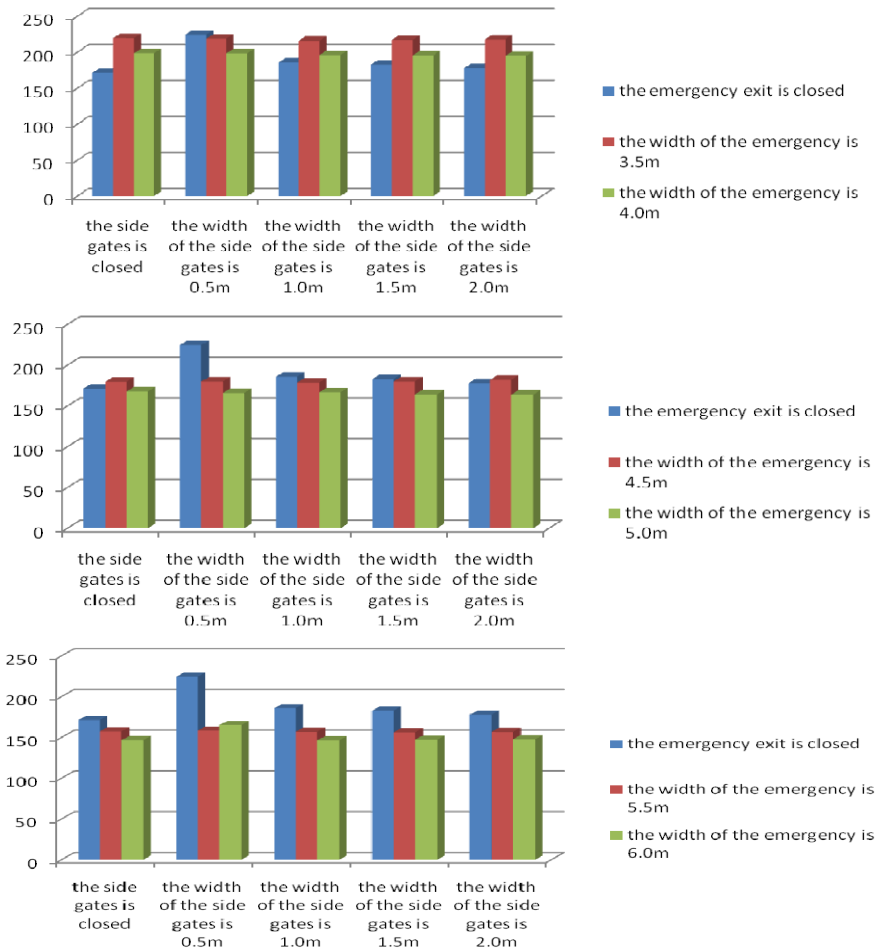


Fig. 5. The comparison of the evacuation time in different scenarios of metro station

6 Conclusion

The corresponding simulation analysis results of traffic capacity for ticket gates of metro station could be concluded as followings:

(1) As for the ticket gates, the obstacle value in software of Building Exodus 4.0 is not clarified in the user's guide. From the field test, this paper specifies the obstacle value of the gates to be 1. The value provides a useful reference to the other design of metro stations under construction.

(2) We could find from the simulation results when the emergency exit is open and the width of the side gates is different, the evacuation time is very close. Hence, it could be concluded that when the emergency exit is open, the evacuation time is mainly controlled by the width of the emergency exit.

(3) The evacuation time decreases monotonously when the width of the side gates increases. If the emergency exit is too small, the evacuation time is larger than that when the emergency exit is closed. Therefore, the appropriate width of the emergency exit should be considered during the design process of the metro station.

(4) There is no doubt that the effect of side gates is important. Generally, the evacuation time is decreased when the width of side gates is increased. When the emergency exit is closed, the side gate with the width of 0.5m is an exception. Thus, the width of side gates should be controlled carefully and comprehensively. From this paper we could get the conclusion that side gates should be opened when passengers evacuate from the metro station during emergency situation. Nevertheless, the width of side gate could not be too small and the width of 0.5m especially should be avoided.

Acknowledgments. The support of the Natural Science Foundation of China (Grant No. 50906063) is gratefully appreciated.

References

1. Congling, S., Zhong, M., Xuran, T.: Fire Experiment and Numerical Analysis of the Underground Station. Science Press, Beijing (2009) (in Chinese)
2. Tian, J., Zhou, X., Li, J.: Influence of Automatic Fare Gate in Subway on Pedestrian Evacuation. *J. Fire Safety Science* 15, 38–43 (2006)
3. Information from software of Building Exodus v4.0 User's Guide (2006)

Analysis of Compensation Strategies for Dynamic Voltage Restorer Based on DSFR-PLL Control*

Xiaoying Zhang^{1,2}, Jun Yan¹, Zhiwei Wen³, and Dan Jin³

¹ College of Electrical and Information Engineering,
Lanzhou University of Technology, Lanzhou, Gansu Province, China

² Key Laboratory of Gansu Advanced Control for Industrial Processes,
Lanzhou, Gansu Province, China

³ Gansu Electric Power Research Institute, Lanzhou, Gansu Province, China
zhxy525@gmail.com

Abstract. Dynamic voltage restorer (DVR) has been considered to be a cost-effective custom power device, which alleviates voltage sags. Recently, the uninterrupted dynamic voltage restorer (UDVR), which is one of the four topologies of DVR, is more used. According to the loads' sensitivity to voltage amplitude and phase jump, UDVR adopts different compensation strategies. This paper presents the pre-sag compensation strategy and in-phase compensation strategy based on decoupled synchronous frame phase-locked loop (DSRF-PLL). Then, an improved integrated compensation strategy is proposed. The simulation test with Matlab/Simulink for single-phase voltage sag compensation verifies that the proposed strategy, compared with in-phase compensation, can mitigate the phase jump accompanied voltage sags. And the simulation results also verify that the proposed strategy can decrease the amplitude of injected UDVR voltage, compared with pre-sag compensation strategy.

Keywords: voltage sag, dynamic voltage restorer (DVR), double synchronization frame, phase-locked loop (PLL).

1 Introduction

Along with the development of modern science and technology, control of most of the advanced manufacturing systems is mainly based on semiconductor devices, which causes such loads to be more sensitive against power system disturbances [1, 2]. Hence, the power quality problems, which encompass a wide range of disturbances such as voltage sags/swells, flicker, harmonic distortion, impulse transient and interruptions, have gained more interest recently. Voltage sags have become the dominant form of power quality problem [3]-[6], and originate from faults on the transmission and distribution systems that are caused by various events, such as animal contact, storms, equipment failure, and insulator failures (due to vandalism) [7]. Thus, it can be seen that voltage sags are inevitable in transmission and

* This work was supported by The National Natural Science Foundation of China (50967001) and The Research Project of Gansu Electric Power Research Institute (2011105010).

distribution systems. However, voltage sags can cause huge losses. For example, an interrupted automotive assembly line cost one US manufacturer US\$250000, and interruptions to semiconductor hatch processing cost US\$30000-US\$1 million per incident [8].

The dynamic voltage restorer (DVR) is known as a cost-effective device to alleviate voltage sags, and conventional DVR consists of energy storage, converter, passive filter and series injection transformer. Utteriorly, the uninterrupted dynamic voltage restorer (UDVR), which is one of the four topologies of DVR, is more used. A UDVR, which can alleviate voltage sags uninterrupted, has a simple construction, and it is easy to control. Thus, the UDVR will be widely used in power systems.

Conceptually, during voltage sags, the UDVR injects a voltage to restore the load supply voltages and exchanges active/reactive power with the surrounding system. It is quite usual for the real power requirement of the UDVR be provided by the energy storage device in the form of a battery, a capacitor bank, or a fly-wheel. However, the larger the capacity of energy storages, the higher the price of device is. Therefore, it is necessary to provide proper compensation strategies in order to reduce the active power exchange of UDVR.

Widely used in present UDVR control are pre-sag compensation strategy and in-phase compensation strategy. Obviously, the injecting voltage of UDVR is relevant to the depth/duration of voltage sags, the load characteristic and the power coefficient. This paper proposes an improved integrated compensation strategy, and the proposed strategy can be implemented based on double synchronous reference frame phase-locked loop (DSRF-PLL).

2 Decoupled Synchronous Reference Frame Phase-Locked Loop

When single-phase voltage sags take place in three-phase power system, they cause a voltage unbalance by generating negative-sequence and zero-sequence voltages. The voltage unbalance will cause an oscillation error in the measurement of the phase angle using conventional software phase-locked loop. Thus, the decoupled synchronous reference frame phase-locked loop (DSRF-PLL) method was used to track the fundamental positive component of the source voltages. Fig. 1 shows the control block diagram of DSRF-PLL. ω_0 is the angular speed of source voltages, ω^* is the angular speed of DSRF-PLL, and ω represents the error between ω^* and ω_0 . DSRF-PLL can detect the phase angle of positive-sequence voltage without oscillation error caused by negative-sequence and zero-sequence voltages, and the mathematic proof is illustrated as follows. Firstly, we can define the instantaneous source voltages as

$$\begin{cases} u_a = u_{a1} + u_{a2} + u_0 \\ u_b = u_{b1} + u_{b2} + u_0 \\ u_c = u_{c1} + u_{c2} + u_0 \end{cases} \quad (1)$$

In (1), u_a^+ , u_b^+ and u_c^+ consist of the positive-sequence components; u_a^- , u_b^- and u_c^- consist of the negative-sequence components; u_0 is the zero-sequence voltage.

Secondly, Clark transformation can be used in order to eliminate the zero-sequence voltage u_0 from three-phase voltages:

$$\begin{bmatrix} u_\alpha \\ u_\beta \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} \tag{2}$$

Here, u_α and u_β can be also separated into positive-sequence components (u_α^+, u_β^+) and negative-sequence components (u_α^-, u_β^-):

$$\begin{aligned} \begin{bmatrix} u_\alpha(t) \\ u_\beta(t) \end{bmatrix} &= \begin{bmatrix} u_\alpha^+(t) \\ u_\beta^+(t) \end{bmatrix} + \begin{bmatrix} u_\alpha^-(t) \\ u_\beta^-(t) \end{bmatrix} \\ &= \sqrt{3}U^+ \begin{bmatrix} \sin(\omega_0 t + \theta^+) \\ -\cos(\omega_0 t + \theta^+) \end{bmatrix} + \sqrt{3}U^- \begin{bmatrix} \sin(\omega_0 t - \theta^-) \\ \cos(\omega_0 t - \theta^-) \end{bmatrix} \end{aligned} \tag{3}$$

Where, U^+ is the rms value of positive-sequence voltage, U^- is the rms value of negative-sequence voltage, θ^+ and θ^- are the initial phase angles of positive-sequence and negative-sequence voltages, respectively.

Delay u_α and u_β for T/4, and finally the positive components (u_α^+, u_β^+) are obtained:

$$\begin{bmatrix} u_\alpha^+ \\ u_\beta^+ \\ u_\alpha^- \\ u_\beta^- \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} u_\alpha(t) \\ u_\beta(t) \\ u_\alpha(t - T/4) \\ u_\beta(t - T/4) \end{bmatrix} \tag{4}$$

Obviously, u_α^+ and u_β^+ are free from the influence of negative-sequence and zero-sequence voltages. It proves that DSRF-PLL can alleviate the oscillation error caused by negative-sequence and zero-sequence voltages.

In Fig.1, u_α^+ and u_β^+ are converted into double synchronous reference frame, by using the output angle θ^* of DSRF-PLL. The u_d^+ and u_q^+ are calculated as

$$\begin{aligned} \begin{bmatrix} u_d^+ \\ u_q^+ \end{bmatrix} &= \begin{bmatrix} \cos \theta^* & \sin \theta^* \\ -\sin \theta^* & \cos \theta^* \end{bmatrix} \begin{bmatrix} u_\alpha^+(t) \\ u_\beta^+(t) \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{3}U^+ \sin(\omega t - \theta^+) \\ \sqrt{3}U^+ \cos(\omega t - \theta^+) \end{bmatrix} \end{aligned} \tag{5}$$

In (5), the positive-sequence components u_d^+ and u_q^+ are decoupled with negative-sequence components. u_d^+ is relevant to ω , and ω may be eliminated by

using a PI controller. Under locked-loop condition, when ω is equal to zero, u_d^+ and u_q^+ can also be described as

$$\begin{bmatrix} u_d^+ \\ u_q^+ \end{bmatrix} = \begin{bmatrix} -\sqrt{3}U^+ \sin \theta^+ \\ \sqrt{3}U^+ \cos \theta^+ \end{bmatrix} \tag{6}$$

Therefore, the DSRF-PLL can track the positive-sequence component of source voltages by regulating u_d^+ . When the source voltages are ideal and balanced, the dq components of source voltages u_d^* and u_q^* appear as dc values. Set u_d^* and u_q^* as the reference value of u_d and u_q , then the output voltage u_{ca} , u_{cb} and u_{cc} represent the injected compensation voltages of UDVR. See Fig. 1.

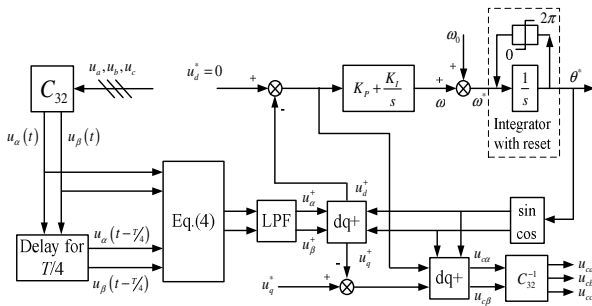


Fig. 1. Control block diagram of SPLL

3 Implement of Compensation Strategies

UDVRs operate to maintain the load supply voltage at its rated value and exchanges active/reactive power with the surrounding system. Hence, it is necessary to provide proper compensation strategies in order to reduce the active power exchange of UDVR. Widely used in present UDVR control are pre-sag compensation strategy and in-phase compensation strategy. Usually, UDVR adopts different compensation strategies according to the loads' sensitivity to voltage amplitude and phase jump. Toward some loads which are very sensitive to voltage amplitude and phase jump, the optimal method is pre-sag compensation strategy; On the other hand, toward the other loads, UDVR could adopt in-phase compensation strategy. Here, the pre-sag compensation strategy and in-phase compensation strategy are implemented based on decoupled synchronous reference frame phase locked loop (DSRF-SPLL). Then, an improved integrated compensation strategy is proposed.

3.1 Pre-sag Compensation

The pre-sag compensation strategy tracks supply voltage continuously and restores load voltage to the pre-sag condition. In this method, the load voltage can be restored ideally, but injected active power cannot be controlled. Fig. 2 shows the single-phase vector diagram of the pre-sag compensation strategy.

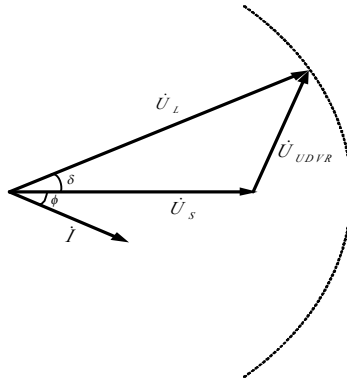


Fig. 2. Pre-sag Compensation strategy strategy

Considering the phase jump accompanied voltage sags, the phase shift angle φ can be described as:

$$\varphi = \arctan \frac{u_d^+}{u_q^+} - \theta^+ \tag{7}$$

In (6), (7) and Fig. 1, if postulate $\theta^+ = 0$ and $u_d^* = 0$, then $u_d^+ = 0$ and $\varphi = 0$. Here, φ cannot reflect the phase jump during voltage sags. On the other hand, if set $u_d^* = u_d^+$, then φ can reflect the phase jump. Hence, when UDVR adopts the pre-sag compensation strategy, make sure u_d^* be equal to u_d^+ .

3.2 In-Phase Compensation

Fig. 3 shows the single-phase vector diagram of the in-phase compensation strategy.

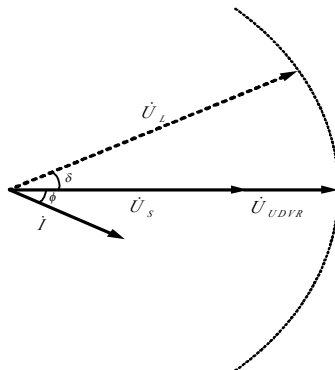


Fig. 3. In-phase compensation strategy

The injecting voltage \dot{U}_{UDVR} is in phase with the supply voltage regardless of the load current and the pre-sag voltage. The advantage of this method is that the magnitude of injected UDVR voltage is minimized for constant load voltage magnitude. However, if UDVR adopts the pre-sag compensation strategy, the injecting active power cannot be controlled, and it could result in poor load ride-through capability for long-duration sags [14]. In Fig. 1, if set $u_d^* = 0$, then UDVR adopts the in-phase compensation strategy, which is the most widely used compensation strategy.

3.3 Improved Integrated Compensation Strategy

UDVR which adopts this method can maintain the amplitude of load voltage to normal value and alleviate the phase jump during the period of voltage sags. Fig. 4 illustrates how the proposed method is able to alleviate the phase jump. In early stage of compensation, the load voltage is in phase with the pre-sag voltage, and in later stage of compensation, the load voltage is in phase with the supply voltage. That's to say, during the compensation, the injected compensation voltage \dot{U}_{UDVR} varies from $\dot{U}_{UDVR(pre-sag)}$ to $\dot{U}_{UDVR(in-phase)}$. With this method, compared with in-phase compensation, the proposed method can alleviate the phase jump accompanied voltage sags. Compared with pre-sag compensation strategy, the method can also decrease the amplitude of injected UDVR voltage.

Fig. 5 shows that the frequency will fluctuate during the compensation if UDVR adopts the method to alleviate the phase jump. However, this frequency fluctuate is allowed by GB/T 15945-2008 (Chinese national standard).

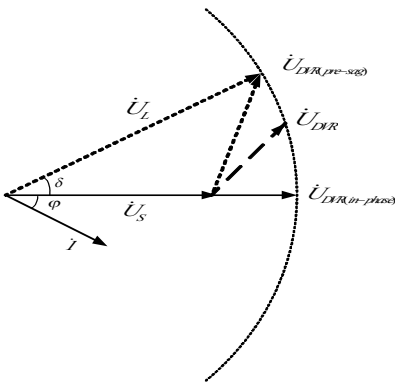


Fig. 4. Improved integrated compensation strategy

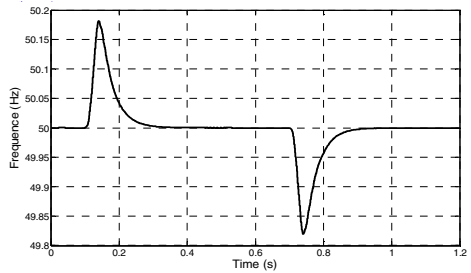


Fig. 5. Frequency fluctuate of load voltage

4 Simulation

To verify the feasibility of the proposed compensation strategy, we conduct the simulation by use of Matlab 7.6. This section gives the simulation results when using the three kinds of compensation strategy.

Fig. 6 shows the waveform of phase-A. According to Fig. 6, the nominal value of foundational component voltage is 220 V (rms), and the initial angle is 0° ; from 0.1s to 0.25s, the rms value declined to 176 V, and the phase jump is 30° . Fig. 7 shows the waveform of load voltage, and Fig. 8 shows the angle between the load voltage and supply voltage. In Fig. 7 and Fig. 8, obviously, there is a phase jump when using the in-phase compensation strategy. It also shows that the proposed strategy can mitigate the phase jump accompanied voltage sags. Fig. 9 illustrates the waveform of compensation voltages, and shows that the amplitude of injected UDVR voltage is highest when using the pre-sag compensation strategy.

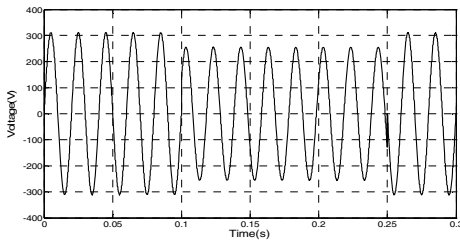


Fig. 6. Voltage waveform of phase A

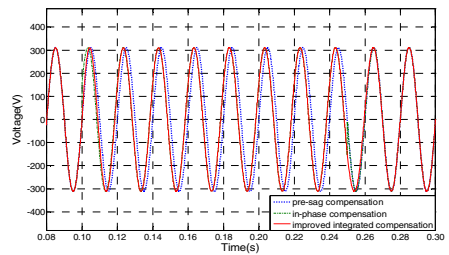


Fig. 7. Compensation results of load voltage waveform

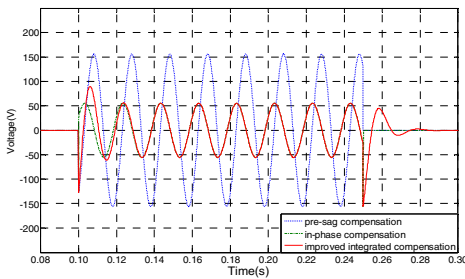


Fig. 8. Compensation voltage waveform of UDVR

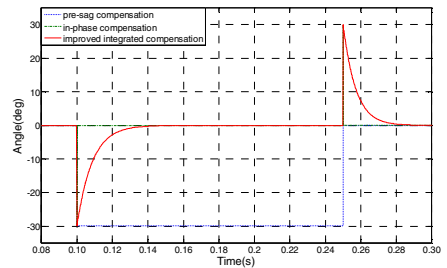


Fig. 9. Angle between the load voltage and supply voltage

5 Conclusion

This paper realizes three voltage sag compensation strategies. The simulation results verify that the pre-sag compensation strategy tracks supply voltage continuously and restores load voltage to the pre-sag condition, but injected active power cannot be controlled. The advantage of the in-phase compensation strategy is that the magnitude of injected UDVR voltage is minimized for constant load voltage magnitude but it could result in poor load ride-through capability for long-duration sags. Therefore, this paper proposes an improved integrated compensation strategy based on decoupled synchronous reference frame phase locked loop (DSRF-PLL). The simulation results verify that the proposed strategy can mitigate the phase jump and decrease the amplitude of injected UDVR voltage.

References

- [1] Kumar, G.V.N., Chowdary, D.D.: DVR with Sliding Mode Control to alleviate Voltage Sags on a Distribution System for Three Phase Short Circuit Fault. In: Proceedings of 3th IEEE International Conference on Industrial and Information Systems, Kharagpur, INDIA, pp. 1–4 (2008)
- [2] Peng, C., Chen, Y., Sun, J.: Study of dynamic voltage restorer and its detecting method. *Electric Power Automation Equipment* 23(1), 68–71 (2003)
- [3] Wang, T., Xue, Y., Choi, S.: Review of dynamic voltage restorer. *Automation of Electric Power Systems* 31(9), 101–107 (2007)
- [4] Hamzah, N., Muhamad, M.R., Arsad, P.M.: Investigation on the effectiveness of dynamic voltage restorer for voltage sag Mitigation. In: Proceedings of 5th Student Conference on Research and Development, Selangor, Malaysia, pp. 1–6 (2007)
- [5] Ge, C., Cheng, H., Wang, X., Zhong, M., Chen, G., Miao, Y., et al.: DVR control algorithm based on minimal energy. *Electric Power Automation Equipment* 29(1), 70–74 (2009)
- [6] Abi-Samra, N., Carnovale, D., Malcolm, W.: The role of the distribution system dynamic voltage restorer in enhancing the power at sensitive facilities. In: Proceeding of WESCON 1996, Anaheim, CA, USA, pp. 167–181 (1996)
- [7] Xiao, X.: Analysis and control of power quality. China Electric Power Press, Beijing (2004)
- [8] Hunter, I.: Power quality issues-a distribution company perspective. *IEEE Power Engineer* 15(2), 75–80 (2001)

Design of H.264 Encoder Based on ADSP-BF561

Qinghui Wang¹, Guian Wu², and Ying Gao

^{1,2} College of Information Engineering,
Shenyang University of Chemical Technology,
Shenyang, China
wangqh8008@vip.sina.com, wga1814@163.com

Abstract. According to the H.264 recommendation, a video encoder is designed to implement high quality video flow transfer based on AD DSP. Hardware platform of the encoder is an image processing equipment based on DSP (BF561) and video capture uses OV7660. Video capture and data format transformation become easier when DMA is used. The result shows that it is possible to implement a real-time video encoder conformed to the H.264 standard on a single BF561 with high quality and low bandwidth.

Keywords: H.264, encoder, BF561, embedded system.

1 Introduction

With the increasing bandwidth of Internet, Internet-based voice image information, especially real-time data transmission has become possible. Video and audio transmission data takes full advantage of Internet bandwidth, and achieve real-time communication and remote meetings possible. Compared to traditional conference calls, video conferencing system takes full advantage of Internet resources, thereby reducing operating costs.

Video Encoder is a major component of the video conference system. Current video compression standards have H.261, H.263, MPEG4 and H.264 and so on. Compared with previous standards, new generation of video coding standard H.264 with a high compression rate, network compatibility is good, video quality is superior and so on. H.264 introduces many of the current video coding new technologies[1], makes the same reconstructed image quality, the coding efficiency than H.263 and MPEG-4 high about 50%. Therefore, video encoder is designed with H.264 coding standard, the encoder mainly used for the U.S. AD's BF561 DSP chip. Currently, the video processing program has also been turned DSP platform from the ASIC solutions[2]. DSP platform for video on the product development have the following advantages: First, the user-developed degree of freedom is greater, support a variety of personal development, can continue to adapt to the market's new demands, improve product performance in the first, and enhance the competitiveness of our products; Second, DSP has a strong processing power, while the DSP in a multi-channel audio and video signal compression at the same time; Third, short development cycle, to achieve rapid technological updates and new generation of products, emerging fast and optimization algorithm can be flexible upgraded.

2 BF561 Introduction

BF561 used symmetrical dual-core architecture, in a BF561 chip integrates two BF533 DSP cores, and two cores can be clocked at up to 600 MHz, support for parallel processing[3]. BF561 features to ensure its powerful digital signal processing capabilities, and support for low-voltage low-current power supply, meet multi-functional digital consumer products for performance and power consumption requirements.

2.1 DMA Introduction

Blackfin processor transfer data using direct memory access (DMA) in memory or between memory and peripherals. DMA controller transfer data between in the memory and chip peripherals (peripheral DMA), and transfer data between L1/L2/L3 (Memory DMA, or MDMA).DMA controller is a key component of the Blackfin processor architecture, completely independent of the kernel, not the cycle stealing, no need to take up the processor core cycle. In an ideal application configuration, the kernel need only set up the DMA controller and the data call process the interrupt.

BF561 has 3 independent DMA controllers DMA1, DMA2 and IMDMA. DMA1 and DMA2 controllers have 12 peripheral DMA channels and 4 DMA channels. IMDMA controller has 4 DMA channels.

2.2 Descriptor-Based DMA

Descriptor-based DMA transfer requires the parameters stored in memory to initialize a DMA queue. Descriptor includes all of the DMA control register to be normal programmed parameters. Descriptor allows multiple DMA queues together. In descriptor-based DMA operation, a DMA channel can be programmed to be completed in the current transmission queue, automatically set and start the other DMA transfer process. In the management of a system DMA transfer process, the descriptor-based model provides maximum flexibility.

Descriptor list model divided into descriptor list "small" model and descriptor list "big" mode. In the descriptor list "small" mode, described section including a 16-bit field to point to the next descriptor entry the low 16-bit address, the high 16-bit address via registers programmed and remain the same, limit descriptors in memory of a specific size of 64 KB pages. When descriptor need cross-page can be used to provide 32-bit entry address of the descriptor list "big" mode.

3 H.264 Video Encoder Constitute

H.264 encoder consists of three parts of video capture, data format conversion and H.264 encoding. Video capture part is responsible for capturing images, and images captured by the PPI interface filled to the specified video frame buffer. Data format conversion part completion of input 4:2:2 format images converted into H.264 encoder to encode the 4:2:0 format date. H.264 encoding part responsible for the 4:2:0 format image coding[4].

During the video encoder design, BF561 A core for running the operating system and protocol stack, while the B core implements H.264 algorithm.

3.1 Video Capture

Video capture using the camera OV7660, OmniVision's OV7660 the United States developed is a CMOS color image sensor chip[5], support VGA, QVGA, GIF and many other resolutions. Video output format Raw RGB, GRB 4:2:2 and YUV/YCbCr (4:2:2). In the design, select the CIF YVYU (4:2:2) format, and need to set up the corresponding register COM1 = 0X00, CLKRC = 0X80, COM7 = 0X30, TSLB = 0X05.

After configuring the camera, open the CPIO the video data into the Blackfin video frame buffer. Descriptor-based PPI DMA descriptor can be easily achieved ping-pong buffering[6], thus ensuring that will not coverage has not been disposed of input data.

Ping-pong buffer shown in Figure 1.

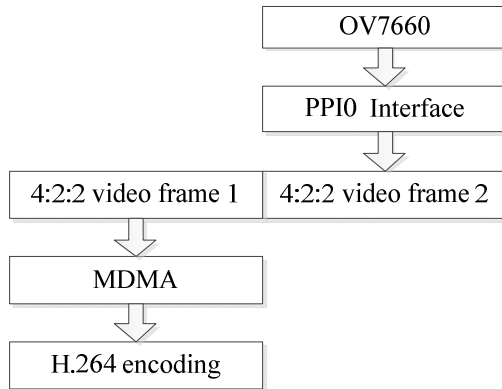


Fig. 1. Processing the first video frame also filling the second video frame

Set the input buffer of two CIF 4:2:2 frame size, First, fill the input data via the PPI interface to 4:2:2 video frame 1, when the first frame data fill is full, processor performs MDMA and compression encoding and other operations for this frame date, at the same time, PPI continues to fill the second video frame. The use of ping-pong buffer, data will continuously fill into between the two video frames.

3.2 Data Format Conversion

Video capture interwoven 4: 2: 2 YUV video data, while the H.264 video coding algorithm for 4:2:0 format video data compression. 4:2:0 data of luminance values buffer and chrominance values buffer are separated, the use of EDMA can be achieved the separation of luminance buffer and chrominance buffer[7].

CIF YVYU (4:2:2) format data shown in Table 1, CIF 4:2:0 format data shown in Table 2. 4:2:2 format image of a macro-pixel consists of 4 Y component, 2 U component and 2 V components. 4:2:0 format image of a macro-pixel consists of 4 Y

component, 1 U component and 1 V components. Y, U and V storage area are separated, Y first, then U, and finally V.

Table 1. CIF YVYU 4: 2: 2

	0	1	2	3	4	5	6	7	...	700	701	702	703
1	Y	V	Y	U	Y	V	Y	U	...	Y	V	Y	U
2	Y	V	Y	U	Y	V	Y	U	...	Y	V	Y	U
3	Y	V	Y	U	Y	V	Y	U	...	Y	V	Y	U
4	Y	V	Y	U	Y	V	Y	U	...	Y	V	Y	U
5	Y	V	Y	U	Y	V	Y	U	...	Y	V	Y	U
...								...					
286	Y	V	Y	U	Y	V	Y	U	...	Y	V	Y	U
287	Y	V	Y	U	Y	V	Y	U	...	Y	V	Y	U

The 4:2:2 format converted to 4:2:0 format, to retain all of the Y, and take 0,2,4, ... line of U and 1,3,5 ... line of V[8]. Based on list of descriptors the "big" mode MDMA, need three pairs of descriptors source_id, dest_y, SOURCE_u, dest_u, SOURCE_v, dest_v, Form the source and purpose of two descriptor list, The Y, U and V from the mixed 4:2:2 source data area into Y, U and V the purpose of separating the data area of the move.

When a CIF 4:2:2 format data filled by the PPI to a Blackfin processor's video frame buffer, generate an interrupt, the interrupt handling routine start MDMA, separates the luminance values and chrominance values, and then to H.264 encoder to encode.

Table 2. CIF YVYU 4: 2: 0

	0	1	2	3	4	5	6	7	...	348	349	350	351
0	Y	Y	Y	Y	Y	Y	Y	Y	...	Y	Y	Y	Y
...									...				
287	Y	Y	Y	Y	Y	Y	Y	Y	...	Y	Y	Y	Y
288	U	U	U	U	U	U	U	U	...	U	U	U	U
...									...				
359	U	U	U	U	U	U	U	U	...	U	U	U	U
360	V	V	V	V	V	V	V	V	...	V	V	V	V
...									...				
431	V	V	V	V	V	V	V	V	...	V	V	V	V

3.3 H.264 Encoding

Waits for a 4:2:0 format data, the frame data executive H.264 encoding, and then continue to wait for 4:2:0 format data generation, until the encoder end.

4 Experimental Results and Analysis

In experiment, cameras capture images encoded by the H.264 encoder, and the compressed bit stream file is stored in the local memory. After encoding, with H.264 decoder for decoding the compressed bit stream, decoded with YUVviewerPlus.exe playing video files.

Encoder can achieve 15 frames per second encoding rate in the experiment. Decoded image clarity compared with the using the Image Viewer View in VisualDSP++5.0[9], the image quality is basically the same, and image is also clear. OV7660 capture images speed is 30 frames per second, H.264 encoder did not meet the real-time encoding.

In the experiment, based on H.264 encoder of BF561 did not achieve the real-time encoding, because there is no optimization for H.264 open source code. As the CIF format image data is relatively large, need to be stored in external memory, the processor to access external memory rather a long time; DCT and motion estimation algorithm function is time-consuming, using C language to achieve. There are some auxiliary functions and print information in the program, thereby affecting the speed of the encoder interface[10].

5 Conclusion

The H.264 BF561-based encoder does not achieve real-time encoding. Focus of future work is to use C compiler of VisualDSP++ development environment for code optimization; Full use of internal resources of storage space as far as possible through DMA, cache and other to reduce the chip and external data scheduling of storage allocation impact on system performance; For some system calls more frequently, more time-consuming modules, such as DCT and motion estimation, full use of the instruction set of BF561, and using assembly language; Remove the source of unnecessary auxiliary functions and print information in order to achieve real-time encoding of H.264 encoder based on BF561.

With the consumer electronic products further enter the ordinary family, the application of H.264 standard video conferencing, video telephony, and unmanned surveillance systems will also be used more and more widely. clear.

References

1. Analog Devices Inc. ADSP-BF561 Blackfin Processor Hardware Reference [EB/OL]. Revision 1.1 (2007)
2. ITU-T REC. H.264 ISO/IEC 14496-10 AVC, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification. Thailand: The 7th Meeting: Pattaya 72, 7214 (2003)

3. H.: 264/ MPEG - 4 Part 10 White Paper [EB/OL], OV7660/ OV7161 CMOS VGA (640×480) Camera Chip Implementation Guide. Rev. 1 (2004), <http://www.vcdex.com>
4. Bi, H.: A new generation of video compression coding standard H.264/ AVC. Posts and Telecommunications Press (2005)
5. OV7660/ OV7161 CMOS VGA (640×480) Camera Chip Implementation Guide. Rev. 1 (2004)
6. Yan, F., Chen, Z., Yuan, J., Katz: Embedded media processing. Electronic Industry Press, Beijing (2007)
7. Analog Devices Inc. Video Framework Considerations for Image Processing on Blackfin Processors (EE-276). Rev.1 (2005)
8. Liu, F.: Video coding technology and international standards. Beijing University of Posts and Telecommunications Press, Beijing (2005)
9. Analog Devices, Inc. VisualDsp++ 5.0 Device Drivers and System Services Manual for Blackfin Processors. Revision 3.0 (2007)
10. Su, H., Kong, D., Yu, S.: ADSP-BF533-based real-time image coding key technologies. Computer Engineering and Applications 4, 88–90 (2005)

Optimized Implementation of H.264 Integer Transform Based on Blackfin 533

QingHui Wang¹, HuaPing Xue², and JingJing Yang³

College of Information Engineering, Shenyang University of Chemical Technology,
Shenyang, 110142, China
wangqh8008@vip.sina.com, xue_hp@yahoo.cn

Abstract. The integer transform implementation is optimized based on the Blackfin533 according to integer transform algorithm in H.264. Blackfin533 system resources and the feature of Blackfin533 instructions are considered. Parallel and vector instructions are used to optimize the algorithm. The optimized program runs in VDSP++4.5 software environment and executive efficiency of the program is tested using the function of testing cycles. It is proved that executive efficiency of the assembler that is optimized is increased significantly compared with the integer transform in X.264.

Keywords: H.264 Standard, Integer Transform, Blackfin533, Assembly Optimization.

1 Introduction

H.264 standard is the latest development standards targeting practical application, jointly developed by the International Telecommunications Union (ITU), International Standards Organization (ISO) and Joint Video Team (JVT) [1]. As the complexity of the algorithm of H.264 standard is higher more than that of H.263 2 times, so algorithm itself has a large amount of computation and Software implementations for DSP chips is very high. It is very necessary to study assembly-optimized implementation of its Core algorithm in the DSP targeting features of H.264. And the computation of Integer transform part In H.264 algorithm accounts for a large proportion of the entire coding process, Therefore, assembly-optimized implementation of Integer transform plays a significant role for improving efficiency of the entire coding algorithm and reducing the complexity of the algorithm.

2 Basic Principles of Integer Transform

Each macro block will be transformed, quantized and encoded in the H.264 standard. According to difference of residuals data type, three kinds of integer transform adopted in the H.264 standard is as follows [2]: 4×4 block of integer DCT transform targeting all of residuals data, 4×4 luma DC coefficient transform (only in 16×16 intra-macroblock prediction mode), 2×2 chroma DC coefficient transform. This paper focuses on 4×4 integer transform of all residual data.

Integer transform is derived from the traditional DC Ttransform, integer DCT transform formula can be derived from a variety of formula [3]:

$$Y = (C_{\mu} X C_f^T) \otimes E_f = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \end{bmatrix} \quad (1)$$

To simplify the implementation of the transform and the transform remains orthogonal, a and b needs to be modified so that: $a=1/2$, $b = \sqrt{10}/5$.

This is the integer transform formula used in H.264, the ‘core’ part of the transform, CXCT, can be carried out with integer arithmetic using only additions, subtractions and shifts. The post-scaling operation $\otimes E_f$ requires one multiplication for every coefficient which can be ‘absorbed’ into the quantisation process. Therefore, the actual H.264 integer transform output:

$$Y = CXCT^T \quad (2)$$

In practical applications, H.264 well factorise matrix multiplication of equation (2) to quadratic one-dimensional integer transform [4]. For example, in the first, one-dimensional integer transform is applied in each row of the image and residual blocks, and also applied in each column of transformed blocks. And each one-dimensional integer transform can be used radix-4 butterfly fast algorithm, this can save computing time.

For the 4x4 pixel block, the original pixel values of defined a row or column are a, b, c and d, and that of transformed pixel are one-dimensional integer. Transformation steps shown in Table 1, where u, v, y and z is intermediate variables generated by One-dimensional transforms, >> indicates a binary shift right.

Table 1. 16-bit integer transform operation process

Transform name	Input value	Intermediate steps	Output value
Integer orthogonal transformation	a, b, c, d	$u = a + d;$ $v = b + c;$ $y = b - c;$ $z = a - b.$	$A = u + v;$ $C = u - v;$ $B = y + (z \ll 1);$ $D = z - (y \ll 1).$

One-dimensional butterfly fast algorithm shown in Fig.1.

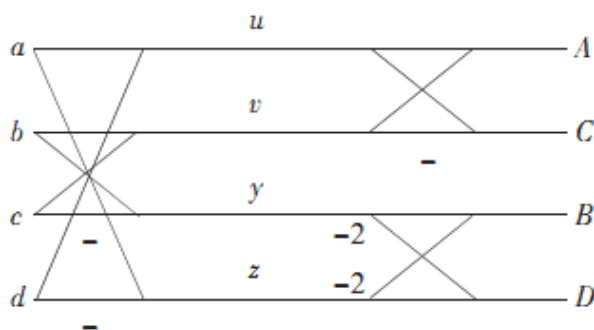


Fig. 1. One-dimensional butterfly fast algorithm

3 Introduction of Blackfin 533 Structure and VDSP++ Software Environment

Blackfin533 [5] is industry-leading embedded multimedia processor of ADI introduced. It has a computing speed of up to 750 MHz DSP core, and very suitable for real-time video codec, especially for power demanding sports video processing applications.

This optimization of the integer transform uses VDSP++4.5 software environment to run and debug assembler optimized integer transform. VDSP++ is a convenient development platform, ADI specifically developed, targeting ADI's DSP devices, it exchange information with the user by the way the graphics window, and operate targeting the processor. VDSP++ integrates two big part: integrated development environment (IDE) and debugger, and provides a more powerful program development and debugging. VDSP++ has a flexible management system, and provides a set of tools for the processor to application and project development, and its development kit integrates a set of tools components the DSP program development needed. The program can compare the difference before and after optimization of the efficiency by using the function of statistical program execution cycles of VDSP++ environment.

4 Optimized Implementation of Integer Transform Based on Blackfin 533

In this paper, X.264 program C code, based on the H.264 standard, integer transform module makes a thorough optimization, final assembly statement takes up the total number of functional blocks about 90% of statements.

The program input is the 16-bit 4×4 integer matrix, which places in the array $dct[4][4]$, defining array $tmp[4][4]$ is used to store temporary variables. The program read value of the input matrix from $dct[4][4]$, firstly 4 row matrix values carry on

one-dimensional fast transform butterfly separately, and the results transpose and store in tmp [4][4]. Then 4 row matrix values of tmp[4][4] array carry on again one-dimensional matrix quickly transform butterfly, and the results transpose and store in dct [4][4], that is the final output of the program. In optimized assembler, pointer register I1 point to the beginning address of dct [4][4] array, pointer register I0 point to the beginning address of tmp [4][4] array, counter P1 = 4 is the cycle length.

The following is the optimization part of the core code in assembler:

```
LSETUP DCT_START1 DCT_END1 LC0=P1;
```

Here the software loop of C program is optimized to hardware loop of assembly. Therefore, Blackfin 533 carries on loop and jump out loop according to the value of loop register, does not need conditional branch instruction.

```
DCT_START1:
```

```
P3=I0;
```

```
I0+=2;
```

The intermediate variable is stored in register, the system is no longer need to allocate memory space for the variable, this saves a lot of time to access memory.

```
R0=R0+I+R1 R1=R0-I-R1 S;
```

```
A1=R0.L*R7.L A0=R0.L*R7.L IS; //I0 is the address of //temporary variable tmp [4][4]
```

```
R4.H= A1-=R0.H*R7.L R4.L= A0+=R0.H*R7.L IS;
```

The external data bus width of Blackfin 533 is 16-bit, it can access 4 byte at a time replace 1 byte or 2 byte, this could improve the execution speed.

```
R2=R1<<1 S;
```

```
R6.H= A1-=R2.H*R7.L R6.L=
```

```
A0+=R2.L*R7.L IS
```

```
||R1.H=W[I1++];
```

```
R4.H=R6.L*R7.L IS ||W[P3++P2]=R4.L;
```

```
R6.L=R3.L*R7.L IS ||W[P3++P2]=R4.H;
```

```
R1.L=W[I1++]+||W[P3++P2]=R6.L;
```

In here, a parallel instruction is equivalent to executive two non-parallel instruction, which saves the number of instructions.

```
DCT_END1:
```

```
W[P3]=R6.H;
```

5 Optimized Results

Table 2 is comparison of before and after optimization. Before carrying out the program rewrite optimization, VDSP++ environment is set to be Compiler Optimize, setting "compile" dialog box through the menu "Project"—"project options".

Table 2. InstructionCycles of DCT function performs

OptimizationMethod	Program without optimization, Compiler settings to Speed optimization	All rewritten as an assembly instruction
InstructionCycles	392	115

Table 2 shows, comparing with X.264 code integer transform module program, the efficiency of Optimized integer transform module assembly code increase approximately 4times.

6 Conclusions

In this paper, the assembly optimization of H.264 algorithm of integer transform algorithm is implemented on ADSP-Blackfin533. In the optimization process, taking full advantage of parallel instructions and vector instructions Blackfin provides to adjust statement structural, which effectively improve Instruction execution efficiency of the integer transform module. The reducing of integer transform execution clock sysles can greatly improve the efficiency of X.264 code processing video image. In order to implement a complete H.264 encoding based on VDSP-Blackfin533, after upon completion of each core assembler optimization algorithm, also needs to perform DMA optimization of data flow and hardware loop, this will be the next most important work

References

1. Li, Q.Y., Wu, R.Q., Feng, F.: Implement of H.264 integer DCT transform and quantization system. Video Engineering 1, 29–32 (2006)
2. Liu, B.L., Liu, G.Z., Su, R.: DSP Implementation of H.264 integer DCT transform and quantization. Microelectronics and Computer, 200–205 (2005)
3. Bi, H.F.: A new generation of video compression coding standard– H. Post & Telecom Press, Beijing (2005)
4. ITU – T REC. H.264|ISO/IEC 14496-10 AVC, Draft ITU – T Recommendation and Final Draf International Standard of Joint Video Specification (2003)
5. Feng, C.: The theory and system design of Blackfin family DSP. Electronic Industry Press, Beijing (2004)

Power Line Communication System Design for LED Transmission

Tai-Ping Sun, Chia-Hung Wang, and Jia-Yu Chen

Department of Electrical Engineering
National Chi Nan University
Nantou, Taiwan (R.O.C.)

tps@ncnu.edu.tw, s95323901@ncnu.edu.tw, s98323505@ncnu.edu.tw

Abstract. LED is one of the best energy efficient components. As a light source, LED presents advantages as lower power consumption, higher brightness and longer lifetime. This paper presents an in-depth description of applying power line communication technology to LED lighting system. By using power lines as communication media to transfer control data, LED is controlled by existing lines on the power system. In this research, we use the LED which contains RGB tricolor in its package, so the change of color dimming can be controlled by a PWM controller.

Keywords: Power Line Communication, On Off Keying, Comparator, Filter, LED, PWM.

1 Introduction

This paper seeks to argue how the LED, with its RGB tricolor, controls the change of color dimming through the application of power line communication technology. The circuits are divided into two parts: one as the power line communication, the other LED control circuit.

In this research, On-off Keying (OOK) modulation and single chip are used to produce 12-bit digital signal, which can be controlled in the power line communication. The digital signal and the 125 KHz of pulse wave signal are multiplied to produce the OOK modulation signal. Inverter which tolerates high drive current is used instead of the line driver. The inverter is connected to the transformer and coupled to the power line. OOK modulation signal is restored to digital signal through transformer, band pass filter, automatic gain control circuit, demodulator circuit and comparator in the receiver circuit.

As for the LED part, 12-bit PLC signals will be divided equally into 3 portions which transfer respectively to the RGB LED. Each color will then generate 16 variations, which will create 4096 colors after mixing. Since the color mixing is achieved by the pulse width modulation (PWM), the PWM width in this case is 6.7 %.

2 Proposed Circuit

The power line communication structure is divided into two parts. The first is a transmitter; the second is the receiver. The transmitter has three parts: OOK signal modulation, a line driver and a line coupler. The receiver has five parts: a line coupler, band-pass filter, AGC, demodulator and comparator, as shown in figure 1. [1]

The LED part structure is divided into four parts. The first is a Function generator (a). The second is a receiver (H). The third is the light-emitting diode channel pulse width modulation circuit (I). The forth is the light - emitting diode control circuit (J).

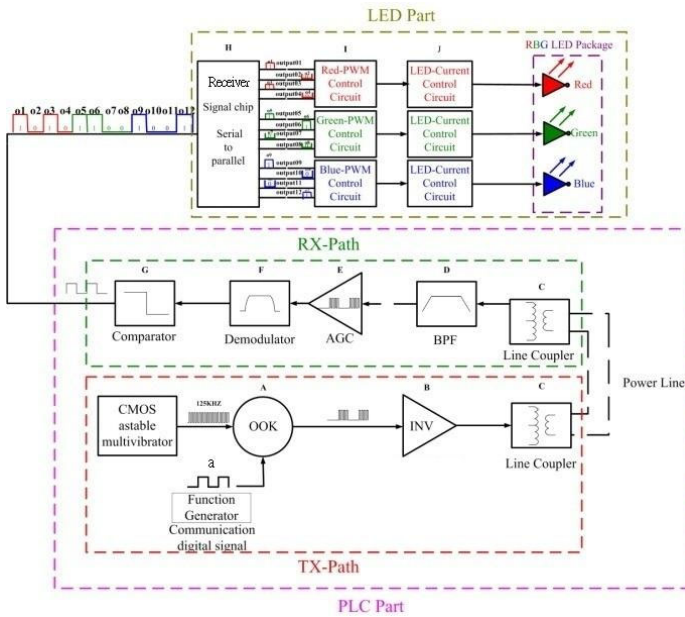


Fig. 1. The schematic diagram of PLC and LED circuit [1]

2.1 Power Line Communication Part Architecture

Transmitter is composed of a square wave generator using a CMOS astable multi-vibrator. The square wave generator is used to generate a 125 KHz pulse signal. To implement the OOK modulation technique, the 0 and 1 signal is coupled using a square wave generator circuit and a digital signal generator. [2] The power line output signals will attenuate at different impedance with changing distance. To solve this problem, the current driving capability can be improved using an inverter circuit in the transmitting stage. [3] The signal is then coupled to the power line. In the receiving stage, the transformer has a feature that can filter out the 60 Hz grid signal and receive the OOK signals. Because the power line impedance during transmission

can cause signal attenuation, the auto gain control stage is used to fix the output signal amplitude. The receiving signal recognition can be improved. Furthermore, the OOK modulation signal will mix with the power line noise, reducing the signal accuracy. Thus, a band-pass filter is used to catch the signals and filter out the noise. A demodulator circuit is used to distinguish the reverting signals to achieve the preliminary comparator work. The demodulator circuit consists of a full-wave rectifier and a peak detector circuit. The OOK modulated signals are restored into digital signals using the comparator circuit hysteresis characteristic. As shown in figure 1.

2.2 LED Part Architecture

In Figure 1, Transmitter (A) (function generator) is responsible for producing the serial signal, which contains the 1-bit start code, 12-bit control data code and 5-bit end code. The data transmission is continuous and periodic. The start code and the end code determine respectively where the data begins and ends. The control data code is designed for the light mixing and color blending of the LED. Since the data is presented in serial signal in power line communication, a converter will be indispensable for transferring the serial signal into parallel signal in order to control the LED. The receiver (H) is responsible for parsing the serial signal. It detects firstly the start code and the end code, and then converts the control data code from serial signal into parallel signal, divides the 12-bit parallel signal into 3 parts and transfers them to the next part. In other words, every color will receive a 4-bit signal, which can produce a brightness spectrum with 4 variations. The PWM technique can expand these 4 brightness variations into a brightness spectrum of 16 variations. The light-emitting diode channel pulse modulation circuit applies the concept of numeral system conversion to enhance the resolution of every color. The numeral system conversion applied here is to convert binary to decimal. Since $2^4=16$, the periodic signal of 4 time interval signals will become 16 time interval signals. Since the chip current is too low and the signals are monitored by the current, perfect circuit control is absolutely necessary to drive the LED and deliver its advantages of fine brightness intensity and a full light emitting spectrum. The light-emitting diode current control circuit is designed to drive adjustable current. This special design is applied to correct the defects that might be caused by the low current provided from the chip, which might cause the malfunction of the LED.

3 Experimental Results

Figure 2 shows how the voltage of power line communication attenuates with different distances. According to Figure 2, from the distance of 10m to 100m, the signal will attenuate accordingly from -6db to -20db.

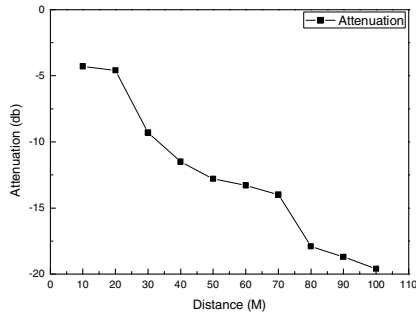


Fig. 2. The attenuated voltage trend

Table 1 is the condition table designed for the PLC. According to Table 1, the modulation scheme is OOK, which delivers 125 kbps data transmission rate within the maximum transmission distance of 100 M. The PLC control data can send 4000 bps. The PLC control data max bits is 18 bits while its pulse width is 254 us.

Table 1. The specification of PLC part

PLC part format	
Modulation scheme	OOK
Modulation data rate	125kbps
Transmission distance	100M
PLC control data rate	4000bps
PLC control data	18bits
PLC pulse width	254us

Table 2 is the condition table illustrating the LED control circuit. According to Table 2, the LED is a RGB LED with 4096 color variations. The PWM color pulse width is 30 us while the PWM control data is 4 per/bits. The LED driver current is 20mA.

Table 2. The specification of LED part

LED part format	
Color type	Red, Green, Blue
Color change	4096
PWM clock pulse width	30us
PWM control data	4 per/bits
LED driver current	20mA

Figure 3 is the CIE1931 chromaticity diagram which shows how the red LED changes accordingly to the PWM from 6.7%% to 100%.

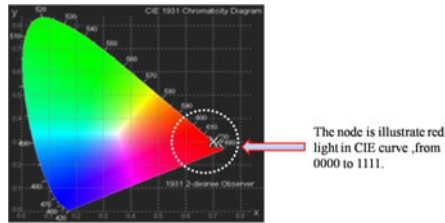


Fig. 3. The CIE1931 of red light

Figure 4 is the CIE1931 chromaticity diagram which shows how the green LED changes accordingly to the PWM from 6.7% to 100%.

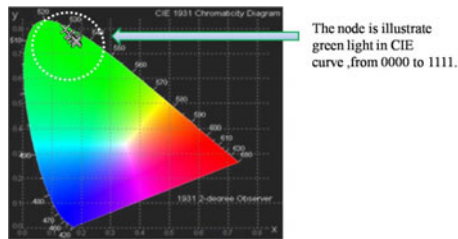


Fig. 4. The CIE1931 of green light

Figure 5 is the CIE1931 chromaticity diagram which shows how the blue LED changes accordingly to the PWM from 6.7% to 100%. According to Figure 3 to Figure 5, PWM cannot affect the position of the color as well as its wavelength. According to Figure 3 to Figure 5, to emit the ideal color of the light, it requires not the alteration of the color intensity of a particular color but the color mixing technique.

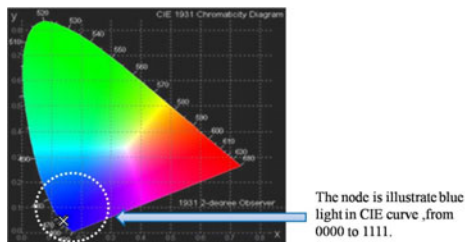


Fig. 5. The CIE1931 of blue light

The upper right corner of Figure 6 shows the PLC serial data and the signal of RGB Channel. The left of Figure 6 shows the data of LED. The bottom right corner of Figure 6 shows the data of CIE curve. Figure 6 shows the serial data format to

produce white light, the white light spectrum diagram, and the CIE1931 chromaticity diagram. When the transmission code is 101010100011, the PWM of red is 60%, the PWM of green 60%, and the PWM of blue 20%.

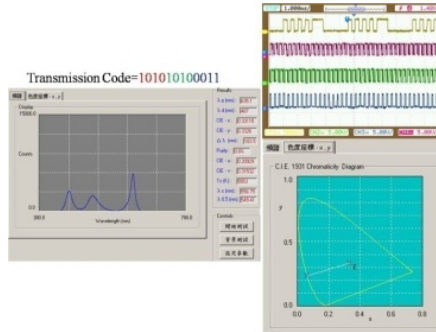


Fig. 6. The specification of white light

Figure 7 illustrates the color of the node from A to H on CIE curve 1931. The color of node A is pink. The color of node B is red. The color of node C is orange. The color of node D is yellow. The color of node E is green. The color of node F is cyan. The color of node G is blue. The color of node H is purple.

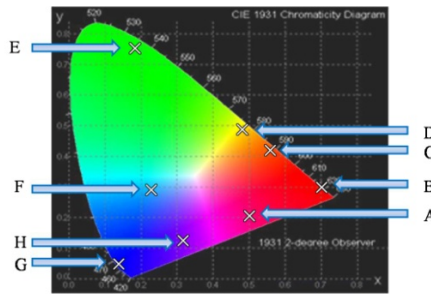


Fig. 7. The node of different color in CIE1931 curve

Tables3 provides additional information about the mixing of different colors. The serial data format for the color of pink is composed of 100100000010. Its ratio of red PWM is 60%, the ratio of green PWM 0%, and the ratio of blue PWM 13.3%. The serial data format for the color of red is composed of 111100000000. Its ratio of red PWM is 100%, the ratio of green PWM 0%, and the ratio of blue PWM 0%. The serial data format for the color of orange is composed of 100000110000. Its ratio of red PWM is 53.3%, the ratio of green PWM 20%, and the ratio of blue PWM 0%. The serial data format for the color of yellow is composed of 100001100000. Its ratio of red PWM is 53.3%, the ratio of green PWM 40%, and the ratio of blue PWM 0%. The

serial data format for the color of green is composed of 000011110000. Its ratio of red PWM is 0%, the ratio of green PWM 100%, and the ratio of blue PWM 0%. The serial data format for the color of cyan is composed of 1000111101111. Its ratio of red PWM is 53.3%, the ratio of green PWM 100%, and the ratio of blue PWM 46.7%. The serial data format for the color of blue is composed of 000000001111. Its ratio of red PWM is 0%, the ratio of green PWM 0%, and the ratio of blue PWM 100%. The serial data format for the color of purple is composed of 100100000101. Its ratio of red PWM is 60%, the ratio of green PWM 0%, and the ratio of blue PWM 33.3%.

Table 3. Experimental results of different color

Node	Data format	Red-PWM	Green-PWM	Blue-PWM	color
A	100100000010	60%	0%	13.3%	Pink
B	111100000000	100%	0%	0%	Red
C	100000110000	53.3%	20%	0%	Orange
D	100001100000	53.3%	40%	0%	Yellow
E	000011110000	0%	100%	0%	Green
F	100011110111	53.3%	100%	46.7%	Cyan
G	000000001111	0%	0%	100%	Blue
H	100100000101	60%	0%	33.3%	Purple

4 Conclusion

This paper seeks to demonstrate how the transmission of LED color mixing can be achieved through power line communication. Given the fact that all appliance application requires power to function, with successful use of the PLC technology, we can lower the cost, reduce the demand for network space, and makes the networks simple. In terms of the LED, PWM controller can expand the data with greater efficiency and enhance the resolution. By applying the current controller, the user can adjust effectively the drive current to his request and achieve the goal of energy saving and reduction of carbon emission.

Acknowledgment. The authors wish to thank the Nation Science Council, Taiwan, R.O.C. for supporting under Grant NSC 97-2120-M-260-001.

References

1. Wang, C.-H., Chen, C.-Y., Sun, T.-P.: Circuit implementation of OOK modulation for low-speed power line communication using X10 standard. In: 2011 13th International Conference on Advanced Communication Technology (ICACT), February 13-16, pp. 248–453 (2011)

2. Kang, I.: Interferometric Operation of an Electroabsorption Modulator for PSK Modulation and OOK Modulation with Performance Enhancements. In: ECOC 2006. European Conference on Optical Communications, September 24-28, pp. 1-2 (2006)
3. Chen, C.S., Ku, T.T., Lin, C.H.: Design of PLC based identifier to support transformer load management in Taipower. In: IEEE Conference Record on Industrial & Commercial Power Systems Technical Conference, May 3-7, pp. 1-6 (2009)

The Design of Multi-parameter and Portable Monitor for Sleep Apnea Status

XiaoMei Liu, Ying Lian, ShaoChun Chen, and ShanShan Tang

College of Information Engineering,
Shenyang University of Chemical Technology, Shenyang, Liaoning, China
liuxiaomei40@126.com, lianying3314@163.com,
csc2@21cn.com, tss7114331@126.com

Abstract. The low-power and multi-parameter monitor has been designed for human sleep apnea status by use of the LPC2132 microcontroller of ARM7. The $\mu\text{C} / \text{OS-II}$ real-time operating system is embedded as the platform of software design and Micro SD card is used to complete the storage management. The monitor is powered by a lithium battery and it could work at least 12 hours. It could achieve the screening of sleep apnea hypopnea syndrome in families and communities and provide a new method for the disease's early detection and diagnosis.

Keywords: LPC2132, $\mu\text{C}/\text{OS-II}$, Physiological Signal, Micro SD Card.

1 Introduction

People spend one-third of their time in sleep every day. Sleep is an indispensable component of the healthy as a necessary process of life and it could eliminate fatigue, restore energy, enhance immunity and delay aging. Some 27 % people around the world have poor sleep according to the World Health Organization survey. Many symptoms could reflect the poor sleep, such as snore, nocturnal awakenings, sleep architecture disorders, daytime sleepiness, accidents and other dangerous occurrences, easily lead to cardiovascular complications and other organ damage. These problems seriously affect people's quality of life and lives [1].

Poor quality of sleep is mainly caused by Sleep Apnea Hypopnea Syndrome (SAHS). According to the pathology, the disease can be divided into three types which are called obstructive sleep apnea, central sleep apnea and mixed sleep apnea [2].

Currently, the international medical community attaches great importance to the study of SAHS. Many countries have established sleep monitoring centers and mainly take use of community medical monitoring and treatment system. Since the mean of monitor is complicated and expensive, there are limitations of such diseases' treatment globally and many patients couldn't get treatment timely. The multi-parameter and portable monitor for sleep apnea status which has described in this paper is suitable for family and community clinic and it is useful for the early detection and diagnosis of sleep breathing disorders.

2 Parts of the Monitor

2.1 General Design

The multi-parameter and portable monitor for sleep apnea status has been designed for household and community clinic. The general design of the monitor is shown in Figure 1. The design includes sleep respiration signals' acquisition and conditioning, data record, LCD display, power management of lithium battery, etc.

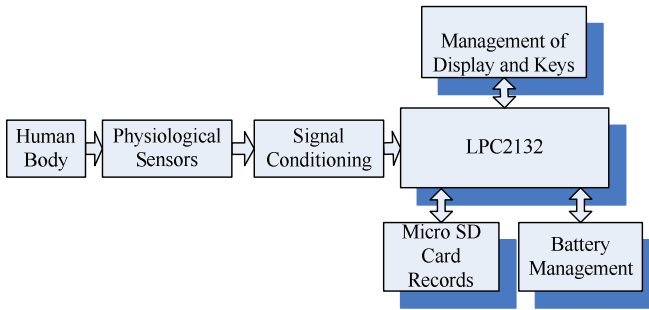


Fig. 1. The general design of the monitor

The LPC2132 microcontroller has been selected and it is based on a 32-bit ARM7TDMI-S CPU with real-time emulation and embedded trace support as the monitor's MCU, the microcontroller is ideal for small system applications due to its tiny size and low power consumption. Since the 10-bit 8-channel ADC meets the multi-channel analog data acquisition need and the serial SPI communication meets the requirement of the Micro SD card storage, the MCU applies to the system[3].

μC / OS-II is highly portable, ROMable, scalable, preemptive real-time, multitasking kernel (RTOS). It has got the safety certification of the U.S. Aeronautics and Space Administration and it can be used in aircraft, spacecraft and other control systems which ask for high security[4]. During the design of multi-parameter and portable monitor for sleep apnea status, the μC / OS-II has been used as the software design platform in order to improve efficiency of the work and improve the monitor's stability. The sleep apnea monitor could record multiple physiological parameters of human body during they are sleeping and it provides the early detection and diagnosis of sleep apnea with effective basis.

2.2 Acquisition and Conditioning of Physiological Signals

The monitor acquires multiple physiological signals, such as Nasal Flow (NAF), Thoracic Respiration, Abdominal Respiration, SaO_2 (%), Pulse Rate, Body Position, Actigraph, and Snore. The combination of this several physiological signals could reflect the patient's sleeping condition completely, including sleeping pose, whether respiration is normal, whether turning actions take place, whether apnea and snore occurs, whether SaO_2 (%) values is normal and so on. Normally, adult's respiratory

rate during sleep is about 12 ~ 16bpm (beats / min). According to Nyquist theorem, when f_s .max (sampling frequency) $\geq 2f_{max}$ (the highest signal frequency) during the analog or digital signals' conversion process, information in the original signal is kept unchanged in the digital signal after sampling [5]. Thus, 20HZ (1200bpm) has been selected as the sampling frequency which satisfies the accuracy of the data processing.

Body's respiratory status is mainly reflected by thoracic movement, abdominal movement, NAF and SaO_2 (%). In addition, pulse rate, body position, actigraph, snore and other auxiliary signals are monitored for more accurate observation. Thoracic and abdominal signals' monitoring uses piezoelectric sensor with chest belts fixed; NAF signal's monitoring uses thermal sensor and will be fixed under the nasal cavity when used; SaO_2 (%) and Pulse Rate signals will be acquired from a measurement module; Switching position sensor provides the body's supine, prone, left lying and right lying signals during sleep; Acceleration-type sensors and microphone sensors provide patient's actigraph and snore signal.

According to the type of sensors' output, sensors will be divided into analog sensor and digital sensor. The monitor uses the microcontroller's A/D converter to acquire analog signals and uses microcontroller's I / O port to get digital signal. The range of analog sensors output is -5 ~ +5 V, since the A/D converter's reference voltage is 3V, the microcontroller's ports could withstand 3V when used as the A / D converter. In this case, the analog sensors' output must be conditioned and filtered moderately to meet the requirement, as shown in Figure 2.

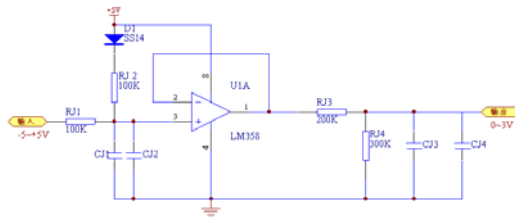


Fig. 2. The Circuit of Analog Signal Conditioning

2.3 Analysis of Sleep Events

Pathological sleep apnea hypopnea syndrome is defined that apnea or hypopnea occurred more than 30 times every night or more than 5 times per hour, the duration of each apnea for 10 seconds or longer. Hypopnea refers to expiratory volume less than 50 % of the normal and accompanied by SaO_2 (%) decreased more than 4% [6].

Normally, NAF, thoracic respiration and abdominal respiration are the same phase, their waveforms are smooth and cycles are basically same. At the same time, pulse rate and SaO_2 (%) are within the normal range, body movements, snore and other events are rarely occurred. Sleep apnea event will occur when respiration has been restricted within a long time, and its type can be distinguished by whether breathing effort exists [7, 8]. The feature of apnea types is shown in Figure 3. Obstructive Apnea

is mainly caused by the obstruction of upper respiratory airway. If obstructive apnea happened, thoracic and abdominal movement won't stop when nasal flow stops, in addition, thoracic and abdominal movement come from the same phase into contradictory movement. Central Apnea is mainly due to the lesions of the respiratory center to the muscle which makes the power of respiratory movement poor. If central apnea happened, thoracic and abdominal movement will stop when nasal flow stops and thoracic and abdominal movement will recover when nasal flow recovers. Mixed Apnea has the pathological mechanism of the both types above. Breathing movements will be stopped completely at the beginning of the mixed apnea, thoracic and abdominal movement will recover first at the end of the mixed apnea when nasal flow has not recovered [9].

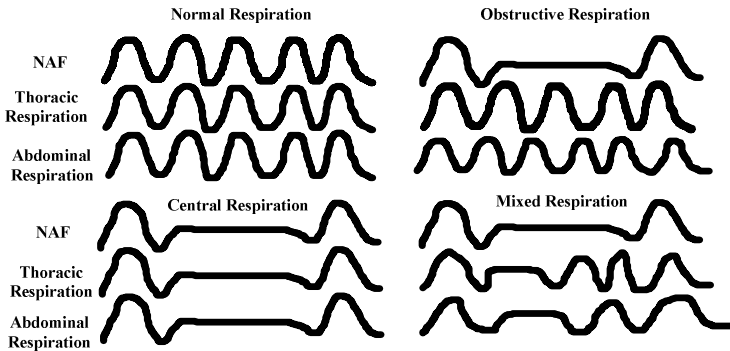


Fig. 3. Types of Apnea

The monitor provides information that the medical analysis needed. According to the diagnostic criteria of sleep respiratory events, the moment, duration and frequency of apnea events will be extracted from the nasal flow signal, and the type of apnea events will be distinguished through the thoracic and abdominal signals' correlation. At the same time, other auxiliary signals could help for recognizing associated characteristics of apnea events.

2.4 Data Record

One 2G Micro SD card has been used as the monitor's storage media. The monitor's data record follows the SPI protocol and has been embedded FAT32 file system in the software design. Micro SD card is the standard large-capacity memory card with a dedicated serial interface and meets the requirement of the data transfer. The SD card is compact (1.55cm * 1.2cm) and meets the monitor's portable design requirement. Data storage follows the SPI protocol, Micro SD card's SPI mode connection is shown in Figure 4. The SPI bus transmits data by bytes through data in and data out channels, it makes the MCU and peripheral devices communicated with serial communication.

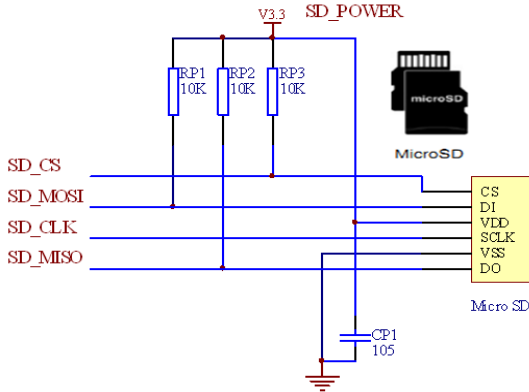


Fig. 4. Micro SD Card’s SPI Mode Connection

As the FAT32 file system has been embedded, data could be stored as a file in the card. In this case, the file could be read, downloaded, moved and deleted on any computer with windows operating system. Several parameters are stored synchronously and they are grouped by time and stored by column. The data will be downloaded from the Micro SD card to a PC to finish post-processing and analysis.

2.5 Power Management

One 3.7V(1000mA) lithium battery has been used to provide the whole system’s power. The lithium battery’s charging circuit is shown in figure 5. It is using the charge control chip LS2815 and a professional constant voltage and constant current charging mode to finish the battery’s charge management. There is protection for the battery’s temperature and the charge current is adjustable.

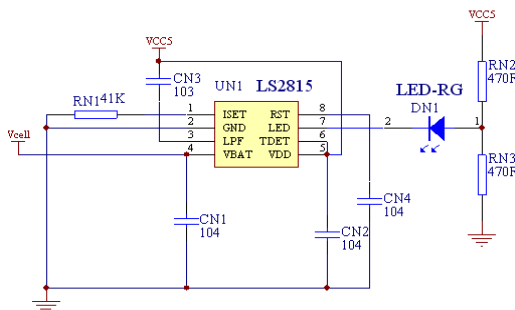


Fig. 5. Charging Circuit

Since the power supply of physiological sensors are 5V and the microcontroller requires 3.3V power supply, MAX1524 has been used to boost the voltage from 3.7V to 5V and precision resistor’s divider method has been adopted to get 3.3V. It

concludes from many experiments that the stable power of the monitor is less than 270mW. In this case, the monitor could record continuously at least 12 ($T=(1000\text{mAh}\cdot 3.7\text{V}) / 270\text{mW}=13.7\text{h}$) hours and could be used for a whole night.

3 Summary

The design of multi-parameter and portable monitor for sleep apnea status meets the market demand. The μC / OS-II has been embedded during the design of the monitor, and the whole system has been proved reliably through many experiments. Because of its flexible shape and low cost, the monitor is ideal for family and community clinic using. Signal wires are easy to wear and they almost do not affect the quality of patient's sleep during night. The monitor records patients' body movement, posture, respiration status, SaO₂ (%) and other signals all night to provide doctors with reliable information. Doctors will finish the sleep apnea syndrome screening and make diagnosis based on these data.

Acknowledgments. This work is supported by Shenyang Science and Technology Program (No.F11-009-2-00).

References

- [1] Young, T., Peppard, P.E.: Epidemiology of obstructive sleep apnea. In: McNicholas, W.T., Phillipson, E.A. (eds.) *Breathing disorders in sleep*, pp. 31–44. W.B.Saynders, London (2002)
- [2] Respiratory diseases group of Chin medical institute respiratory diseases sleep. Guidelines of Obstructive sleep apnea syndrome diagnosis and treatment. *Chin. J. Intern. Med.* 42(8) (2003) (in chinese)
- [3] PHILIPS. LPC2131/2132/2138 User Manual (2004)
- [4] Labrosse, J.J.: *The Grid: MicroC/OS-II the Real-Time Kernel*. Beibei Shao translated, Beijing (2003) (in Chinese)
- [5] Gockler, H.G., Groth, A.: *The Grid: Multi rate System and Sample rate conversion and digital filter*. DeHai Wang translated, Beijing (2009) (in Chinese)
- [6] He, Q., Chen, B.: *The Grid: Sleep Disordered Breathing*, Beijing, pp. 89–92 (2009) (in chinese)
- [7] American Academy of Sleep Medicine Task Force. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. *Sleep* 22, 667–689 (1999)
- [8] Clark, S.A., Wilson, C.R., Satoh, M., Pegelow, D., Dempsey, J.: Assessment of inspiratory flow limitation invasively and noninvasively during sleep. *Am. J. Respir. Crit. Care Med.* 158, 713–722 (1998)
- [9] He, Q., Chen, B.: *The Grid: Sleep Disordered Breathing*, Beijing, pp. 309–335 (2009) (in chinese)

The Design of Body Sensor Network Based on CC430

XiaoMei Liu, ShanShan Tang, ShaoChun Chen, and Ying Lian

College of Information Engineering, Shenyang University of Chemical Technology,
Shenyang, Liaoning, China

{liuxiaomei40,tss7114331}@126.com, csc2@21cn.com,
lianying3314@163.com

Abstract. With the development of microelectronics technology and wireless communication technology, the body sensor network has been widely applied to varied fields. Health monitoring is one of the most interested fields. BSN (Body Sensor Network) based on CC430 is proposed for human health monitoring, which can accurately detect human physiological parameters. It has good expansibility, low power consumption and flexibility, and provides more space for health monitoring research and development.

Keywords: BSN, CC430, physiological parameters.

1 Introduction

Human physiological parameters contain rich human health status information to the human body, so the accurate detection of physiological parameters can reduce the incidence of the disease, and achieve the effect of prevention and treatment in advance [1]. The traditional monitoring system needs the patient to be detected in hospital with multiple sensor wires which limits the patients' activity and comfort and thus negatively influences the measured result. Therefore it is the trend of technological development to develop the monitoring system of the human body parameters with wearable, tiny structure, light weight and low power consumption. It's also the subject that domestic and foreign scholars are being overcome.

This paper presents a body sensor network for human health monitoring. A number of tiny sensors, strategically placed on the human body, create a body sensor network that can monitor physiological parameters, providing real-time feedback to the user or the medical staff.

2 The BSN

With the development of biosensor technology, wearable technology and wireless sensor network technology, body sensor network gradually plays an increasing role in the medical field. It is composed of monitoring nodes, central processing nodes, and a receiving device which is mainly used in medical care, disease surveillance and biomedical fields [2].

Body sensor network is divided into brain region, chest region and arm region according to a body torso and nearby principle, contains multiple sensor nodes. There

are nasal flow sensor and snoring sensor in brain region. Chest region includes thoracic-abdominal respiratory waveform sensors, ECG sensor, etc. Oxygen saturation, temperature and blood pressure sensors are in arm region. These sensor nodes are managed according to their location.

2.1 The BSN Protocol

TI's Simplicii TI protocol is adopted, which is a simple RF protocol for the low frequency network. The amount of code is small, low cost [3].

The Simplicii TI protocol packages the network into several types of API functions. The point to point communication can be realized through directly calling the API function. Simplicii TI can run at low cost hardware resources. In addition to flash required by the program and RAM used by random variable, Simplicii TI does not require any other resource; it doesn't even need a timer [4].

2.2 The BSN Structure

A peer-to-peer topology is adopted, as shown in Figure 1. Following the principle of proximity, the same microcontroller CC430 is used to collect thoracic and abdominal respiratory waveform in the chest region. The chest node is used as the center node to receive the data of another two nodes. Then the data would be sent to PC from the center node by wireless transmission. The software for medical analysis is installed in PC, which can record and analyze the collected data. So users can determine a patient's health status easily and accurately, and then put forward effective treatment.

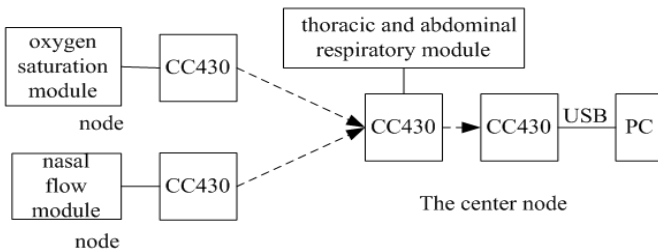


Fig. 1. The structure diagram of BSN

2.3 The Specific Application of the BSN

According to BSN's characteristics and design requirements, three regions are used—the brain region, the hand region and the chest region. Brain region has a nasal airflow sensor. There is an oxygen saturation sensor in hand region. Chest region contains thoracic and abdominal respiratory waveform sensors. Oxygen saturation sensor is placed in the middle finger.

Thoracic-abdominal respiratory waveform sensors are put in belts which are tied to people's chest and abdomen. Nasal airflow sensor is placed under the person's nose.

The receiving device is made into a small rectangular box tied to the body's waist. Specific wear is shown in Figure 2. Doctors may add or remove sensor nodes to reconfigure the network according to the need of medical doctors.

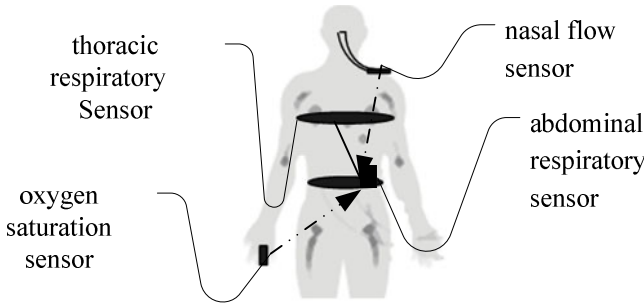


Fig. 2. Schematic diagram of sensors wear

3 The Node Design of the BSN

3.1 The Node Structure

Node is responsible for collecting physiological parameters and transmitting collected data to the central node through the wireless way. A node includes sensors, signal conditioning circuits, a processor module, and wireless communication module^[5], as shown in Figure 3. Physiological parameters acquisition module is made up of sensors and conditioning circuits.

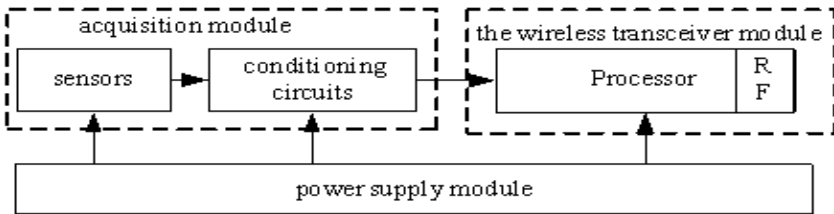


Fig. 3. The node structure diagram

Sensors is to collect physiological parameters, and conditioning circuits adjust signal to meet the AD converter requirements. Processor module is responsible for coordinating the work of the various nodes, processing information and preservation etc. Wireless communication between the central node and others node is realized by wireless communication module. The power module provides necessary energy for normal operation.

3.2 The Hardware Design of the Node

This network composed of tiny portable nodes equipped with many sensors. For the hardware design of nodes, the following points must be considered. Firstly, nodes are small enough to be worn comfortably for a long time. Secondly, their energy consumption should also be optimized so that the battery is not required to be changed regularly. Thus, tiny, light weights, low power consumption of chips have been selected.

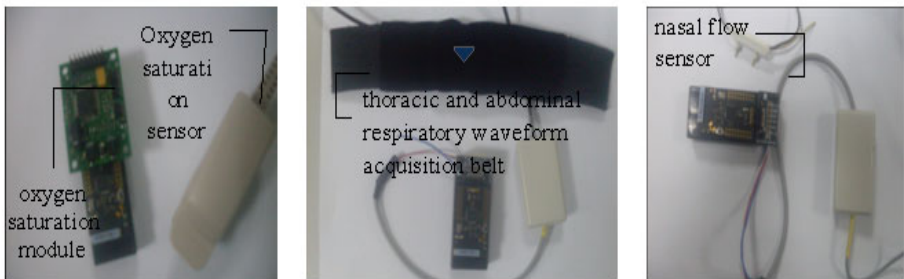
3.2.1 The Processor

The processor is the heart of the system. The system is based on CC430 single chip microcomputer with powerful processing function, high processing speed and low power consumption. In the clock operating conditions of 1MHz, the current will be 200-400uA. When the clock is shutdown mode, the current is only 0.1uA [6].

CC430 chip has built-in wireless module CC1101. This design is set in the 433MHz frequency band. RF transceiver is integrated with a highly configurable modem. This modem supports different modulation formats, the data transfer rate can up to 500kbps, and software may modify the baud rate. It can be widely used in various short distance wireless communication field [6].

3.2.2 The Physiological Parameters Acquisition Module

The physiological parameters acquisition module is an important part of the design which consists of three sub-modules, module of oxygen saturation, thoracic-abdominal respiratory waveform, and nasal flow, as shown in Figure 4.



(a) Oxygen saturation module (b) thoracic and abdominal module (c) nasal flow module

Fig. 4. The diagram of the physiological parameters acquisition module

A high sensitive finger clip sensor is used in oxygen saturation module. Its internal photodiode converts the light signal to the electrical signal. The signal is a mixed quantity, both red signal and infrared signal, which must be sent to conditioning circuits, shown in Figure 5. The signal is amplified 11 times by AD620, then successfully separated red light and infrared light through the positive and negative

switch circuit and the CD4053 switch. The two signals are sent to low-pass filter circuit and high-pass filter circuits to get DC components and AC components, components are sent into AD conversion, and then calculate the oxygen saturation value.

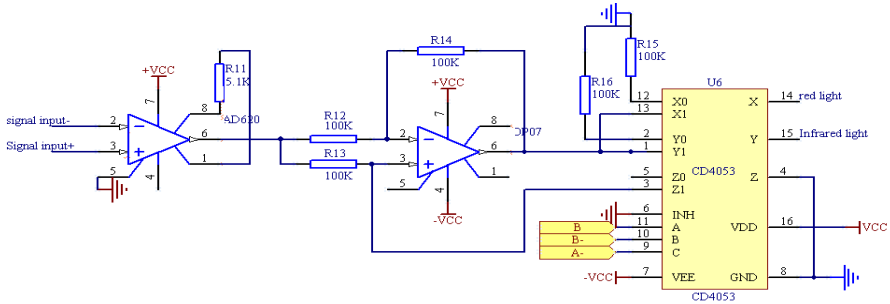


Fig. 5. The circuits of oxygen saturation module

Thoracic abdominal respiratory waveform sensors are piezoelectric sensors, sensitivity > 500mV/300g (grams). Nasal airflow sensor is a thermal three-wave sensor with high sensitivity, light weight, easy to use which is placed below nostril.

Waveform amplitude of thoracic-abdominal respiratory waveform sensors and nasal flow waveform sensor is different according to different people which are amplified through amplifiers to the range between -5V~+5V, and then adjusted by conditioning circuits to meet AD conversion’s requirements.

3.3 The Software Design of the Node

Software design includes node and the center node design, as shown in Figure 6.

Nodes follow energy conservation mode-sleep, wake up and work. After powered up, each node is in a dormant state, the processor stops working in a low current receiving state. When receiving a command sent by the central node, the node is awakened. The processor determines according to nodes’ number, if the command is the current node, the node enters the working state, begin to collect physiological parameters. If the command is not the current node, then back to sleep.

The central node completes initialization of hardware and interface and then sleeps to wait for receiving messages. When the center node received a message to identify the sensor node according to the number of the node, then to store. The central node goes back to sleep after completion.

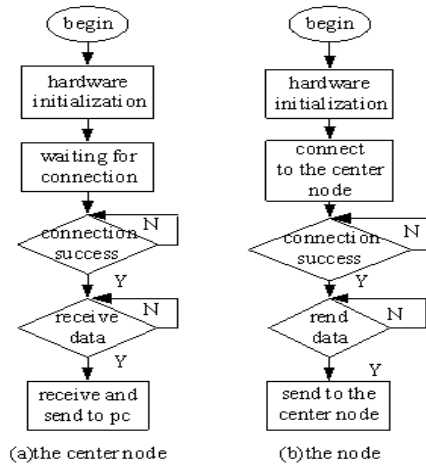


Fig. 6. Flow chart of the body sensor network software

4 Summary

This paper describes a body sensor network based on CC430. It can monitor human physiological parameters timely through sensors which are placed on body. And it has many advantages such as: low power consumption, low interference, high flexibility, scalability, etc. Not only can it be widely used in hospitals, clinics, communities, but also can enter people's homes. In future work, the system's software and hardware design will further be developed and improved. This design meets modern society demands for high-quality health trend and has great research value.

Acknowledgments. This work is supported by the science and technology project of Shenyang City (No. F11-009-2-00).

References

1. Shnayder, V., Chen, B.-R., Lorincz, K., Fulford-Jones, T.F., Welsh, M.: Technical Report TR-08-05, Division of Engineering and Applied Sciences, Harvard University (2005)
2. Otto, C., Sanders, C., Jovanov, E.: *The Grid: Journal of Multimedia*, New York (2006)
3. TI company. *SimpliciTI Sample Application User's Guide* (2009)
4. TI company. *Application Note on SimpliciTI Frequency Agility* (2007)
5. Krishnamachari, B.: *The Grid: Networking Wireless Sensors*. Cambridge University Press, New York (2005)
6. TI company. *CC430f5135 User Manual* (2010)

Historical and Status Quo Analysis on the Production and Consumption of China's Electric Power Energy

Jie Yao, Jin Zhu, and Wei Gao

School of Economy and Management, Northeast Dianli University, 132012, Jilin, China
{Leeyooab, zhujin09}@163.com, 13944224885@126.com

Abstract. This article makes deep research work on the history and status quo of the production and consumption of China's electric power energy by analyzing the data in detail. It studies the change tendency of the total production of China's electric power energy from 1990, and studies the ratio of the amount generated by different sources of electric power to the total and their changing tendency. It also makes detailed statistical analysis on the industrial distributions of China's electric power energy and on the effects that electric power consumption of different regions has had on regional output. Through the analysis above, the article can help the readers understand the development situation of China's electric power energy and its major roles in promoting regional economic growth profoundly.

Keywords: electric power energy, energy production, energy consumption.

1 Introduction

Energy is one of the material conditions for the people to survive, and it is also one of the most important resources of economic development and social progress. After modern industrialization and social division of labor, energy begins to influence all facets of human production and human life widely and profoundly. As an important form of energy, the production and consumption of electric power energy has become the focus and hot issue of academic research in recent years.

China is one of the countries with the fastest economic growth, at the same time it is also the country who produces and consumes large capacity of electric power energy. Therefore, China's great achievements in economic development, high speed production and consumption of electric power energy have become the focus of the world.

2 Analysis on the Change Tendency of China's Energy Production and Consumption

Since China was founded, especially since the reform and opening up, China has made remarkable achievements in energy development. China has formed the new pattern of energy supply in which coal supply is the main body, electric power supply

is the core, and at the same time companied with the overall development of oil, gas and renewable energy. Since 1978, China's energy production and consumption has shown a growth trend year by year. Till 2008, China has become the largest energy producer in the world, with the total primary energy production of 2.6 billion tons of standard coal. By 2009, China's energy production and consumption has reached 2.75 billion tons and 3 billion tons of standard coal respectively.

From the constitution of energy production and consumption, we can find that coal is still in the dominant position of China's energy production and consumption. The amount of coal production and consumption accounts for nearly 70% of that year's energy production and consumption. The amount of crude oil accounts for about 20%. The other components of energy production and consumption are renewable energy sources, such as natural gas, hydropower, nuclear power and wind power. Their proportion is relatively small. The proportion of natural gas is generally 2%-3%. Hydropower, nuclear power and wind power have developed very fast in recent years. In 2009, the proportion of their production and consumption is 8.7% and 7.8% respectively.

From the ratio of individual industry's energy consumption to the total, the amount of the secondary industry has accounted for the vast majority of energy consumption, with the ratio of more than 71.5%. And it has also shown a growth trend. The second is the third industry. With the continuous improvement of people's living standard, the ratio of the third industry in total energy consumption is also showing an upward trend year by year. The energy consumption of the first industry hasn't developed as fast as that of the second and third industry, and its ratio in the total has also declined. The ratio of first industry energy consumption has dropped while the ratio of second and the third industry has risen. To some extent, it can reflect the changes in China's industrial structure.

3 Historical and Status Quo Analysis on the Production and Consumption of China's Electric Power Energy

3.1 Analysis on the Production of China's Electric Power Energy

A country's energy structure of electric power refers to the proportion of the capacity generated by all kinds of power in the total capacity. Whether the structure is reasonable or not will directly affect the sustainability of this country's power supply. Until now, thermal power is still in the leading position in most countries' power energy structure.

Based on the analysis on the changes of China's electricity production from 1990 to 2008, this article finds that China's capacity of electricity production has been increasing year by year since 1990. And according to the power energy structure, the capacity generated by thermal power, hydropower and nuclear power, has increased individually in various degrees every year, just as shown in Figure 1 below.

The changes of China's total electric power production are coincided with China's rapid economic development in recent years. The increasing supply of electric energy year by year provides a solid foundation for economic development. As shown in Table 1, thermal power has been in the dominant position in China's electric power production with the percentage of up to 80%. Hydropower is in the second place of China's power energy structure, but its proportion has been showing a declining trend, with the percentage of 16.94% in 2008. Nuclear power takes the third place in the power energy structure, the average ratio is about 2%. Other new energies are still so small in scale in China's power energy structure.

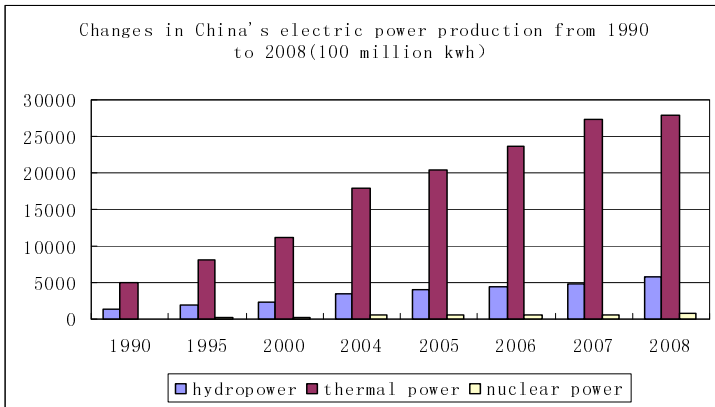


Fig. 1. Changes in China's Electric Power Production from 1990 to 2008

Table 1. The Ratio of the Amount that Hydropower, Thermal Power and Nuclear Power Produce Annually to that Year's Total Amount from 1990 to 2008

	1990	1995	2000	2004	2005	2006	2007	2008
Hydropower	20.34%	19.01%	16.51%	16.09%	15.92%	15.24%	14.83%	16.94%
Thermal power	79.37%	80.24%	82.70%	81.72%	82.09%	82.89%	83.24%	80.78%
Nuclear power	0.00%	1.28%	1.24%	2.30%	2.13%	1.92%	1.90%	1.98%

Overall, China's power energy structure depends on the thermal power excessively. Since coal belongs to the non-renewable resources and its utilization rate is low, this kind of energy structure is deemed to be unsustainable. With the development and implementation of China's environmental laws and regulations, and with China's accession into a number of international environmental conventions, environmental pollution caused by thermal power plants will get more and more strict restriction. In order to ensure the sustainability of power energy production, we must optimize the power energy structure. The cost of nuclear power, coal, oil and gas, electric power published by American Nuclear Society (NEI) in 2004 is 1.68, 1.90, 5.39 and 5.87 cents / KWh respectively. From the international point of view, the production cost of nuclear power has the most competitive advantage. The environmental cost of

conventional coal-fired power plants is 10.92 RMB cents / KWh, that of nuclear power is 0.32 RMB cents / KWh. It is clear that the highest environmental cost is caused by coal, natural gas takes second place, the environmental cost caused by nuclear power is the lowest. According to "The Regulations on Collection and Use of Sewage Charges" and "Management Approaches On The Standard Of Collecting Sewage Charges", 0.6 RMB yuan will be charged for each equivalent pollution, which means that the external costs of power generation are internalized into the accounting costs of thermal power plants, and thus further weakens the competitiveness of thermal power [1].

3.2 Analysis on the Consumption of China's Electric Power Energy

Analysis on the Industrial Distribution of China's Power Energy Consumption.

By analyzing the data in detail, this article gets the industrial distribution and overall tendency of China's power consumption from 1990 to 2008, as shown in Figure 2. we can find that the secondary industry, including industry and construction, accounts for a relatively large proportion in the power consumption, and the annual electric power consumption of industry is far beyond that of other industries, accounting for more than 70% of total power consumption. And it has also shown a gradually increasing trend.

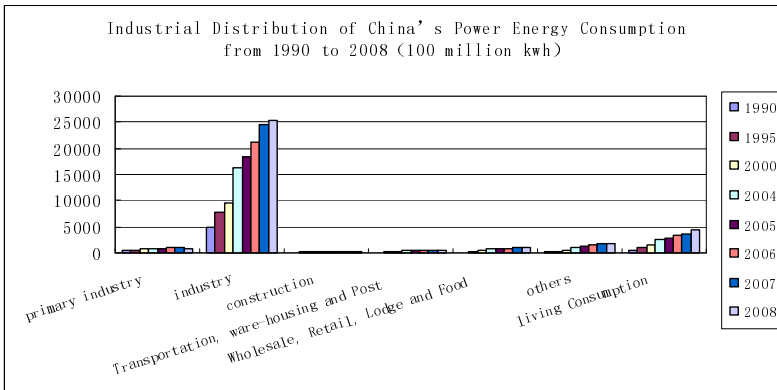


Fig. 2. Industrial Distribution of China's Power Energy Consumption from 1990 to 2008

There are close relationships among electric power consumption, industrial scale and economic output. Gradually increasing consumption of electric power also reflects the great development of China's industry. The large amount of electric power consumption is determined by the nature of the industry. Industry is closely related with many national economic lifelines. It provides basic life materials for residents, and provides raw materials for different industries. It belongs to the high energy-consuming industries. At the same time, from the perspective of industrial structure, the industries closely related with the life of residents are all sectors with high power energy consumption, such as manufacturing. From 2006 to 2008, the

power energy consumption of manufacturing sector has accounted for nearly 73% of the total amount in industry, and that of electricity sector, gas and water supply sector has accounted for about 20% of the total. The percentage of the extractive industry is relatively much smaller than the others.

It is not only the industry that has consumed large amount of electricity, but also the tertiary industry that has consumed a lot of electricity. This part of electricity has been used on people's daily consumption to meet people's daily needs. The amount of electric power consumed on daily life has also been showing an upward trend year by year, from 7.72% of total electricity consumption in 1990 to 12.73% of the total in 2008, as shown in Table 2. This fully demonstrates that Chinese living standards are continuously improved and people's consumption capacities have been significantly improved along with the improvement of living standards. Another industry that has consumed more electricity is forestry, animal husbandry, fishery and water conservancy, which belongs to the primary industry. The amount of electric power consumption of the primary industry is also rising, but the ratio of it to the total amount of that year's electric power consumption has been decreasing.

Table 2. The Ratio of the Amount of Electric Power Consumption in Different Industries to that Year's Total Amount from 1990 to 2008

	1990	1995	2000	2004	2005	2006	2007	2008
Agriculture, Forestry, Animal husbandry, Fisheries and Water conservancy	6.85%	5.81%	5.00%	3.68%	3.51%	3.31%	2.99%	2.57%
Industry	78.22%	76.42%	71.66%	73.98%	74.10%	74.32%	75.30%	73.50%
Construction	1.04%	1.59%	1.15%	1.01%	0.94%	0.95%	0.94%	1.06%
Transportation, ware-housing and Post Industry	1.70%	1.82%	2.09%	2.05%	1.73%	1.63%	1.63%	1.66%
Wholesale, Retail, Lodge and Food business	1.22%	1.99%	2.92%	3.35%	3.02%	2.96%	2.84%	2.95%
Others	3.25%	2.34%	4.77%	4.72%	5.38%	5.44%	5.22%	5.54%
Living Consumption	7.72%	10.03%	12.41%	11.22%	11.33%	11.37%	11.07%	12.73%

Analysis on the Regional distributions of China's Electric Power consumption in 2009. By analyzing the data in "China Statistical Yearbook of 2010 ", we can draw the distribution curves of regional electric power consumption and GDP, as shown in Figure 3. It can be seen that the amount of the following six areas' electric power consumption is among the best in China in 2009, they are Guangdong, Jiangsu, Shandong, Zhejiang, Hebei and Henan province etc. The ratio of the amount of these areas' electric power consumption to that year's total is summed to be up to 45.8%, that of Guangdong is 9.86%, that of Jiangsu is 9.06%, that of Shandong is 8.04%, that of Zhejiang is 6.75%, that of Hebei is 6.4%, and that of Henan is 5.69%. At the same time, we can notice these areas' regional GDP is also among the best in China.

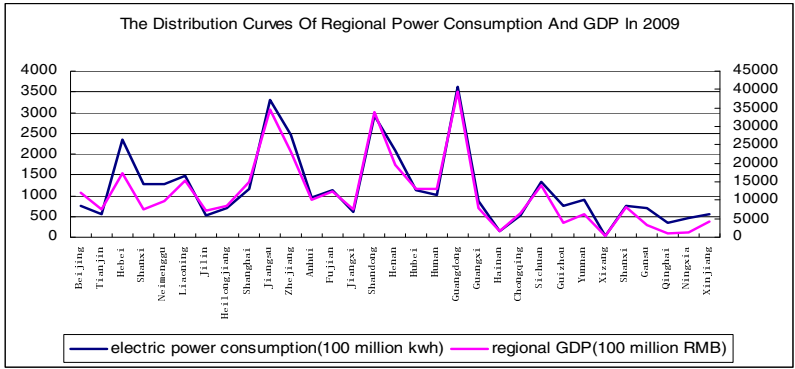


Fig. 3. The Distribution Curves of Regional Power Consumption and GDP in 2009

In addition, we can also find that the curve shape of China's regional electric power consumption is almost the same as that of regional GDP in 2009 from Figure 3. By analysis, the correlation coefficient between these two factors is calculated to be as high as 0.96. Thus, the amount of regional electric power consumption is closely related to regional GDP. The amount of the region's electric power consumption reflects its economic development level to some extent, and the region's economic development level also has a direct effect on the demand of electric power energy.

4 Conclusion

Through the analysis above, we can see that the supply and demand of electric power energy not only is the major macro-economic issue that affects the stability and growth of China's economy, but also has a significant impact on the future sustainable economic development. The formulation of reasonable short-term and long-term power developing policies is closely related to people's livelihood. When formulating the specific policies for developing electric power, we must pay attention to improve energy efficiency, increase the conversion of primary energy to secondary energy, improve the existing power structure, make an overall plan for the operation and construction of hydropower and thermal power, and improve the proportion of electric power consumption in the structure of total energy consumption and in the structure of production, so as to ensure the sustainability of electric power production.

References

1. Yang, Y., Sun, T.: Policy Analysis on Sustainable Energy Development of Electric Power Production. *J. Scientific Management Research* 6, 52–55 (2008)
2. Yin, J.-H., Wang, Z.-H.: The Relationship between Energy Consumption and Economic Growth in China—Based on the Data in the Period of 1953–2008. *J. Science Research Management* 7, 122–129 (2011)

A High Performance Audio Watermarking Algorithm Based on Substitution in Wavelet Domain

Huan Hao¹, Liang Chen¹, Yipeng Zhang¹, Yinguang Zhang¹, and Bo Zhang²

¹ Department of Electronic Information Engineering ICE, PLAUST, Nanjing, China, 210007

² Department of Mechanical and Electrical Engineering, XMU, Xiamen, China, 361005

Abstract. An audio watermarking scheme in wavelet domain which takes advantage of substitution is proposed in this paper. First the host audio is decomposed by wavelet transform, and then convert low frequency coefficients to segments. Finally embed the watermark XOR with a pseudo-random sequence into the host audio by substitution of the segments. The experimental results illustrate that the suggested scheme has achieved high imperceptibility and is robust against common audio signal processing such as additive noise, low-pass filtering, median filtering, re-quantifying and re-sampling.

Keywords: audio watermarking, wavelet transform, pseudo-random sequence, substitution.

1 Introduction

With the broad use of multi-media products these years, a very crucial issue for copyright protection has emerged. Digital watermarking is considered as an effective way to resolve the copyright protection problems of digital products.

Audio watermarking methods exploit the insensitivity of the human auditory system (HAS) in various techniques such as embedding algorithms based on low-bit coding, echo, rational dither modulation, patchwork, Fourier transform, wavelet transform or spread spectrum and interpolation. Interpolation techniques are often designed to provide a good perceptual quality from known sample values. [1] In Reference [2], substitution-by-interpolation is used in time domain. Reference [3] brings substitution-by-interpolation in wavelet domain for audio watermark and achieves better performance, comparing with Reference [2]. A fatal weakness for [3] is that it's almost impossible to perform a 5-level wavelet transform for a long audio because the storage and computation required are huge. From our experimental results, [3]'s algorithm, however, seems not adaptive to Gaussian white noise. Moreover, the low-frequency coefficients of the 5-th level are insensitive for our ears. So the substitution-by-interpolation is meaningless. These motivate us to propose a more adaptive and high-efficient watermarking technique to reduce the false recovery for watermarked audio against common attacks.

Our watermarking techniques are mainly based on substitution to further improve the transparency and the performance of audio watermarking technique with

robustness against common attacks. Due to the subsection for the host audio, it's high-efficient to realize wavelet transform. Comparing with [3], our scheme can embed more watermarks into the same audio by substitution with fore-and-aft segments. Furthermore, our scheme has achieved high accuracy, imperceptibility and robustness by using substitution in wavelet domain.

2 Review of Basic Concept for Wavelet

Wavelet transform is a powerful tool for analysis and processing of non-stationary signal. It's expanded with wavelet basis formed by localized function, which has many special features and benefits. It has been widely used in science, engineering, mathematics and computer science. Most notably, it is used for signal processing, to represent a discrete signal in a more redundant form.

The basic idea of the wavelet transform is to represent any arbitrary function f as a superposition of wavelets. Any such superposition decomposes f into different scale levels, where each level then further decomposed with a resolution adapted to the level. The Discrete Wavelet Transform (DWT) can be showed:

$$W_k[f(x)] = \frac{1}{2^k} \int_{-\infty}^{+\infty} f(t)\psi * (\frac{x-t}{2^k})dt \tag{1}$$

Where $\psi(t)$ is the mother wavelet.

An original audio signal S will be decomposed into two parts: the approximate part and the detail part. Depending on the application and the length of the signal, the low frequencies part might be further decomposed into two parts of high and low frequencies. Fig.1 shows a 3-level DWT decomposition of signal S .

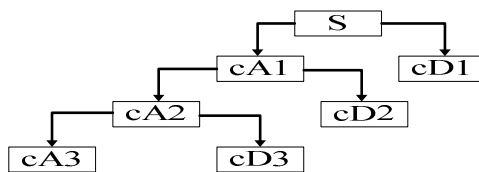


Fig. 1. 3-level DWT decomposition of signal S

Due to its excellent spatio-frequency localization properties, the DWT is very suitable to embed watermark. The frequency of audio is 300 Hz -3400 Hz and frequency lower than 300 Hz is imperceptible for our ears. So we can embed watermark into this region by taking advantage of this characteristic.

3 Our Watermarking Scheme

3.1 Watermark Embedding

The watermark embedding is described in the following steps:

Step1: divide the host audio S into frames and the length of each frame is N . In this way we can reduce storage and computation. Moreover, N can be the key to extract watermark for the receiver.

Step2: perform DWT for every frame to embed watermark into the low-frequency coefficients and the wavelet basis adopts “db4”. We get $A_k^L, D_k^L, D_k^{L-1}, \dots, D_k^1$, where A_k^L is the L^{th} level approximate part and $D_k^L, D_k^{L-1}, \dots, D_k^1$ are the detail parts of 1^{st} to L^{th} level.

Step3: the size of watermark is $M1 \times M2$. In order to increase the robustness to various attacks in the channel, we should eliminate the relation in the watermark. [4] XOR the watermark W with a pseudo-random sequence, which is generated by a “seed”. In this way, we get W' and enhance the security of watermark at the same time. Then convert the two-dimensional watermark W' into one-dimensional sequence W'' . The watermark before and after XOR with a pseudo-random sequence are shown in Fig.2.

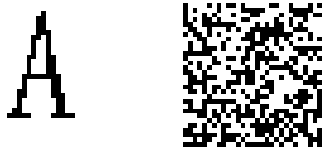


Fig. 2. Original watermark and XOR with pseudo-random sequence watermark

Step4: further divide A_k^L into non-overlapping sets, the length of each set is l . $\varphi_1, \varphi_2, \varphi_3 \dots$ are unchanged and $\psi_1, \psi_2 \dots$ would be substituted by $\varphi_1, \varphi_2, \varphi_3 \dots$. Specific rules are as follows:

$$\begin{aligned} \psi_j' &= \begin{cases} \varphi_j, b_i = 1 \\ \varphi_{j+1}, b_i = 0 \end{cases} \\ \varphi_j' &= \varphi_j \\ \varphi_{j+1}' &= \varphi_{j+1} \end{aligned} \tag{2}$$

Considering the AWGN (zero-mean and variance σ^2) attacks in the channel, the fatal recovery for watermark will increase. In embedding scheme, we adopt an adaptive method to improve it. The embedding algorithm is summarized as below:

```

for i=1:N
  for j=first set to last one
    if max(E_{\varphi_j}, E_{\varphi_{j+1}}) < \alpha_1 * mean(k)  \varphi_{j+1} = (1 + \alpha_2) * \varphi_j ; end

    Err = (\varphi_j - \varphi_{j+1}) * (\varphi_j - \varphi_{j+1})^T ;

    if Err < \alpha_3 * mean(k)  \varphi_{j+1} = (1 + \alpha_4) * \varphi_j ; end

    if b_i = 1  \psi_j' = \varphi_j ; else  \psi_j' = \varphi_{j+1} ; end
  end
end
end

```

Where E_{φ_j} and $E_{\varphi_{j+1}}$ are the energy of φ_j and φ_{j+1} , $mean(k)$ is the mean energy of the k^{th} frame, Err is the square error of φ_j and φ_{j+1} , b_i is the i^{th} bit of secret bit stream in a frame, $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are thresholds, here we adopt 0.4, 1.5, 2, 0.3. ψ_j' is the marked value of ψ_j . The main parameters of this algorithm are the length of frame N , the length of segment l and thresholds $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, which define the maximum allowed ratio error caused by attacks in the channel. Fig.3 illustrates the embedding algorithm.

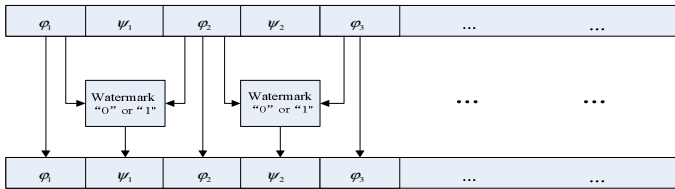


Fig. 3. Embedding algorithm

Step5: perform Inverse Discrete Wavelet Transform (IDWT) on modified $A_k^L, D_k^L, D_k^{L-1}, \dots, D_k^1$. This operation produces audio S' embedded with watermark.

3.2 Watermark Extraction

The watermark extraction can be performed without the original audio. The details are described in the following steps:

Step1: divide the watermarked audio S' into frames and the length of each frame is N .

Step2: perform DWT for every frame and we get $A_k^L, D_k^L, D_k^{L-1}, \dots, D_k^1$.

Step3: further divide A_k^L into non-overlapping sets $\varphi_1', \psi_1', \varphi_2', \psi_2', \varphi_3', \dots, \varphi_l'$ and calculate the square error $Err1$ and $Err2$. We can extract watermark W_1 . The algorithm is summarized as follows:


```

for i=1:N
  for j=first set to last one
    Err1 = ( $\varphi_j'$  -  $\psi_j'$ ) * ( $\varphi_j'$  -  $\psi_j'$ )'; Err2 = ( $\varphi_{j+1}'$  -  $\psi_j'$ ) * ( $\varphi_{j+1}'$  -  $\psi_j'$ )';
    if Err1 < Err2  $b_j = 1$ ; else  $b_j = 0$ ; end
  end
end
end

```

Fig.4 illustrates the extraction procedure.

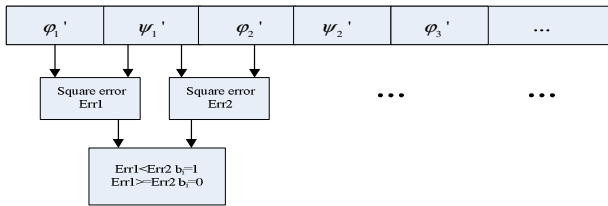


Fig. 4. Extraction procedure

Step4: convert the one-dimensional watermark W_1 into two-dimensional sequence W_1' , and then XOR it with the pseudo-random sequence generated by “seed”. Thus we get extracted watermark W_1'' .

4 Experimental Results

An audio sampled at 8kps with the length of 36s in WAVE format was chosen as test signal. Watermark W is a binary image with size 30×30 . The length of frame N is 9600. Variance of White Gaussian noise σ^2 is 0.01.

4.1 Imperceptibility Test

Fig.5 shows the comparison between original audio and watermarked audio, from which we can see that our scheme has little distortion in waveform.

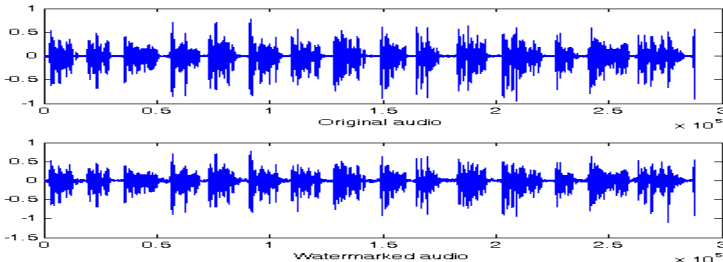


Fig. 5. Waveform comparison before and after embedding watermark into an audio

We evaluate the distortion by mean opinion score (MOS), which is a subjective measurement, and achieves transparency between imperceptible and perceptible but not annoying, MOS=4.4. We test the signal noise ratio (SNR) at the same time, which is an objective measurement. It is defined as below:

$$SNR = 10 \log \left\{ \frac{\sum_{n=1}^l x(n)^2}{\sum_{n=1}^l ((x(n) - x'(n))^2)} \right\} \tag{3}$$

From formula (3), we get the SNR before and after embedding watermark is 20.315. Thus our scheme has a high transparency and imperceptibility not only in subjective measurement, but also in objective one.

4.2 Robustness Test

The detection performance depends on bit accuracy rate (BAR), which is defined as:





$$BAR = \frac{\text{Number of Bits Correctly Detected}}{\text{Number of Bits Embedded}} \tag{4}$$

In order to test the performance, the watermarked audio suffers from different kinds of attacks including additive noise, re-sampling, re-quantifying, median filtering, low-pass filtering. The simulation results, including BAR and extracted watermark are shown in Table 1. Seen from Table 1, our scheme is robust against different types of common attacks.

Table 1. Resistance to different types of attacks

Attack	Parameter	State	BAR	Extracted watermark
None attack	-	-	1	
Addictive noise	White Gaussian noise(mean=0, $\sigma^2=0.1$)	best	1	
Addictive noise	White Gaussian noise(mean=0, $\sigma^2=0.1$)	worst	0.9989	
Resample	8k->16k->8k	-	1	
Re-quantify	16bit->8bit->16bit	-	1	
Median filter	Length of window is 3	-	1	
Low-pass filter	<=1.5k	-	0.9967	

Table 1. (continued)

Low-pass filter	$\leq 2k$	-	0.9967	
Low-pass filter	$\leq 2.5k$	-	0.9978	
Low-pass filter	$\leq 3k$	-	0.9989	
Low-pass filter	$\leq 3.4k$	-	1	

5 Conclusion

A high performance audio watermarking scheme based on substitution in wavelet domain is proposed in this paper. The technique takes advantage of the HAS and time-frequency characteristics of wavelet transform. Simulation results show that the proposed method is blind and has good performance in both subjective and objective measurement. Furthermore, with 3.5dB loss in SNR than [3], the proposed method is more robust to various attacks. Besides, it is more efficient (To embed the same watermark W , the length of host audio required is only 66.7%, storage needed is 2.3% and computation is about 30%).

Acknowledgments. This work is supported by National Natural Science Foundation of China (Grant No. 61072042).

References

1. Fallahpour, M., Megias, D.: High capacity audio watermarking using FFT amplitude interpolation. *IEICE Electronics Express* 6(14), 1057–1063 (2009)
2. Deshpande, A., Prabhu, K.M.M.: A substitution-by-interpolation algorithm for watermarking audio. *Signal Processing*, 218–225 (2009)
3. Bai, Y., Bai, S., Bao, J.: Audio Watermarking Algorithm Based on Substitution by Interpolation in Wavelet Domain. *Journal of Chongqing University of Technology* (2011)
4. Chen, X., Li, X., Hu, G., Shen, J.: Adaptive digital audio watermarking method based on DCT and lifting wavelet transform. *Journal of Computer Applications* 31(2) (2011)

Generalization Model of State Observers Based on Delay-Quantization Method

Leiming Liu and Chaonan Tong

School of Automation & Electrical Engineering,
University of Science and Technology Beijing, Beijing 100083, China
Leiming.liu@163.com

Abstract. For the state observers control problem over networks, delay-quantization method is applied and a discrete-time Markov jump linear system (DMJLS) model is established. When the parameters expressing delay take different values, the model with long (or short) Markov delay and independent delay are obtained respectively. And an approach to get the initial stabilizing controllers is proposed. Simulation on a cart and inverted pendulum show that designed controllers with state observers can make the system stable.

Keywords: state observers control, delay-quantization method, Markov delay.

1 Introduction

NCS (Networked Control Systems) is a networked real-time closed-loop feedback control system. The communication network inevitably raises new challenging problems such as packet losses, and delays. They will present determinate characters or stochastic, constant[1] or time varying, long[2~4] or short, and independent[4,5] or governed by Markov chains[3,5]. In this paper, delay-quantization method is applied in NCS with state observers to establish a DMJLS model. Simulation on a cart and inverted pendulum shows that the model and the algorithm have practicality.

2 The Structure and Model of NCS

The general structure of NCS is shown in Fig.1. Let delay from sensor to controller be $\tau_{sc}(t)$, and delay from controller to actuator be $\tau_{ca}(t)$. Hereby, let the plant model be

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t - \tau_{sc}(t)) \end{cases}, \mathbf{x}(t) \in \mathbf{R}^n, \mathbf{u}(t) \in \mathbf{R}^r, \mathbf{y}(t) \in \mathbf{R}^m \quad (1)$$

and $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are matrices with appropriate dimensions. Let the controller model be

$$\begin{cases} \dot{\mathbf{v}}(t) = \mathbf{D}\mathbf{v}(t) - \mathbf{L}\mathbf{y}(t), \quad \mathbf{w}(t) = \mathbf{K}\mathbf{v}(t) \\ \mathbf{u}(t) = \mathbf{w}(t - \tau_{ca}(t)) \end{cases}, \mathbf{v}(t) \in \mathbf{R}^q, \mathbf{w}(t) \in \mathbf{R}^r, \mathbf{D} = \mathbf{A} + \mathbf{L}\mathbf{C} + \mathbf{B}\mathbf{K} \quad (2)$$

If $\mathbf{A} + \mathbf{L}^* \mathbf{C}$ is stable for \mathbf{L} , we should select controller \mathbf{K} when there are delays in the NCS to satisfy some demands of system performance. For model (1) and (2), we

should give the equivalence discrete form. Therefore we will first make the following assumptions and driving fashion[5]:

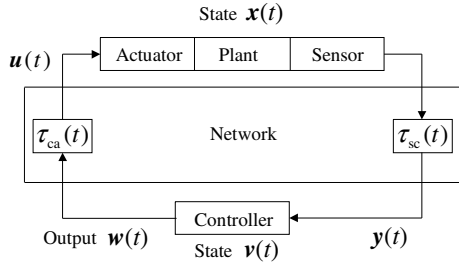


Fig. 1. General structure of NCS with delays

1. Different nodes of the control system are clock synchronization, the data packets have timestamps.
2. The sensor node is clock-driven, and have constant sampling period h . The sampling instant is $t_k = kh, k = 0, 1, 2, \dots$. The controller nodes and actuator nodes are all event-driven.
3. Buffers are used at the controller node and at the actuator node, receiving the latest data packets from the sensors and controllers respectively.
4. The controller always uses the latest data to calculate the control signal.
5. The actuator always uses the latest control signal, until a new control signal arrives; if the control signal $u(t_m)$ arrives after $u(t_n)$, and $m < n$, $u(t_m)$ should be ignored by the actuator. For the output y and w , the same thing is done.
6. Stochastic delays are governed by finite state time-homogeneous Markov process. This paper will use symbols as follows:

- a) Function $\delta : \delta(0) = 1$ and $\delta(i) = 0$, when $i \neq 0$, and function $\bar{\delta} : \bar{\delta}(i) = 1 - \delta(i)$.
- b) The mark I denotes a identity matrix; 0 denotes null matrix or zero; $x_k \equiv x(t_k)$.
- c) Let $\bar{C}_i = [I\delta(i) \ I\delta(i-1) \ \dots \ I\delta(i-d)]$.

In this section, first we present a delay-quantization discrete-time model of continuous-time systems (1), (2). From assumption 6, let

$$\tau_{sc}(t) < s \cdot h, \tau_{ca}(t) < a \cdot h, h_1 = h / \sigma, h_2 = h / \alpha. \tag{3}$$

Where $s, a, \sigma > 1$ and $\alpha > 1$ are all positive integer. From (3), for any sampling instant t_k , there exist nonnegative integers $s_k, \sigma_k, a_k, \alpha_k$, satisfying $\sigma_k h_1 \leq \tau_k^s - s_k h < (\sigma_k + 1)h_1, \alpha_k h_2 \leq \tau_k^a - a_k h < (\alpha_k + 1)h_2$. Where $\tau_k^s = \tau_{sc}(t_k), \tau_k^a = \tau_{ca}(t_k), s_k \in \mathcal{V}_s = \{0, 1, \dots, s-1\}, \sigma_k \in \mathcal{V}_\sigma = \{0, 1, \dots, \sigma-1\}, a_k \in \mathcal{V}_a = \{0, 1, \dots, a-1\}, \alpha_k \in \mathcal{V}_\alpha = \{0, 1, \dots, \alpha-1\}$. Let transition probability matrices (TPM) of delays s_k, σ_k, a_k and $\alpha_k, k = 0, 1, 2, \dots$ be stochastic matrices respectively:

$$P^s = (p_{ij}^s), P^\sigma = (p_{ij}^\sigma), P^a = (p_{ij}^a), P^\alpha = (p_{ij}^\alpha). \tag{4}$$

Where $p_{ij}^\mu = \text{Prob}\{u_{k+1} = j | u_k = i\}$, $i, j = 0, 1, \dots, u-1$ and $u = s, \sigma, a, \alpha$ respectively.

According to the above assumption and description, we can get formula (5) and (6) as follows

$$\mathbf{x}_{k+1} = A_1 \mathbf{x}_k + B_0^{\alpha_k} \mathbf{w}_{k-a_k} + B_1^{\alpha_k} \mathbf{w}_{k-a_k-1}, A_1 = e^{Ah}, B_0^{\alpha_k} = \int_0^{h-\alpha_k h_2} e^{A\rho} B d\rho, B_1^{\alpha_k} = \int_{h-\alpha_k h_2}^h e^{A\rho} B d\rho \tag{5}$$

$$\mathbf{v}_{k+1} = D_1 \mathbf{v}_k + E_0^{\sigma_k} C \mathbf{x}_{k-s_k} + E_1^{\sigma_k} C \mathbf{x}_{k-s_k-1}. \tag{6}$$

Where $D_1 = e^{Dh}$, $E_0^{\sigma_k} = -\int_0^{h-\sigma_k h_1} e^{D\rho} L d\rho$ and $E_1^{\sigma_k} = -\int_{h-\sigma_k h_1}^h e^{D\rho} L d\rho$.

Let the augmented vector be respectively

$$\tilde{\mathbf{x}}_k = [\mathbf{x}_k^T \ \mathbf{x}_{k-1}^T \ \dots \ \mathbf{x}_{k-s+\delta(\sigma-1)}^T]^T, \tilde{\mathbf{v}}_k = [\mathbf{v}_k^T \ \mathbf{w}_{k-1}^T \ \dots \ \mathbf{w}_{k-a+\delta(\alpha-1)}^T]^T, \hat{\mathbf{x}}_k = [\tilde{\mathbf{x}}_k^T \ \tilde{\mathbf{v}}_k^T]^T \tag{7}$$

We present Theorem 1 as follows:

Theorem 1. System (5)~(7) is equivalent to the following closed-loop system (8),

$$\hat{\mathbf{x}}_{k+1} = A_{s_k, \alpha_k}^{\sigma_k} (K) \hat{\mathbf{x}}_k, A_{s_k, \alpha_k}^{\sigma_k} (K) = \begin{bmatrix} \tilde{A} & \tilde{B}^{\alpha_k} \tilde{F}_{a_k} \\ \tilde{E}^{\sigma_k} \tilde{C}_{s_k} & \tilde{D} \end{bmatrix} \tag{8}$$

Where

$$\tilde{A} = \begin{bmatrix} A_1 & 0 & \dots & 0 & 0 \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix}, \tilde{B}^{\alpha_k} = \begin{bmatrix} B_0^{\alpha_k} & B_1^{\alpha_k} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \tilde{E}^{\sigma_k} = \begin{bmatrix} E_0^{\sigma_k} & E_1^{\sigma_k} \\ 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \tilde{D} = \begin{bmatrix} D_1 & 0 & 0 \\ K & & \\ 0 & I & 0 \end{bmatrix}$$

$$\tilde{C}_{s_k}^{\sigma_k} = \begin{bmatrix} C \bar{C}_{s_k} \\ C \bar{C}_{s_k + \bar{\delta}(\sigma_k)} \end{bmatrix}, \tilde{F}_{a_k} = \begin{bmatrix} K \delta(a_k) & I \delta(a_k - 1) & \dots & I \delta(a_k - a) \\ K \delta(a_k + 1) & I \delta(a_k) & \dots & I \delta(a_k - a + 1) \end{bmatrix}.$$

Proof. This theorem can be proved through directly calculation. □

Remark 1. We can express the matrix of (8) as (9) where parameters of controllers are separated from parameters of the plant.

$$A_{s_k, \alpha_k}^{\sigma_k} (K) = \hat{A} + \hat{B}^{\alpha_k} \cdot \hat{K}_{a_k}^{\sigma_k} \cdot \hat{C}_{s_k}^{\sigma_k}. \tag{9}$$

Where $\hat{A} = \begin{bmatrix} \tilde{A} & 0 \\ 0 & 0 \end{bmatrix}, \hat{B}^{\alpha_k} = \begin{bmatrix} 0 & \tilde{B}^{\alpha_k} \\ I & 0 \end{bmatrix}, \hat{K}_{a_k}^{\sigma_k} = \begin{bmatrix} \tilde{D} & \tilde{E}^{\sigma_k} \\ \tilde{F}_{a_k} & \tilde{G}_{a_k} \end{bmatrix}$ and $\hat{C}_{s_k}^{\sigma_k} = \begin{bmatrix} 0 & I \\ \tilde{C}_{s_k}^{\sigma_k} & 0 \end{bmatrix}$.

In the four matrices of (9), the controllers parameters K are only contained in $\hat{K}_{a_k}^{\sigma_k}$, This is the simplicity of the model. □

Remark 2. According to assumption 6 and formula (4), formula (8) is a DMJLS model, the delay variable $s_k, \sigma_k, a_k, \alpha_k$ are jump parameters of the system. The state set of $(s_k, \sigma_k, a_k, \alpha_k)$ is Cartesian product $\theta = \vartheta_s \times \vartheta_\sigma \times \vartheta_a \times \vartheta_\alpha$ of $\vartheta_s, \vartheta_\sigma, \vartheta_a$ and ϑ_α , and the delay state TPM is Kronecker product $P = P^s \otimes P^\sigma \otimes P^a \otimes P^\alpha$ of P^s, P^σ, P^a and P^α . Therefore, (8) can be expressed as $\hat{x}(t_{k+1}) = A_{s_k, \alpha_k}^{\sigma_k, a_k}(K)\hat{x}(t_k)$, $(s_k, \sigma_k, a_k, \alpha_k) \in \theta$. We record the number of elements of the set θ as $|\theta|$, and let $|\theta| = d + 1$. Furthermore, many-subscript set can become single subscript set, so (8) can also be expressed as

$$x_{k+1} = A_{\zeta_k} x_k, \zeta_k \in \theta = (0, 1, \dots, d), k = 0, 1, 2, \dots, x_k = \hat{x}(t_k) \tag{10}$$

Remark 3. If the probability distribution of delays is $\pi = [\eta_0 \ \eta_1 \ \dots \ \eta_d]$, when we let TPM $P^T = [\pi^T \ \pi^T \ \dots \ \pi^T]$, model (8) can be used to express state observers feedback control issue[4,5]. □

3 Stability Analysis of DMJLS

We have presented a DMJLS model (8) or (10) of the NCS with quantized delays . For every time index k , there are $d+1$ matrices, which are A_0, A_1, \dots, A_d , this is a DMJLS with $d+1$ states. Assume the initial states ζ_0 and x_0 of ζ_k and x_k are known, suppose probability distribution vector of delay state ζ_k at the instant k be $\pi(k) = [\pi_0(k), \dots, \pi_d(k)]$, $k = 0, 1, 2, \dots$, and $\pi(0) = \pi_0$ is probability distribution for ζ_0 , so $\pi(k+1) = \pi(k)P$, $k = 0, 1, 2, \dots$. Now we shall discuss the stability problem of the system (10). In [7], a necessary and sufficient condition for stochastic stability of DMJLS was presented. In [6], a necessary and sufficient condition for mean square stability (MSS) of DMJLS was presented, and at the same time, some other equivalent necessary and sufficient conditions were also presented, such as the condition of the spectral radius of an augmented matrix being less than one. This paper refers to two theorems of [6]. Based on these, we propose Theorem 4 which get a feasible initial solution of the stability problem of the system (10).

Definition 1. We say that system (10) has MSS, if for any initial condition $x_0, \zeta_0 \in \theta$ satisfying

$$\lim_{k \rightarrow \infty} E\{\|x_k(x_0, \zeta_0)\| | x_0, \zeta_0\} = 0. \tag{11}$$

Theorem 2[6]. The MSS of system (10) is equivalent to the following: There exist some symmetric positive definite matrices Q_0, \dots, Q_d , satisfying

$$Q_i - \sum_{j=0}^d p_{ij} A_j^T Q_j A_j > 0, i = 0, 1, \dots, d. \tag{12}$$

Theorem 3[6]. Let spectral radius of matrix A be $\rho(A)$. Define $A = XY$, where $X = P^T \otimes I_{n^2}$, $Y = \text{diag}[\hat{A}_1 \hat{A}_1 \cdots \hat{A}_d]$, $\hat{A}_i = A_i \otimes A_i, i = 0, 1, \dots, d$. Then MSS of (10) is equivalent to $\rho(A) < 1$.

Now we shall propose Theorem 4, which will be used to solve issue (10).

Theorem 4. If some A_j of $A_j (j = 0, 1, \dots, d)$ satisfying $\rho(A_j) < 1$, and every row of P is $(\eta_0, \eta_1, \dots, \eta_d)$, then there exists nonnegative number $\eta_0, \eta_1, \dots, \eta_d, \sum_{i=0}^d \eta_i = 1$, making $\rho(A) < 1$ of Theorem 3. (omit proof) □

4 Design and Algorithm Guaranteeing Stability of the System

According to Theorem 2, we should find $Q_j > 0, A_j, j = 0, 1, \dots, d$, satisfying (12). For issues (1) and (2), when $\tau_{sc}(t) \equiv 0$ and $\tau_{ca}(t) \equiv 0$, we can establish an augmented closed loop system. Then we can get K in (2) guaranteeing system mentioned above be asymptotic stable by eigenstructure assignment method^[8], thus part of initial values for A_0, \dots, A_d , in(12) is obtained. Finally, based on these, we should find Q_0, \dots, Q_d and make them, together with A_0, \dots, A_d , be feasible initial solutions of (12). According to Theorem 4, let special stochastic matrices P_0 be initial values of P , then we can get initial feasible solutions of (12). Hereinafter we shall solve for parameters K to make the system (10) be mean square stable corresponding to P_E , which are the practically measured values of P .

5 Examples

Consider the NCS problem of the cart and inverted pendulum shown in Fig.2, the mass of the cart is $M = 2\text{kg}$, the mass of the pendulum is $m = 0.1\text{kg}$, the length of the pendulum is $l = 0.5\text{m}$, the position of the cart is x (m), and the inclined angle of the pendulum is θ (rad). Let $x_1 = \theta, x_2 = \dot{\theta}, x_3 = x$ and $x_4 = \dot{x}$, set $y = [x_1, x_3]^T$. Let $q = 4, s = 2, \sigma = 2, a = 1, \alpha = 3$ then $d = 11$. Set the sampling period $h = 0.05\text{s}$. By the method in [8], we can find the initial values of K . Hereby we can get initial values of matrices (9). Let TPM be

$$ps = \begin{bmatrix} 0.80 & 0.20 \\ 0.79 & 0.21 \end{bmatrix}, pm = \begin{bmatrix} 0.7075 & 0.2925 \\ 0.6925 & 0.3075 \end{bmatrix}, pa = [1], pn = \begin{bmatrix} 0.72 & 0.2800 & 0 \\ 0.70 & 0.2800 & 0.0200 \\ 0.70 & 0.2825 & 0.0175 \end{bmatrix}.$$

Hereby, we can get the controller parameters as follows:

$K_0 = [43.2754 \ 35.5115 \ 0.0486 \ 0.6009]$, $K_1 = [49.3787 \ 31.2823 \ 0.0549 \ 0.6897]$, which make the system be mean square stable. When the initial state

$x_0 = [0.1 \ 0 \ 0 \ 0]^T$ and the initial probability distribution vector π_0 : $\pi_s = [0.65 \ 0.35]$, $\pi_m = [0.75 \ 0.25]$, $\pi_a = [1]$, $\pi_n = [0.55 \ 0.3 \ 0.15]$, we can obtain the initial condition response curves of the closed-loop system shown in Fig.3.

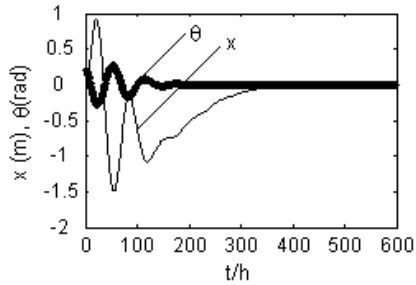
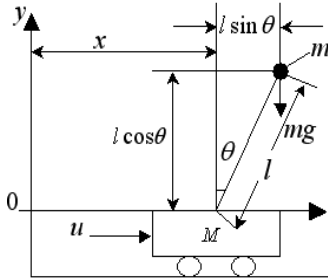


Fig. 2. Cart and inverted pendulum

Fig. 3. Initial condition response of NCS with delays

6 Conclusions

For the NCS with delay-quantization method, a DMJLS model on state observers is presented. For the reason of applying delay-quantization method, the model presented is fit for solving the problem of NCS with independent delay and Markov delay. The simulation shows applicability of the model and validity of the algorithm.

References

1. Luck, R., Ray, A.: An Observer-Based Compensator for Distributed Delays. *Automatica* 26(5), 903–908 (2003)
2. Zhang, W.: Stability Analysis of Networked Control Systems, PhD Dissertation, Department of Electrical Engineering and Computer Science, Case Western Reserve University, USA (2001)
3. Xiao, L., Hassibi, A., How, J.P.: Control with Random Communication Delays via a Discrete Time Jump Systems Approach. In: Proceedings of the 2000 American Control Conference, pp. 2199–2204. Chicago, USA (2000)
4. Hu, S.S., Zhu, Q.X.: Stochastic Optimal Control and Analysis of Stability of Networked Control Systems with Long Delay. *Automatica* 39(11), 1877–1884 (2003)
5. Nilsson, J.: Real-Time Control Systems with Delays. PhD Dissertation, Department of Automatic Control, Lund Institute of Technology, Sweden (1998)
6. Costa, O.L.V., Fragoso, M.D.: Stability Results for Discrete-Time Linear Systems with Markovian Jumping Parameters. *Journal of Mathematical Analysis and Applications* 179(1), 154–178 (1993)
7. Shu, Z., Lam, J., Xiong, J.: Static Output-Feedback Stabilization of Discrete-Time Markovian Jump Linear Systems: A System Augmentation Approach. *Automatica* 46, 687–694 (2010)
8. Duan, G.R.: Theory of Linear Systems. Harbin Institute of Technology Press, Harbin (1998)

Design of Doherty Amplifier in Digital TV Transmitter*

Chen Jun and Su Kaixiong

Institute of Physics and Information Engineering of Fuzhou University,
Fuzhou 350002 China

Abstract. OFDM modulation is used in the digital TV transmitter. this signal has some characteristics such as wide bandwidth, multi-carrier, high PAPR, etc. The efficiency of transmitter power amplifier will be made in very high demand. In this paper people will pull Doherty amplifier in the front-end power amplifier. Introduced the principle and basic structure of Doherty amplifier. Through computer simulation people can verify the performance of this amplifier circuit in the digital television transmission system. Also we design the actual circuit and test and achieve real results. Through the results, we can find that the Doherty amplifier has solved of the low efficiency problem in digital TV front-end transmitter. So it has great relevance.

Keywords: digital television, Doherty, efficiency, PA.

1 Introduction

The process of development to the 4G, the biggest challenge is how to support a variety of existing mobile communication standards including GSM, GPRS, WCDMA and HSDPA. At the same time, it must support 100Mb / s ~ 1Gb / s data rate and support OFDMA modulation, support MIMO antenna technology, as well as support for the VoWLAN network. Therefore, during the RF signal chain design process, how to reduce power consumption and improve RF power amplifier efficiency microwave power amplifier design have become the key technology needed to solve. Currently there are three routes which can improve the amplifier efficiency: amplifier with CMOS technology, increases integration and improves PA efficiency indirectly; amplifier with InGaP technology, achieves low power consumption and high power amplifier efficiency; improving the efficiency of the amplifier by using SiGe BiCMOS process technology [1] [2]. From the three technology above, they are all from the microscopic point of view to achieve improved power amplifier efficiency. In practical applications, in addition to considering the angle from the amplifier core technology, we can consider increasing the amplifier efficiency from the amplifier circuit structure. The most classic is the Doherty amplifier. The Doherty amplifier can improve the average efficiency from 15% to 20%. Doherty amplifier is used more and more [3] [4].

* Fund: Fujian major projects (2010HZ0004-1) Fuzhou City, school science and technology cooperation (2011-G-105) Fuzhou University Technology Development Fund (2011-XY-23)

2 Doherty Power Amplifier

Doherty amplifier consists of a main amplifier and an auxiliary amplifier. The main power amplifier works in Class B or Class AB. The auxiliary amplifier works in Class C. Two amplifiers don't work in rotation. The main amplifier has been working all the time. The auxiliary amplifier doesn't work until it reaches the peak which has been set before (this amplifier is also called the peak amplifier). 90° quarter-wavelength line behind the main amplifier is an impedance transformation. It's function is to reduce the apparent impedance of the main amplifier when the auxiliary amplifier works. So the main amplifier's output current becomes large. Because of the main amplifier with a quarter wavelength lines, in order to make the two amplifier outputs in same phase, the auxiliary amplifier is also needed a 90° phase shift in front of it. Shown in Fig.1 [5] [6].

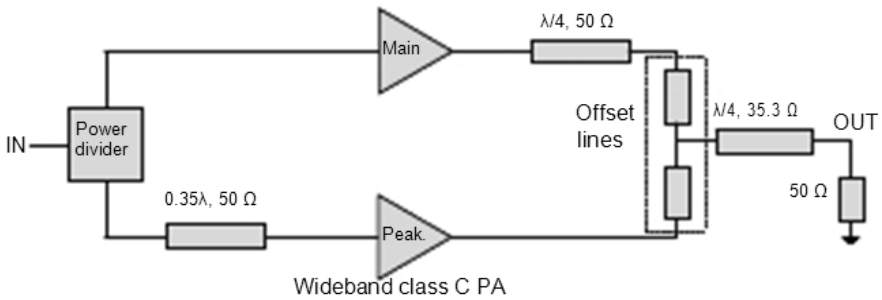


Fig. 1. Block diagram of the designed wideband Doherty PA

Main power amplifier works in Class B. When the input signal is relatively small, only the main amplifier is active. When the peak of output voltage of the tube reaches the saturation point, the theoretical efficiency can reach 78.5%. If we encourage in twice, the tube's output voltage reaches half of the peak and saturates. The maximum efficiency is 78.5%. The auxiliary amplifier begins to work together with the main amplifier (C class, the threshold is set to half the excitation voltage signal) at this time. Making the point of view from the main amplifier, the auxiliary amplifier reduces the load because the auxiliary amplifier acts as a negative serial impedance. So even if the main amplifier's output voltage is saturated and constant, the output power has continued to increase (the current flowing through the load becomes larger) because of the load The decrease. When it reach the peak of excitation, the auxiliary amplifier has reached it's maximum point of efficiency. So the two power amplifier's efficiency is much higher than the efficiency of a single class B amplifier. The single class B power amplifier's largest efficiency of 78.5% is in peak, but now the efficiency of 78.5% is in half of the peak. Therefore, this system architecture can achieve very high efficiency (each amplifier to achieve maximum output efficiency) [7].

3 Doherty Power Amplifier Simulation

In microwave circuit simulation, people use ADS software to simulate. There is the use of ADS simulation software to achieve full match doherty power amplifier simulation. Shown in Fig.2

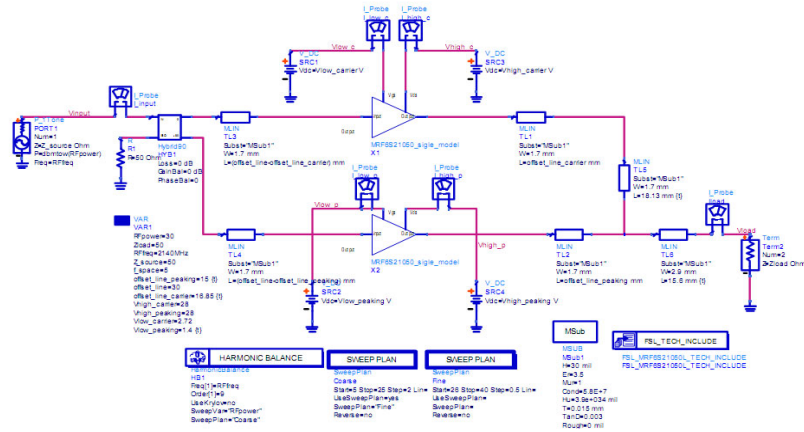


Fig. 2. Diagram of ADS software simulation

The following figure is the simulation results of doherty amplifier power gain and efficiency of transport.

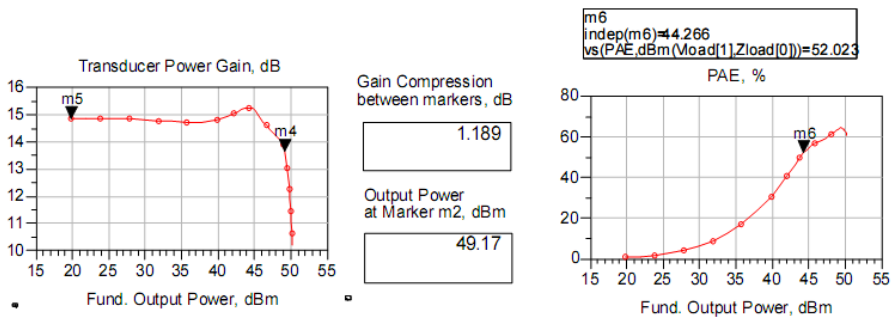


Fig. 3. The simulation results of Transmission power gain and efficiency

The transmission efficiency gain from the diagram can be seen that the amplifier power is 15dB, 1dB point is 49dbm, 3dB compression point is 50dbm. From the efficiency graph, it is clear that the power efficiency of 6dB back is at 52%. It is far higher than traditional Class AB amplifiers, which fully embodies the advantages of Doherty power amplifier.

4 Design of Doherty Power Amplifier’s Actual Circuit

According to the theoretical analysis, we design a practical application of doherty amplifier circuit. We Use Freescale’s power amplifier chip. Specific design of the circuit shown in Fig.4.

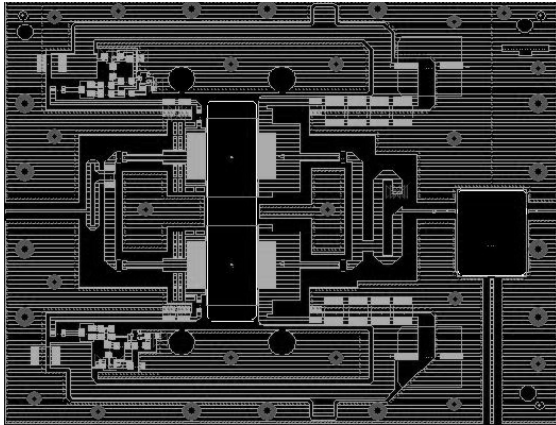


Fig. 4. Doherty power amplifier’s actual application

5 Analysis of Test Results

For the average RF board, the dielectric constant will work with the heat caused by the amplifier temperature changing. It causes that the loss of the entire circuit is not stable

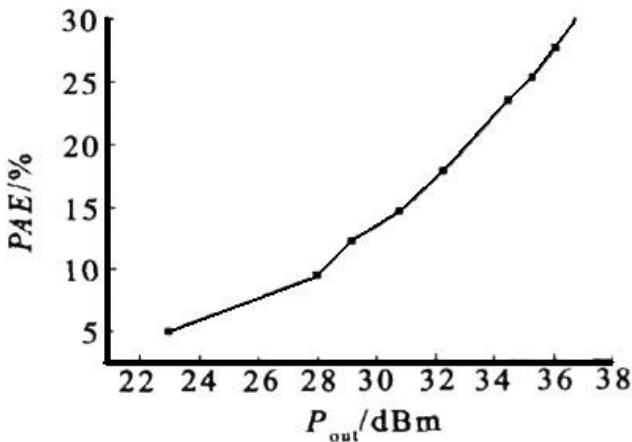


Fig. 5. Doherty power amplifier’s efficiency

and affects circuit 's performance. Using RF35 for PCB circuit board, the dielectric constant can be more stable when the amplifier's temperature changes. We use frequency spectrum analyzer and signal generator to test the power amplifier. The results obtained of the doherty power amplifier output power of 36 dBm in the case, the efficiency is 28%. There is a certain gap between the actual results and simulation results, but the result was more consistent.

6 Conclusion

Doherty power amplifier will be used in digital TV front-end transmitter system. It greatly improving the efficiency of front-end Transmitter systems. Through a combination of the corresponding linearization techniques, such as feed-forward technique, we can make digital television transmitter front-end efficiency and overall performance linearity to be fully guaranteed. Test results show that doherty technology meets the modern wireless communication systems of power amplifier requirements. It will have very broad application prospects.

References

1. Wang, Y., Fu, H., et al.: Nonlinear amplifier effects on the OFDM signal. Xidian University 29(4), 504–509 (2002)
2. Albullet, M.: RF Power Amplifiers, pp. 346–360. Noble Publishing Corporation. Atlanta (2001)
3. Gonzalez, G.: Microwave transistor amplifiers analysis and design (Microwave Transistor Amplifiers Analysis and Design). Tsinghua University Press, Bai Xiaodong (2003)
4. Freescale Semiconductor. Wireless RF Product, Freescale Semiconductor Device Data, DL110, Arizona, vol. 5, pp. 508–521 (2004)
5. Doherty, W.H.: A new high efficiency power amplifier for modulated waves. Proc. IRE 24(9), 1163–1182 (1936)
6. Bohge, M., Gross, J., Wolisz, A., et al.: Dynamic Resource Allocation in OFDM Systems: an Overview of Cross-layer Optimization Principles and Techniques. IEEE Network 21(1), 53–59 (2007)
7. Bumman, K., Jangneon, K., Ildu, K., et al.: The Doherty power amplifier. IEEE Microwave Magazine 7(5), 42–50 (2006)

Fusion Method of Feature Weight Based on Evidence Support Matrix

Guochu Chen¹, Rui Miao², and Yue Li²

¹ Electrical college, Shanghai Dianji University, Shanghai 200240, China

² School of Information Science and Engineering,
East China University of Science and Technology, Shanghai 200237, China
chgcsh@yahoo.com.cn, asdmiaorui@126.com

Abstract. This paper proposes a new fusion method that is based on weight of evidence support matrix eigenvector. The two factors that are conflict coefficient and evidence distance affecting evidence conflict are analyzed firstly. Evidence support matrix is constructed, and characteristic vector of the max eigenvalue can be calculated as the weight of evidence. Then the evidences with weight are fused using evidence combination formula. The method is proved reasonably and effectively according to experiment.

Keywords: evidence theory, evidence distance, synthesize conflict coefficient.

1 Introduction

The evidence theory was first proposed by Dempster who was a specialist of America in control fields in 1987, then further improvement and development was researched by his student Shafer, The evidence theory is widely applied in fault diagnosis[1,2], target recognition [3], prediction [4,5], data fusion[6], and so on. While evidence combined formula deals with high conflict evidence, a operated against natural result is gained. One is improved evidence combined formula, there are many kinds of improved methods [7-9]; another is improved evidence, Murphy introduces a way based on evidence-weighted average, which settles conflict of evidences [10].

Two factors are analysed in this paper: conflict factor and evidence distance[11], derived measure of the degree of different evidence, then gained improved evidence using conflict coefficient, last fused by evidence combined formula and compared with literature[7]and [12].

2 Basic Concepts of DS Theory

Assuming Θ is a identification framework , a mapping is defined on 2^Θ which is power set of Θ , $m : 2^\Theta \rightarrow [0,1]$, for $\forall A \subset \Theta$, s.t. $m(A) \geq 0$, $\sum_{A \subset \Theta} m(A) = 1$, It

reflects support level of evidence to identification framework's subset A under mathematical sense.

Assuming m_1, m_2 are different evidence defined on identification framework Θ , then combined formula as follows:

$$\begin{cases} m(A) = (1-k)^{-1} \sum_{\substack{B_i \cap C_j = A \\ A \neq \emptyset}} m_1(B_i)m_2(C_j) \\ m(\emptyset) = 0 \end{cases} \tag{1}$$

here $k = \sum_{B_i \cap C_j = \emptyset} m_1(B_i)m_2(C_j)$ is called conflict factor of evidence

3 Measurement of Evidence Confliction

In evidence combined formula, k is a physical quantity which measures evidence conflict. Here giving another example, assuming identification framework,

$$\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}, m_1(\theta_1) = m_1(\theta_2) = m_1(\theta_3) = m_1(\theta_4) = m_1(\theta_5) = 0.2$$

$m_2(\theta_1) = m_2(\theta_2) = m_2(\theta_3) = m_2(\theta_4) = m_2(\theta_5) = 0.2$, we can get $k = 0.8$, it reflects that high conflict is existed among the evidence, this is wrong obviously. From example. we can safely draw the conflict factor is one important diathesis that weights measurement of evidence conflict yet, not all big k stands for exist of confliction.

Conflict factor can't describe measurement of evidences, Here, we quote "distance function" [11], and give a new measure method.

Def. 1. (distance function) : assuming Θ is a complete identification framework including n different propositions, m_1 and m_2 are two different evidence vectors, distance function of m_1 and m_2 is :

$$d(\overline{m_1}, \overline{m_2}) = \sqrt{\frac{1}{2}(\overline{m_1} - \overline{m_2})^T D(\overline{m_1} - \overline{m_2})} \tag{2}$$

Distance function can also represent confliction of evidence, through define of distance function, we can acknowledge that distance function stands for diversity between evidences. Because of limitations of conflict factors, I propose a new measure method using conflict factor and distance function.

Def. 2. (Synthesize conflict coefficient): assuming d_{ij} is distance function of m_i and m_j , k_{ij} is conflict factor.

$$c'_{ij} = \sqrt{d_{ij}k_{ij}} \tag{3}$$

$$c''_{ij} = \frac{1}{2}(d_{ij} + k_{ij}) \tag{4}$$

$$c_{ij} = \lambda_{ij1}c'_{ij} + \lambda_{ij2}c''_{ij} \tag{5}$$

Called c_{ij} synthesize conflict coefficient of m_i and m_j .

This define not only settle the problem that the traditional conflict factor can't accurately describe confliction measurement, but also consider uncompatibility and diversity of focus element under identification framework. The define is reasonable via experiment.

4 New Method of Evidence Fused Based on Synthesize Conflict Coefficient

Assuming Θ is a complete identification framework, there are N independent evidences m_i , c_{ij} is synthesize conflict coefficient of m_i and m_j . So we can get cross correlation function sup_{ij} of m_i and m_j using $c_{ij} : \text{sup}_{ij} = 1 - c_{ij}$,

Then get $N \times N$ square matrix based on $\text{sup}_{ij} = 1 - c_{ij}$

$$S = \begin{pmatrix} \text{sup}_{11} & \cdots & \text{sup}_{1N} \\ \cdots & \cdots & \cdots \\ \text{sup}_{N1} & \cdots & \text{sup}_{NN} \end{pmatrix} \tag{6}$$

Then get evidence i support that all evidences sustain

$$\text{sup}_i = \sum_{j=1}^N \alpha_j \text{sup}_{ij} \tag{7}$$

Because S is a irreducible non-negative symmetric matrix, S has a largest eigenvalue λ_{\max} according as P-F theorem[13], and components of eigenvector x belonging to the largest eigenvalue λ_{\max} are positive numbers. So: $Sx = \lambda_{\max} x$

$$\text{Unfold it } x_1 \text{sup}_{i1} + \cdots + x_N \text{sup}_{iN} = \lambda_{\max} x_j \tag{8}$$

Comparing the right of formula (7) with the left of (8), we can making $\alpha_j = x_j$, get

$$\text{sup} = (\lambda_{\max} x_1, \lambda_{\max} x_2, \cdots, \lambda_{\max} x_N) \tag{9}$$

Normalization can get weight vector W of all evidences: $W = (\omega_1, \omega_2, \cdots, \omega_N)$. Here $\omega_i = (\sum_{j=1} x_j)^{-1} x_i$

5 Numerical Experiment

For the convenience of comparison, this paper uses experiment data of document [12],and compares my method with document [7] and [12].

Identification framework $\Theta = \{A, B, C\}$, all evidences as follows, fusion result sees table 1. $m_1(A) = 0.98, m_1(B) = 0.01, m_1(C) = 0.01, m_2(A) = 0,$

$$m_2(B) = 0.01, m_2(C) = 0.99, m_3(A) = 0.9, m_3(B) = 0, m_3(C) = 0.1,$$

$$m_4(A) = 0.9, m_4(B) = 0, m_4(C) = 0.1$$

Table 1. Comparison this method with others'

	m(A)	m(B)	m(C)	m(Θ)
DS	0	0	1	0
Yager ^[7]	0	0	0.000099	0.999901
DY ^[12]	0.7980	0	0.0401	0.2019
This	0.8106	0	0.0151	0.1743

Through table 1, result of DS theory is wrong clearly,and yager can not make sure which one is the result. If the largest minus the second largest, result must be greater than 0.6, method of document [12] is not good. The method of this text not only judges the right evidence , but also meets the higher decision criterion.

6 Conclusion

This paper uses conflict factor and distance function which represents measurement of evidences confliction. Then evidence support matrix is constructed using distance function and advanced a new method based on evidence support matrix. This way is reasonable and effective by experiment. In fact, other factors which effect evidence conflict can be further discussed, such as similarity between focus elements. Different measure method can be constructed making use of diverse factors, and then compare those methods. It has great theory value as well as practical significance.

Acknowledgements. This work is supported by Shanghai Municipal Education Commission funded research and innovation in key projects under Grant No.09ZZ211; Key Discipline of Shanghai Education Commission under Grant No.J51901

References

1. Xu, X.-B., Wen, C.-L., Wang, Y.-C.: Information fusion Algorithm of Fault Diagnosis Based on Random Set Metrics of Fuzzy Fault Futures. *Journal of Electronics & Information Technology* 31(7), 1635–1640 (2009)
2. Bu, L.-P., Liu, K.-P., Hou, X.-G.: Fault Diagnosis of Rotor Based on D-S Evidence Reasoning. *Journal of Vibration, Measurement & Diagnosis* 31(1), 23–26 (2011)
3. Bing, H., Mao, S.-Y., Zhang, Y.-W., Li, S.-H.: Evidence Combination and Decision Based on DS Evidence Theory and Evidence Classification. *Journal of Electronics & Information Technology* 21(7), 894–899 (2002)
4. Bei, H., Hu, C.-H., Jiang, X.-P.: Iterative Multi-step Prediction Model Based on Theory of Evidence. *Control Theory & Applications* 27(12), 2013–2019 (2010)
5. Li, P., Yan, Y., Zheng, W.-J., Jia, J.-F., Bai, B.: Short-term Load Forecasting Based on Quantum Neural Network by Evidential Theory. *Power System Protection and Control* 38(16), 49–53 (2010)
6. Wang, Y.-L., Luo, J.-Q., Yue, H.-W.: An Approximation Algorithm of D-S Evidence Theory and Its Application in Data Fusion. *Modern Radar* 30(6), 28–31 (2008)
7. Yager, R.: On the dempster shafer framework and new combination rules. *Information Science* (41), 93–137 (1987)
8. Lefevre, E., Colot, O., Vannoorenberghe, P.: Belief function combination and conflict management. *Information Fusion* 3(3), 149–162 (2002)
9. Xiang, Y., Shi, X.-Z.: Modification on Combination Rules of Evidence Theory. *Journal of Shanghai Jiaotong University* 33(3), 357–360 (1999)
10. Murphy, C.K.: Combining belief functions when evidence conflicts. *Decisive Support Systems* 29(1), 1–9 (2000)
11. Jousselme, A.L., Grenier, D., Bosse, E.: A new distance between two bodies of evidence. *Information Fusion* 2(1), 91–101 (2001)
12. Wen, J., Zhang, A., Deng, Y.: A Novel information Fusion Method Based on Our Evidence Conflict Representation. *Journal of Northwestern Polytechnical University* 28(1), 27–32 (2010)
13. Zeng, F.-P., Min, L.: A Few Notes of Perron-Frobenius. *Journal of Guangxi University* 21(3), 222–223 (1996)

Implementation and Application of DSP-LCD Driver Based on Framebuffer

Qinghui Wang, Xiaoying Chang, Ansong Feng

College of Information Engineering,
Shenyang University of Chemical Technology,
Shenyang, China
wangqh8008@vip.sina.com

Abstract. The implementation of LCD driver based on Framebuffer in uClinux was introduced. Firstly describe the Framebuffer processing mechanism and the underlying driver interface functions, and conduct a concrete realization on ADSP-BF561 platform; Secondly, based on Framebuffer, the embedded graphical user interface MiniGUI was described briefly, and conduct a transplant on ADSP-BF561 platform; Finally, the actual test results verify the correctness of the work.

Keywords: ADSP-BF561, uClinux, LCD, Framebuffer, MiniGUI.

1 Introduction

In recent years, embedded systems and achieved rapid development in the consumer electronics and industrial control intelligent instruments has been widely used. With the rise of Linux technology, uClinux embedded field gradually occupies an important position[1]. uClinux driver structure and structure of the standard Linux driver for a similar, The only difference is uClinux does not support memory management unit[2]. Therefore, in order to take full advantage of uClinux, uClinux driver's side must be Type to wise of embedded hardware performance, Make embedded devices running fine on a graphic user interface possible. MiniGUI is a well-known embedded open-source graphical interface software, The purpose is to introduce lightweight graphical environment to run on small deives embedded uClinux. The basis to achieve all this is Framebuffer. FrameBuffer under the Linux console is a common graphical interface, it has good platform-independent, can support most of the hardware, it also can build a graphics system very easily, thus has more and more manufacturers support. This design uses ADSP-BF561 Blackfin family of processors as a development platform. Wireless video surveillance system in man-machine interface need to achieve the development platform of LCD Framebuffer driver, and can complete its graphical interface.

2 LCD Drive Mechanism Underlying

2.1 Hardware Platform

ADSP-BF561 is the development platform of ADof based on DSP at the core, integrate a variety of peripherals, including TFT color LCD touch-screen and LCD

controller. LCD controller's function is to generate display drive signals, to drive LCD displays. Users only need to read and write a series of registers, and complete the configuration and display control.

LCD controller through DMA data into the memory of the display board to display graphics. The TFT color LCD of ADSP-BF561 is 256-color, and 320 × 240 pixels, 3.5 inch LCD, LCD controller supports monochrome / color LCD display. The important step of Configuring the LCD controller is to specify the display buffer, the contents are read by the buffer, the size is determined by screen resolution and number of display colors.

2.2 Framebuffer Implementation Mechanism

In uCLinux, Framebuffer is a way to extract the graphics hardware and it is also a good user and graphical interface. Framebuffer is the memory of a device after the abstract, It allows the upper application in graphical mode directly to the display buffer reads and writes[2]. Physical memory location, feed mechanism, and so the details to be completed by the Framebuffer device drivers[2]. Display according to the specified memory block data corresponding to display the corresponding graphical interface, the LCD controller complete the corresponding drivers.

Framebuffer display buffer state in the core address space uCLinux. While in uCLinux, Each application has its own virtual address space, in application can't directly access the physical buffer address. Thus, uCLinux operating file-operations in the file structure provides mmap function. File contents can be mapped to user space. For frame buffer devices, can mapped by mapping operation. The screen buffer to the user's physical address mapping a virtual address space, After this the user can read and write access to the screen buffer virtual address, Draw on the screen[3]. Framebuffer memory block in the distribution shown in Figure 1:

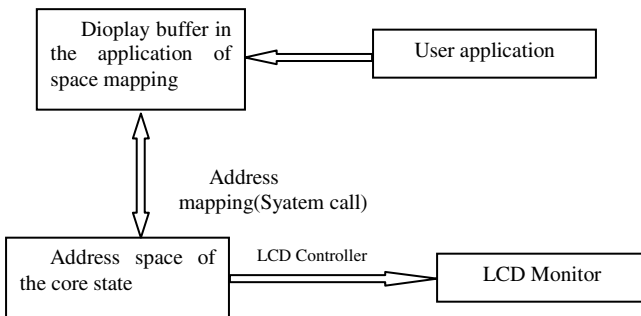


Fig. 1. Memory block distribution

3 Framebuffer Driver Implementation

Application of control by the kernel Framebuffer, there are the following three ways:

- 1) Read / write / dev / fb. The equivalent of read / write the screen buffer.
- 2) map (map) operation., By mapping, the screen can be mapped to the physical address of the buffer section of the user virtual address space, then the user can read and write access to this virtual address of screen buffer, drawing on the screen.
- 3) I / O control. For the frame buffer device, the device file ioctl operations can read / set the display device and screen parameters, such as resolution, display colors, screen size and so on. ioctl is controlled by the underlying driver to complete.

Therefore, Framebuffer driver to complete the work is assigned the size of memory, LCD control register initialization, set var to modify hardware information and fix the corresponding information. Framebuffer device a “file level - driver layer” interface mode. In the file level, uCLinux defined a series of read and write control operation:

```
static struct file-operations fb-fops = {
owner: THIS-MODULE,
read: fb-read, /*read*/
write: fb-write, /*write*/
ioctl: fb-ioctl, /* control operation*/
mmap: fb-mmap, /*mapping operation*/
open: fb-open, /*open operation*/
release: fb-release, /*shutdown*/
};
```

Application basis having a file open frame buffer device, if successfully opened, the frame buffer device can read, write and other operations [4]. The framebuffer device to read and write the most important task is to obtain frame buffer in memory space and the corresponding physical address some of the characteristics of LCD.

Figure 2 shows the ADSP-BF561 platform to write applications that display graphics frame buffer device the whole process. Basis described above, based on the ADSP-BF561 platform writing framebuffer driver needs to do the actual work is as follows: 1) write the initialization function. Initialization function initializes the LCD controller, set the display mode by writing to register and display the number of colors, and then assign LCD display buffer. Through the kmalloc function in the Linux distribution of a continuous space.

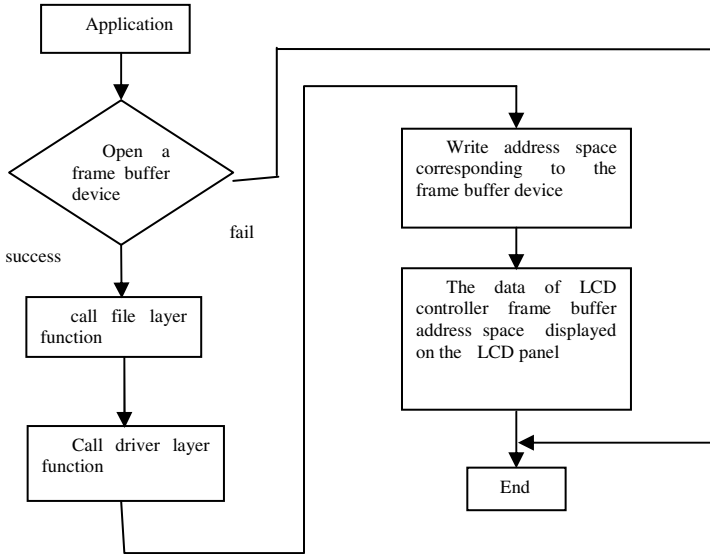


Fig. 2. driver implement block diagram

This ADSP-BF561 platform, using a 24-bit true-color digital screen. The approach is based on 8-bit data into the incoming RGB mode. According to the formula: Image data size = (image width image height * number of bits of recorded pixels) / 8, namely the need to allocate a buffer for the display $(320 \times 240 \times 24) / 8 = 225\text{K}$ bytes, the buffer is usually assigned in the large-capacity external SDRAM, the start address stored in the LCD controller registers. Finally, initialize a fb-inf Structure, filling one of the member variables, and call register-framebuffer (& fb-info), the fb-info registration into the kernel. Here is the frame buffer device initialization code:

```

int init ADSP-BF561fb-init(void)
{
    .....
    struct known-lcd-panels * p-lcd;//LCD Panel
    feature
    /* Init LCD Controller */
    lcd-Init(p-lcd->bpp);
    lcd-DispON(); //Open LCD display
    /* Init Framebuffer Parameter */
    fb-parameter-init(p-lcd);
    memset(&fb-info.gen, 0, sizeof(fb-info.gen));
    /* Print some information */
    printk("ADSP-BF561framebuffer driver: Init
    ready\n");
    .....
    return 0;
}

```

2) the preparation of structure fb-info function pointer fb-ops in the corresponding member functions, and in accordance with the hardware configuration register.

struct fb-ops in include / linux / fb.h define these functions are used to set / get fb-info structure member variable. When the application file to the device operating time Ioctl will call them. Only need the following three functions can be achieved:

```
struct fb_ops {
    .....
    int (*fb-get-fix)(struct fb_fix_screeninfo
    *fix, int con, struct fb_info *info);
    int (*fb-get-var)(struct fb_var_screeninfo
    *var, int con, struct fb_info *info);
    int (*fb-set-var)(struct fb_var_screeninfo
    *var, int con, struct fb_info *info);
    .....
};
struct fb-ops in include / linux / fb.h defined
```

These functions are used to set / get fb-info structure member variable. When the application file to the device operating time Ioctl will call them.

ADSP-BF561 based write driver, the main work based on the ADSP-BF561 hardware set the appropriate register. Main settings: BF561fb_mach_info structures that define the display of some of the information used, such as clock, size, etc.; BF561fb_activate_var function on register settings related to the ADSP-BF561 LCD controller of the set, these registers set according to the use of TFT-LCD screen to set it up; BF561fb_set_controller_regs and BF561fb_lcd_nit function settings, CPU and LCD related to the physical connection, LCD and need to connect the CPU to set specific individual CPIO register.

3) test driver. The completed LCD driver statically to the kernel which, after the image downloaded to the development board. Upper left corner of the LCD screen will display a small penguin icon. View device file:

```
[root]# ls -al /dev/fb/0
```

4 Experimental Results and Analysis

In this experiment, uClinux provides mmap function in the file-operations structure, The contents of the file directly mapped to user space. By mapping operations, Framebuffer map the physical address of screen buffer to a virtual address of the user's space, users by reading and writing this virtual address to access the screen buffer, drawing on the screen. Transplant MiniGUI system, need to realize uClinux kernel Framebuffer driver. Corresponds to the ADSP-BF561 in LCD controller, this part of the driver program must be statically compiled into the kernel, at system startup by the startup parameters passed into the kernel to activate the device. In the experiments, according to the ADSP-BF561 hardware set corresponding register, and each configuration setting is very critical, should be strictly in accordance with hardware platform and system, otherwise it will cause image distortion or can not display graphical interface.

5 Conclusion

Tested, the LCD driver in this article have been achieved successfully on the ADSP-BF561 platform. Framebuffer as a basis graphics facility, abstract memory into device, through this device to read and write directly to operate memory. MiniGUI transplanted in the process of ADSP-BF561, play a key role. Make small in-class graphics interface build on DSP platform successfully, achieve the man-machine interface of wireless network video surveillance system.

References

1. Yu, M., Chen, X., Fang, H.: Linux Programming Definitive Guide. Machinery Industry Press, Beijing (2001)
2. Buell, A.: FrameBuffer How To 1(1), 22 (1999)
3. Ze, T.: Embedded system development and application. Beijing University of Aeronautics Astronautics University Press, Beijing (2005)
4. Zou, S.: Embedded Linux Design and Application. Tsinghua University Press, Beijing (2002)
5. Beijing Feynman Software Technology Co., Ltd. MiniGUI technical white paper (2005), <http://www.minigui.org>

A Novel Energy Efficient Cluster-Based Algorithm for Wireless Sensor Networks

Bin Jiao¹ and Youping Xiong²

¹ Electric School, Shanghai DianJi University,
No.690, Jiang Chuan Rd., Min Hang District, Shanghai, 200240, China
jiaob@sdju.edu.cn

² School of Information Science and Engineering,
East China University of Science and Technology, No.130,
Mei Long Rd., 200237, Shanghai, China
youping918@sina.com

Abstract. The energy, memory and processing power of the sensor nodes in Wireless Sensor Network(WSN) are constrained. Many researchers have proposed energy-efficient routing algorithms to save energy consumption for WSN. In this paper, an Energy Efficient Clustering Algorithm(EECA) is proposed. This algorithm chooses Cluster Heads(CHs) according to the residual energy and number of neighbor nodes. When transmitting data packets between CHs, both distance and residual energy are considered. Simulation results show that this algorithm greatly reduces the energy consumption of the sensor network, and also the number of data packets transmitted to the sink node increases.

Keywords: Wireless Sensor Network, routing algorithm, energy consumption.

1 Introduction

The wireless sensor network is a new kind of sensing system that is used to monitor events in the area which is always difficult to access. It has been widely used in military, medical, and environmental applications[1]. With the development of sensors, embedded systems and wireless communication, wireless sensor network technology has been developed rapidly. The wireless sensor network has extended the ability to acquire information, making it possible to connect the physical information in real world into networks.

The typical sensor network is composed of many sensor nodes and one sink node. The sink node is also called base station. The sensor nodes have the ability to transmit, receive, and process data and can communicate with each other, and sink node is a special node which has stronger processing power and more energy. Those tiny sensors have many restrictions, such as limited energy, memory, and processing ability, and energy is an important factor for the network because it is almost impossible to recharge or replace the battery of the sensors. Routing protocols must be energy efficient to prolong the lifetime of the network. The main issue of designing routing protocols for WSN is energy consumption. So how to use limited energy to maximize the lifetime of WSN is the focus of our discussion. Currently, many energy-efficient

routing algorithms have been proposed to prolong the lifetime of the sensor network[2]. But the main problem of the existing protocols is that it chooses an optimal path frequently, as a result the sensors on the path die in a short time. To overcome the shortages of recent routing algorithms, an energy efficient clustering algorithm, which is based on clusters, is proposed in this paper. To evaluate the performance of the algorithm, the number of data packets received at the base station is also taken into consideration.

In this paper, the main purpose of our research is to prolong the lifetime of the sensor network and to maximize the number of transmitting data packets. The proposed algorithm adopts a new way to select Cluster Heads(CHs). Data transmission between CHs is realized in a new way. The rest of the paper is organized as follows. Section 2 describes the previous research for WSN. Section 3 describes the radio energy dissipation model. Section 4 states the proposed algorithm. Section 5 compares the performance of the proposed algorithm with the popular routing algorithm LEACH[3], and finally section 6 represents the conclusion.

2 Related Work

The existing routing protocols are classified as flat routing protocols and hierarchical routing protocols by the topology of the sensor network. During the hierarchical routing protocol, the sensor nodes are divided into clusters. Each cluster consists one Cluster Head(CH) and many cluster members. The cluster member only communicates with CH in each cluster. CHs aggregate data received from its members and remove redundant data, and then send the data to the sink node. The hierarchical routing protocols have many advantages, such as reducing control messages, bandwidth re-usability, and good scalability. Recently many hierarchical routing algorithms have been proposed for WSNs. One famous work is Low Energy Adaptive Clustering Hierarchy(LEACH), which was proposed by Heinzelman and his colleagues. LEACH is energy efficient and has better performance compared with traditional routing protocols.

To distribute the overload to each sensor node of the network, LEACH changes the CHs each round periodically. Cluster member sends data to CH, and then CH aggregates the data and sends the data to sink node. Each period is called a round. A round contains two steps: cluster set up phase and data communication phase. During set-up phase, each node randomly generates a number between 0 and 1, if this number is less than $T(n)$, then it broadcasts a message to its neighbors that it becomes a new CH. The cluster members choose the most suitable CH according to the signal strength received, and join the cluster.

However, the selection of CHs is random in LEACH. It does not take energy into consideration. It may select a low-energy node as a CH, and this is not suitable for the application. Another, when data aggregation is finished by a CH, it directly sends data packets to the sink node without routing. As a result, the CHs which are far away from the sink node consume much energy. To overcome the shortcomings of LEACH, many improved algorithms are proposed. Lindsey et al. introduced chain into clustering and proposed PEGASIS(power-efficient gathering in sensor information systems)[4]. In this algorithm, nodes are connected in a chain and data is transmitted through the chain.

It avoids the energy consumption caused by the selection of CHs. However, if one of the nodes in the chain dies, data transmission may fail. To overcome the shortcomings of the conventional hierarchical routing algorithms, an improved energy-efficient cluster-based algorithm is proposed in this paper. Energy consumption and number of transmitting packets are taken into consideration in order to improve the performance of the WSN.

3 Radio Energy Dissipation Model in WSN

The energy dissipation model is the same as depicted in paper [5]. When a node transmits a k -bit message to its neighbor with a distance d , E_{Tx} is used by this node and the dissipated energy is given as follows:

$$E_{Tx}(k, d) = \begin{cases} E_{elec} \times k + E_{amp} \times k \times d^2, & d < d_0 \\ E_{elec} \times k + E_{amp} \times k \times d^4, & d \geq d_0 \end{cases} \quad (1)$$

When a node receives the k -bit message, E_{rx} is used by the node and the dissipated energy is given as follows:

$$E_{Rx}(k) = E_{Rx} \times k \quad . \quad (2)$$

And when the CH aggregates the k -bit message, E_{fuse} is used and the dissipated energy is given as follows:

$$E_{fuse}(k) = E_{fuse} \times k \quad . \quad (3)$$

Where E_{elec} represents the power dissipated to run the transmitting or receiving unit in sensor nodes, and E_{amp} represents the power dissipated to run the transmitter amplifier.

4 Proposed Algorithm

The proposed algorithm is a novel clustering algorithm. The algorithm is divided into two major phases: cluster setup phase and data communication phase.

4.1 Cluster Setup Phase

In the cluster setup phase, a new cluster selecting formula is designed to make the algorithm energy-efficient, and sensor nodes choose CH according to another new formula. The procedures can be described as follows:

(1) The sink node broadcasts a cluster-constructing message to all the nodes to start clustering. When the nodes receive the cluster-constructing message, each node acquires the information including the distance and residual energy of its neighbor nodes.

(2) Each sensor node calculates its P value according to the formula below. The nodes which have the biggest P value then become CH in this round. In this paper, the number of selected CHs is \sqrt{n} , where n is the total number of the living nodes in each

round. The node which has more residual energy and more neighbors has a higher probability to become CH. And the node which has become a CH broadcasts a message to its neighbor that it has become a Cluster Head.

$$P_i = (1 - a) \bullet E_i + X_i \bullet a . \quad (4)$$

where a is the coordinating argument, $0 < a < 1$, X is the number of neighbor nodes of node i .

(3) After receiving the message from CHs, each of the non-cluster nodes joins the most suitable Cluster Head according to the formula below. The CH which has the biggest c value becomes the node's cluster head. This means CHs which have more energy and has a short distance from the node can become this node's cluster head.

$$c_{ij} = \frac{E_j^\alpha}{r_{ij}^\beta} . \quad (5)$$

where r_{ij} is the distance between node i and CH node j , E_j is the residual energy of CH node j , and α , β are the coordinating arguments.

4.2 Data Communication Phase

Data communication is done between CHs. When a CH needs to send data, it chooses another CH as next hop, and this procedure is repeated until all the data packets are sent to sink node. Those which will become next CH should have a shorter distance from the sink node. Another, the selection of next CH needs to consider energy and distance as well. So, we can reuse formula(5) to select next hop of the CH. And the coordinating arguments should be modified coordinately to satisfy the needs of the environment. When the data packets are all sent to the sink node from the source node, data communication is finished in one round. Each round is repeated as this until all data packets are sent to the sink node.

5 Performance and Evaluation

To evaluate the performance of the proposed algorithm, we simulate the algorithm in contrast to LEACH. The specific environment is 100 nodes are randomly deployed in a square which is 100 meters long and 100 meters wide, the coordinate of the sink node is (50,150). The initial energy of each node is given as 0.5J. The value of other parameters is given as follows: $E_{elec} = 50\text{nJ/bit}$, $E_{amp} = 100\text{pJ}/(\text{bit} \times \text{m}^2)$, $E_{fuse} = 5\text{nJ/bit}$, and the distance of communication range is 20 meters. We assume that each data packet contains 4000 bits. Our simulation is carried out under the circumstance mentioned above.

Figure 1 shows the total energy consumption of the sensor nodes in each round, we can see that the proposed algorithm EECA consumes less energy than LEACH. EECA chooses CHs based on the residual energy of the nodes, while LEACH selects CHs randomly. In LEACH data communication is through single-hop, while in EECA data communication is based on multi-hop. Therefore, energy consumption is reduced when

a CH broadcasts a message to its members in EECA. The simulation result shows that EECA greatly reduces energy consumption.

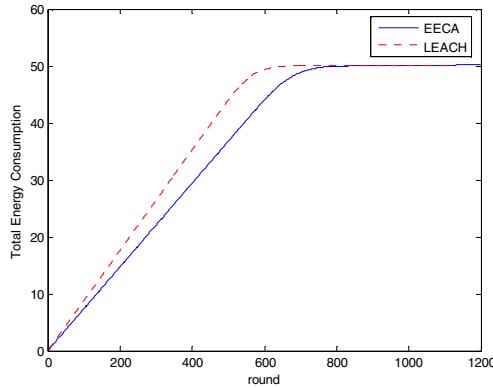


Fig. 1. Total energy consumption of the sensor nodes in each round

Figure 2 shows the number of received data packets at the base station. As the figure shows, the number of data packets received at the base station increases while using EECA compared to LEACH. EECA is more efficient as it does not only take distance into consideration and it uses multi-hop communication. By using the proposed idea, the proposed algorithm finally achieves good performance. The circumstance of Data communication is better.

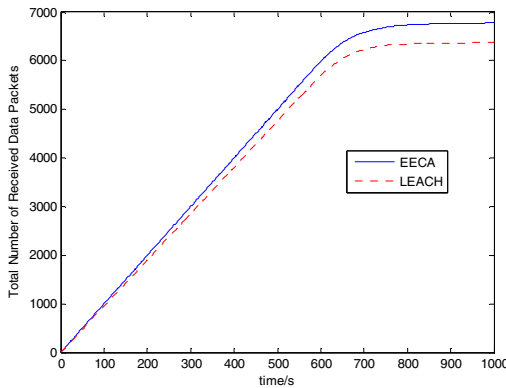


Fig. 2. Number of received data packets

6 Conclusion

To improve the performance of WSN, we have proposed a novel energy efficient clustering algorithm. This algorithm chooses Cluster Heads according to the residual

energy and distance between sensor nodes. The nodes with more energy have a higher probability to become CH. When a CH needs to transmit data to the destination, we adopt the same idea in data communication. Simulation results show that the proposed algorithm effectively reduces energy consumption, indicating that it can prolong the lifetime of WSN. The number of data packets received at the base station increases by using this algorithm. The performance of this algorithm is more efficient than traditional hierarchical algorithms.

Acknowledgement. This work was supported in part by Shanghai Municipal Science and Technology Commission. (Grant No. 10JC1405800), Project of Science and Technology Commission of Shanghai Municipality(08DZ1200505), Project of Shanghai Municipal Economic and Information Commission(09A118), Key Discipline of Shanghai Municipal Education Commission(J51901), and Shanghai Dianji University(09C401).

References

1. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y.: A Survey on Sensor Networks. *IEEE Communications Magazine* 40(8), 102–114 (2002)
2. Camilo, T., Carreto, C., Silva, J.S., Boavida, F.: An Energy-Efficient Ant-Based Routing Algorithm for Wireless Sensor Networks. In: Dorigo, M., Gambardella, L.M., Birattari, M., Martinoli, A., Poli, R., Stützle, T., et al. (eds.) ANTS 2006. LNCS, vol. 4150, pp. 49–59. Springer, Heidelberg (2006)
3. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: *Proceedings of the 33rd Hawaii International Conf. on System Sciences*, pp. 3005–3014. IEEE Computer Society, Maui (2000)
4. Lindsey, S., Raghavendre, C.: Pegasis: power-efficient gathering in sensor information systems. *IEEE Transactions on Parallel and Distributed Systems* 13(9), 924–932 (2002)
5. Heinzelman, W.R.: An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans. on Wireless Communication* 1(4), 660–670 (2002)

A 3D Temperature Field Reconstruction Algorithm Based on Acoustic Travel-Time Tomography

Hua Yan, Kun Li, and ShanHui Wang

School of Information Science and Engineering,
Shenyang University of Technology, Shenyang, China
yanhua_01@163.com, likun0606@126.com

Abstract. In order to improve the reconstruction accuracy of 3D temperature field, a new 3D temperature field reconstruction algorithm based on acoustic travel-time tomography is proposed. Four typical temperature field models, i.e. one-peak symmetry model, one-peak bias model, two-peak symmetry model and four-peak symmetry model are used to verify the effectiveness of the new algorithm. 3D display of reconstructed temperature fields and reconstruction errors are given. According to the simulation result, we can conclude that the new algorithm has high reconstruction accuracy and good anti-noise ability.

Keywords: temperature field, reconstruction algorithm, acoustic tomography.

1 Introduction

Temperature field reconstruction based on acoustic travel-time tomography [1, 2, 3] uses the dependence of sound speed in materials on temperature along the sound propagation path. At present, the research on temperature field reconstruction by acoustic travel-time tomography is mainly conducted in two-dimensions. 3D temperature field reconstruction^[4, 5] was discussed by only few researchers. And the reconstruction algorithm they used was least square method, which requires that the number of travel-times is greater than the number of spatial points where the fields are being reconstructed. This limits the spatial resolution of the reconstructed fields.

In this paper, a new reconstruction algorithm of 3D temperature field based on acoustic travel-time tomography is proposed; its number of spatial points of the reconstructed fields is much greater than the number of travel-times. As a result, high reconstruction accuracy can be achieved.

2 The Principle of Temperature Field Reconstruction Based on Acoustic Travel-Time Tomography

Temperature measurement by acoustic method is based on the principle that the sound velocity in a medium is a function of the medium temperature. The sound velocity c in a gaseous medium at an absolute temperature T is given by[3, 4, 5]

$$c = B\sqrt{T} \quad (1)$$

where B is a constant decided by gas composition. The value of B for air is 20.45.

To survey the temperature distribution in a space, several acoustic sensors (sources/receivers) should be installed on its periphery. In this paper, the space to be measured is 10m×10m×10m, and 32 acoustic sensors are amounted on its periphery. Fig.1 shows the position of 32 sensors and the 172 effective sound wave paths. On the basis of sensor positions and the travel-time measurements along each effective sound wave path, the temperature field can be reconstructed by using suitable reconstruction algorithms [1, 2, 3].

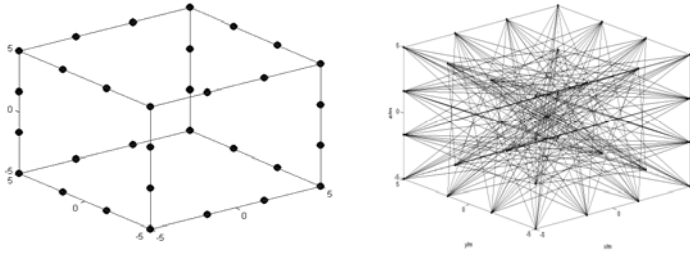


Fig. 1. A space surrounded by 32 sensors and the 172 effective sound wave paths

3 The Principle of the New Reconstruction Algorithm

Supposing that the distribution of temperature is $T(x,y,z)$, the distribution of reciprocal of sound speed is $f(x, y, z)$, then $T(x, y, z)$ can be presented as:

$$T(x, y, z) = 1/[B \cdot f(x, y, z)]^2 \tag{2}$$

The acoustic travel-time g_k along the path k can be expressed as

$$g_k = \int_{p_k} f(x, y, z) dp_k \quad (k = 1, 2, \dots, K), \tag{3}$$

where K is the number of the effective sound travel paths, that is, the number of the effective sound travel-times.

Divide the space to be measured into M pixels (cells), which means the number of spatial points of the reconstructed fields is M . Let the center coordinates of pixel m is (x_m, y_m, z_m) . Expand $f(x,y,z)$ over a finite set of basis functions as follows:

$$f(x, y, z) = \sum_{m=1}^M \epsilon_m \phi_m(x, y, z) \tag{4}$$

Where ϵ_m is the modulus determining the distribution of sound speed reciprocal, and $\phi_m(x, y, z)$ is an exponential basis function:

$$\phi_m(x, y, z) = e^{-\beta \sqrt{(x-x_m)^2 + (y-y_m)^2 + (z-z_m)^2}} \tag{5}$$

Where β is the shape parameter of exponential basis functions, it is related to the size of the space and the layout of the sound sources/receivers.

Combining equation (3), (4) and (5), we have:

$$g_k = \sum_{m=1}^M \epsilon_m \int_{p_k} \phi_m(x, y, z) dp_k = \sum_{m=1}^M \epsilon_m a_{km}, \quad a_{km} = \int_{p_k} \phi_m(x, y, z) dp_k \quad (6)$$

Define: $A = (a_{km})_{k=1, \dots, K, m=1, \dots, M}$, $g = (g_1, \dots, g_K)^T$, $\epsilon = (\epsilon_1, \dots, \epsilon_M)$

The equation (6) can be expressed as:

$$g = A\epsilon \quad (7)$$

Since matrix A is ill-conditioned, we solve equation (7) by means of the singular value decomposition (SVD)^[6] of matrix. Using SVD, matrix A can be written uniquely as

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^T, \quad \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r) \quad (8)$$

Where $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ are the non-zero singular values of matrix A , U and V are orthogonal matrixes and their columns are eigenvectors of AA^T and $A^T A$ respectively. Then the pseudo inverse of A can be expressed as:

$$A^+ = V \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^T, \quad \Sigma^{-1} = \text{diag}(1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_r) \quad (9)$$

So we have:

$$\epsilon = A^+ g \quad (10)$$

Combining equation (4) and (10), the sound speed reciprocal of each pixel can be gained. And then the temperature of each pixel can be obtained by using equation (2). Finally, the detailed temperature distribution of the space can be calculated by 3D interpolation method.

3 Verification of the New Algorithm

In order to test and verify the effectiveness of the new algorithm, four typical temperature field models are reconstructed. They are

One-peak symmetry model defined as:

$$T(x, y) = \frac{2000}{0.05(x^2 + y^2 + z^2) + 1} \quad (11)$$

One-peak bias model defined as:

$$T(x, y) = \frac{350}{0.05[(x + 2)^2 + (y - 2)^2 + (z - 2)^2] + 1} \quad (12)$$

Two-peak symmetry model defined as:

$$T(x, y) = \frac{350}{0.2[(x+3)^2 + y^2 + (z-1.6)^2] + 1} + \frac{350}{0.2[(x-3)^2 + y^2 + (z-1.6)^2] + 1} \tag{13}$$

Four-peak symmetry model defined as:

$$T(x, y) = \frac{350}{0.5[(x+2.5)^2 + y^2 + (z-2.5)^2] + 1} + \frac{350}{0.5[(x-2.5)^2 + y^2 + (z-2.5)^2] + 1} + \frac{350}{0.5[(x+2.5)^2 + y^2 + (z+2.5)^2] + 1} + \frac{350}{0.5[(x-2.5)^2 + y^2 + (z+2.5)^2] + 1} \tag{14}$$

To evaluate the quality of the reconstructions, following reconstruction errors are used. They are

The root-mean-squared error of the reconstructed field

$$E_{rms} = \frac{\sqrt{\frac{1}{M} \sum_{j=1}^M [T(j) - \hat{T}(j)]^2}}{T_{ave}} \times 100\% \quad j = 1, \dots, M \tag{15}$$

The error of the average temperature reconstructed

$$E_{ave} = \frac{T_{ave} - \hat{T}_{ave}}{T_{ave}} \times 100\% \tag{16}$$

The error of the maximum temperature reconstructed

$$E_{max} = \frac{T_{max} - \hat{T}_{max}}{T_{max}} \times 100\% \tag{17}$$

where $T(j)$ is the model temperature of pixel j , $\hat{T}(j)$ is the reconstruction temperature of pixel j , T_{ave} is the average temperature of the model field, \hat{T}_{ave} is the average temperature of the reconstruction field, T_{max} is the maximum temperature of the model field, \hat{T}_{max} is the maximum temperature of reconstruction field.

In this paper, the measured space is divided into $M=10 \times 10 \times 10$ pixels, the shape parameter β in equation (5) is 0.0001. Correspondingly, the condition number of matrix A is 9.48×10^6 . The travel-time data corresponding to above typical models are calculated, exactly and with simulate measurement errors. Then, the temperature distributions of these models are reconstructed with these data and the new algorithm.

The model fields and the reconstructed fields obtained by the new algorithm when the travel-times are noise-free are shown in Fig.2, the reconstruction errors are given in Table 1. In reference [4], One-peak symmetry model is reconstructed by using noise-free travel-times and the least square method, the reconstruction errors are $E_{ave} = 10.17\%$, $E_{max} = 17.11\%$. Obviously, the reconstruction accuracy of the new algorithm is much higher than that of the least square method.

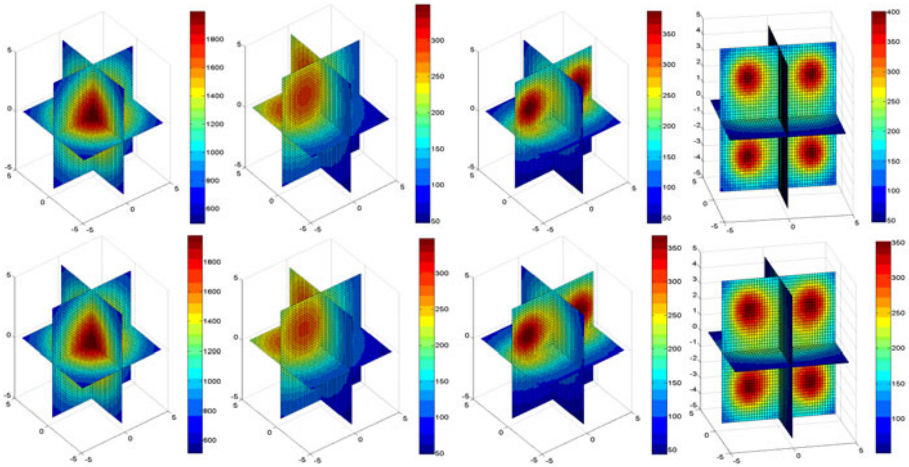


Fig. 2. The model temperature fields and the reconstructed temperature fields when the travel-times are noise-free, the pictures above are the model fields, and the pictures below are the reconstruction fields.

Table 1. Errors of the reconstruction fields when the travel-times are noise-free

Error (%)	E_{ave}	E_{max}	E_{rms}
One-peak symmetry model	0.79%	0.86%	1.47%
One-peak-bias model	0.69%	0.52%	1.36%
Two-peak symmetry model	1.16%	4.14%	2.34%
Four-peak symmetry model	1.79%	11.67%	5.11%

In order to test the anti-noise ability of the new algorithm, corrupted travel-time data are produced by adding Gaussian noise to the exactly data. The mean of the noise is zero and the standard deviation of noise is $1e-4$, $1e-3$, $1e-2$ respectively. The root-mean square reconstruction errors of the four models under different noise levels are given in Table 2, and the reconstructed fields of the four-peak model are shown in Fig.3 as examples.

Table 2. Errors of the reconstruction fields when the travel-times are corrupted

Standard deviations of the noise	E_{rms}			
	One-peak symmetry model	One-peak bias model	Two-peaks symmetry model	Four-peaks symmetry model
$1e-4$	1.33%	1.38%	2.37%	5.14%
$1e-3$	1.65%	1.78%	2.72%	5.49%
$1e-2$	9.51%	10.39%	11.57%	12.68%

Simulation results show that the new algorithm can reconstruct the model fields from the exact travel-times quite accurately, no matter the model is simple with one-peak or complicated with four-peaks; the new algorithm can withstand the noise in travel-times effectively as long as the standard deviation of the noise does not exceed $1e-2$.

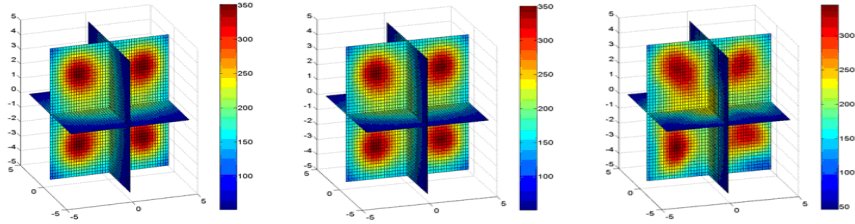


Fig. 4. The reconstructed temperature fields of the four-peak symmetry model with corrupted data, from left to right, the stand deviation of adding noise is $1e-4$, $1e-3$ and $1e-2$

5 Conclusion

A new reconstruction algorithm of 3D temperature field based on acoustic travel-time tomography is proposed. Unlike the well-known least square method, the number of spatial points of a reconstruction field obtained by the new algorithm is much greater than the number of travel-times. As a result, a reconstruction accuracy, which is much better than that of the least square method, can be achieved. Simulation results show that the new algorithm has good ability in 3D complex temperature field reconstruction and anti-noise.

Acknowledgments. This work is supported by National Natural Science Foundation of China (grant no. 60772054), Specialized Research Fund for the Doctoral Program of Higher Education of China (20102102110003) and Shenyang Science and Technology Plan (F10213100).

References

1. Barth, M., Raba, A.: Acoustic topographic imaging of temperature and flow fields in air. *Measurement Science and Technology* 22(3) (2011)
2. Holstein, P., Raabe, A., Muller, R., et al.: Acoustic tomography on the basis of travel-time measurement. *Measurement Science & Technology* 15(6), 1240–1248 (2004)
3. Bramanti, M., Salerno, E.A., Tonazzini, A., Gray, A.: An acoustic pyrometer system for tomographic thermal imaging in power plant boilers. *IEEE Transactions on Instrumentation and Measurement* 45(1), 159–161 (1996)
4. Liu, L.L.: Research on acoustic reconstruction of three-dimensional temperature fields and velocity. Northeastern University (2007)
5. Wang, R.X.: Investigations on Acoustic Detecting Techniques and Simulation of 3-D Temperature Field. Daqing Petroleum Institute (2007)
6. Guo, W.B., Wei, M.S.: Singular value decomposition and its application in the theory of generalized inverse, pp. 14–15. Science Press, Beijing (2008)

Retraction: Evaluating Gigabit Switches Using Perfect Communication

JianMing Zhao and YongNing Guo

Department of Mathematics and Computer Science
Fuqing Branch of Fujian Normal University, Fuqing 350300, China
{zhaoke_cn, guoyn}@163.com

Several conference proceedings have been infiltrated by fake submissions generated by the SCIdgen computer program. Due to the fictional content the chapter “Evaluating Gigabit Switches Using Perfect Communication” by “JianMing Zhao and YongNing Guo” has been retracted by the publisher. Measures are being taken to avoid similar breaches in the future.

Application of Virtual Prototype Technology to Simulation Test for Airborne Software System

Yang Wang¹, Lifeng Wang¹, and Zewei Zheng²

¹ School of Aeronautic Science and Technology, Beihang University, Beijing

² School of Automation Science and Electrical Engineering,

Beihang University, Beijing 100191, China

zhuoyue771988@126.com

Abstract. This paper develops a detailed analysis of embedded software testing method and virtual prototype technology. Then it makes advantage of virtual prototype platform, including universality, extensibility, easy upgrade, high intelligence and short development cycle, and applies them to airborne software testing. In this way, it makes up for the shortcomings, such as long development cycle, poor controllability and high cost of the embedded software platform and dedicated external equipment. In this paper, an airborne software testing technology based on the virtual prototype is proposed, combining with the practical work of software testing.

Keywords: virtual prototype, airborne software, simulation, software test.

1 Introduction

Along with the wide use of the technology of micro-electronics, computer and automation in military equipment, the degree of autoimmunization and complexity of military equipment have been developed continuously, which makes technical logistics safeguard more difficult and brings more requirements to the personnel and equipment [1]. Obviously, these areas have very strict requirements on software reliability, because the software failure will cause enormous economic damages and security dangers to human beings [2]. Software testing as an important means of software quality assurance is widely used to improve software quality. In recent years, the new challenge of aviation embedded software testing is proposed because of the constant turmoil of aerial embedded software market and the situation of intense competition. In this paper, the virtual prototype technology thoughts and methods are applied to aerial embedded software testing, which is used to solve the above problems.

2 Concept of Virtual Prototype Technology

Virtual Prototype Technology is a kind of digital design method of computer simulation model based on product. These digital models also known as virtual

prototype can help us quickly analyze and compare various design schemes in the stage of conceptual design. It combines development models in different fields of engineering to simulate real product in the aspects of appearance, function and behavior. Virtual prototype supports concurrent engineering. The basic idea is that conducting simulation in the virtual prototype to forecast, evaluate and optimize the product features, performance, appearance, can improve product quality, reduce the cost of development, and shorten the development cycle before the implementation of physical prototype.

The target of VPT is to reduce or even eliminate the physical prototype, so it is especially suitable for this condition that manufacturing cycle of the physical prototype is long and expensive. Fig. 1 gives a comparison with two product development models in which one based on virtual prototype (solid line) and the other one based on the physical prototype (dotted line).

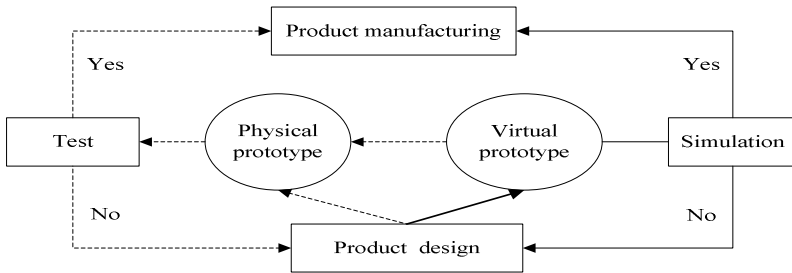


Fig. 1. Comparison of two development modes based on virtual prototype and physical prototype

From the above discussion, we can see that the application of virtual prototype technology has the following advantages. Firstly, it can reduce the design cost and shorten the product development process required in time. Secondly, it can help the physical prototype to verify and test the design. Thirdly, Meantime,more designs schemes can be tested ,which is incomparable for physical prototype.It helps to improve the quality of products. Finally, it can make communication between the design groups more convenient, for virtual prototype technology supports concurrent design.

3 The Characteristics of Embedded Software

The software testing is an important means to guarantee the software quality, reliability and security of military embedded software [3]. This kind of software is extremely important and is classified as safety critical software [4]. Most of embedded systems have features of hard real time and resource-constrained and usually require high reliability [5]. Seeing about the military embedded software testing, we discovered that comparing with the non-embedded mainframe software,

embedded software has its own characteristics which has decided the particularity and the complexity of its test.

- 1). Embedded software does not run unless its particular hardware environment is built successfully.
- 2). Embedded software testing is not only to ensure high reliability in a specific environment, but also to ensure the real-time characteristics. Especially in the airborne software, it may cause great losses, if embedded software does not have the ability of real-time response in certain environment.
- 3). Test quality, process and cost are challenged, while the work is huge.
- 4). Existing process issues.

For small and medium sized software projects, they are effected little by the above process issues, because they are easy to develop and convenient to use and maintain. But for large and medium-sized projects, the impact of these issues are more prominent, because of its long development cycle, complexity, and difficulty of maintenance.

4 Application of Virtual Prototype to Airborne Software Test

To sum up, the issues of airborne software testing including running environment, quality, schedule, cost and so on, can be improved by the simulation test methods based on virtual prototype technology.

4.1 Establishment of Virtual Prototype Model

Virtual prototype is a mathematical model to replace the physical prototype. It is an aggregate of embedded systems or equipment in the field of digital model, containing all key features of the real system or equipment [6]. The establishment of virtual prototype is a continuous refinement and improvement process based on model. Establishment of virtual prototype is shown in Fig 2.

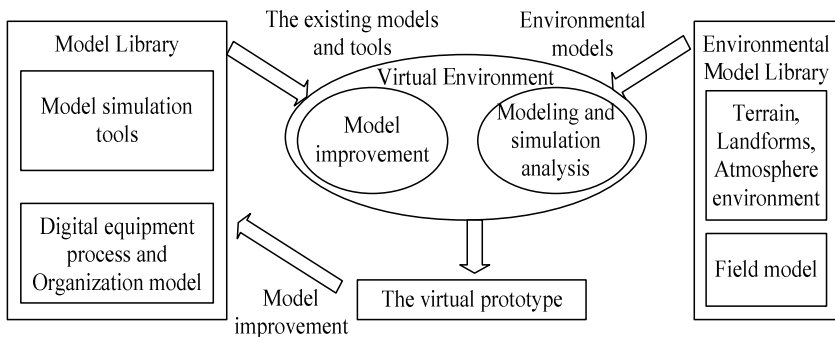


Fig. 2. The establishment of Virtual Prototype

4.2 Structure of General Simulation Platform

There are two main issues to be solved for simulation and test system. The first one is that we need to simulate all kinds of change rules of the control objects in the system by means of simulation, so that the system can be run in a "real" environment. The second one is that various key parameters of the running system should be tested by means of test, in order to give a conclusive suggestion for the various technical indicators after real time analysis. Therefore, the simulation and test system should have the following functions: the real running environment signal of the object to be tested is generated and the simulation signal is tested.

In order to achieve the basic functions, general simulation testing system mainly consists of three parts which are the functional simulator, test resource and data analyzer. The core part of a simulation and test system is the functional simulator, whose main function is to simulate a variety of changes rule of the tested equipment, to generate actions or behavior trajectories, to receive control commands and to simulate dynamic parameters of the equipment according to simulation model of the tested equipment. Test resource includes the test of hardware and software, and execution of system test. Data analyzer mainly accomplishes the real-time analysis and post hoc analysis of the simulation and test system, make comprehensive analysis on tested equipment and draw some useful conclusions.

4.3 Simulation and Test Platform Based on the Virtual Prototype

According to the above analysis, this paper puts forward to a construction scheme of airborne software simulation and test platform based on virtual prototype, as shown in Fig. 3.

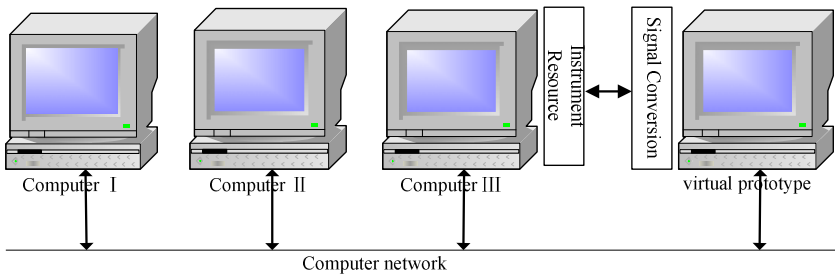


Fig. 3. General simulation and test system

Computer one is a control computer of the system, which controls and manages the whole system. It can also control to simulate all kinds of rules of the tested equipment according to the needs of test. Then, the virtual prototype produces actions or behavior trajectories. Computer two is a control computer of the test system. It does not only accomplish test tasks, but also controls the instrument resources to complete the signal simulation and fault injection. The analog signal is to simulate the real

running environment of the test object. Computer three is a real-time and post hoc analysis workstation of the system. It can display and analyze the signal, and monitor instrument running environment and state of the instrument. So it is able to accomplish behavior analysis and judgment. As the signals between the instrument resource and the tested equipment are not the same, there is a need for an intermediate system to convert them into the same signals. The signal conversion system is just to do it.

5 Virtual Test

After the establishment of the model and simulation system, the virtual prototype can be used to test reliability of the real prototype. We can save a lot of development cost or other potential costs, for the work has been done before establishing a complex system or the real prototype. Virtual test mainly includes functional test, physical test.

5.1 Functional Test

Functional test are used to testify the operation of system, including the accuracy of system, the interaction between subsystem and system, and other operation conditions. All kinds of simulation function of virtual prototype should be used, such as discrete events, continual system simulation and parameter analysis. At the same time, there are several kinds of technologies in the functional test, for example, equivalence partitioning, boundary value analysis and error guessing and so on. The functional test is mainly used to discover the following mistakes:

- 1). If there exists any incorrect or omissive function?
- 2). Can the input be correctly accepted at the connector? Is the outcome correct?
- 3). Is there any data structure mistake or external information visiting mistake?
- 4). Can the performance satisfy requirements?

5.2 Physical Test

Physical test includes the analysis of fatigue, intension, electromagnetic effect and attrition degree. The contents of testing include long-time working test, pressure test, destructive test and so on. For example, we need to detect equipments such as airscrew, collector, temperature probe and valve, which are involved in the test of this airborne software.

6 Conclusion

The paper expounds an effective way of software test, by combining with the characteristics of embedded software and the technology of virtual prototype.

First, the simulation test based on virtual prototype has advantages, such as shorter construction cycle, safer test and repeatability of test.

Second, compared with physical simulation test, the simulation test based on virtual prototype can save manufacturing expenses of real outskirts equipments.

At last, the software code doesn't need to make any modification. And it can finish the invasive test of software to be tested and discover problems which is not easily found in hardware test technology.

References

1. Pengcheng, F.U., Qibin, D., Junbin, D., et al.: Studies on Software Structure of Military Universal Automatic Test System. In: 7th International Symposium on Testing, Beijing (2007)
2. Zhen, L., Bin, L., Ning, M., et al.: Formal Testing Applied in Embedded Software. In: 2009 8th International Conference on Reliability, Maintainability and Safety, Chengdu (2009)
3. Libya, J., Tiening, W., Ronghui, W.: Study on Military Embedded Software Testing Based on Concurrent Engineering. In: 8th International Symposium on Test and Measurement, Chongqing (2009)
4. Xinlei, Z., Jia, L., Xing, S.: The Test and Estimation Method of Software Reliability Based on State Analysis. In: 8th International Conference on Reliability, Maintainability and Safety, Chengdu (2009)
5. Jiang, C., Liu, B., Yin, Y., et al.: Study on Real-Time Test Script in Automated Test Equipment. In: 2009 8th International Conference on Reliability, Maintainability and Safety, Chengdu (2009)
6. Jianping Cai, C.: Practical Technology of Embedded Software Testing (嵌入式软件测试实用技术). Tsinghua University Press, Beijing (2010)

Retraction: Deconstructing the Ethernet

Zhong Chen and YongNing Guo

Department of Mathematics and Computer Science
Fuqing Branch of Fujian Normal University, Fuqing 350300, China
{czcat, guoyn}@163.com

Several conference proceedings have been infiltrated by fake submissions generated by the SCIGen computer program. Due to the fictional content the chapter “Deconstructing the Ethernet” by “Zhong Chen and YongNing Guo” has been retracted by the publisher. Measures are being taken to avoid similar breaches in the future.

Comment on a Research and Analysis Four-Prime RSA

Yongning Guo and Chenglian Liu*

Department of Mathematics and Computer Science
Fuqing Branch of Fujian Normal University, Fuqing 350300, China
guoyn@163.com, chenglian.liu@gmail.com

Abstract. In 2009, Zhang and Liu proposed a four-prime scheme which it based on RSA algorithm. However, for this scheme, it has some error. We will point out this situations.

Keywords: RSA Algorithm, Chinese Remainder Theorem, Four Prime Number.

1 Introduction

Fujioko et al. [2] presented an electronic cash scheme using a modulus $N = p^2q$, Okamoto and Uchiyama [4] designed a public-key scheme which it proved as secure as factoring a modulus $N = p^2q$. In 1998, Takagi [6] introduced a fast CRT-RSA variant using modulo of the form $N = p^kq$. Later, Boneh et al. [1] showed that schemes with modulo of the form $N = p^r q$ are more susceptible to attacks that leak bits of p than the original RSA-scheme. May [3] also proposed a secret exponent attacks on RSA-type scheme with moduli $N = p^r q$. In 2009, Zhang and Liu [7] proposed a novel scheme four-prime modulus which it also based on RSA algorithm. However, for this scheme, it has some incorrect. We simply introduce the Zhang-Liu scheme in section 2, and section 3 is our comment. The conclusion draws in final section.

2 Review of Zhang-Liu Scheme

2.1 The RSA Algorithm Description

RSA [5] algorithm describe following:

Step 1. Generate two large prime number p and q , the p and q did not published.

Step 2. Compute n where

$$n = pq. \quad (1)$$

Step 3. Randomly select an integer e , where

$$(e, \phi(n)) = 1. \quad (2)$$

Step 4. Compute an integer d , where

$$cd \equiv 1 \pmod{\phi(n)}. \quad (3)$$

It exist an integer d , thus e relatively prime with $\phi(n)$.

* Corresponding author.

Step 5. Publish n and e , but d must keep secret.

If n is enough huge, even known n and e , it is very difficult to factor p and q , thus the secret d can not be derived.

2.2 The RSA Algorithm Proof

Proof. 1). If $(m, n) = 1$, then $(m, p) = 1$ and $(m, q) = 1$, from Step 4, it exists a integer b where

$$(e \cdot d) = 1 + b(p-1)(q-1). \quad (4)$$

$$\begin{aligned} c^d \pmod{n} &\equiv m^{ed} \pmod{n} \\ &\equiv m^{1+b(p-1)(q-1)} \pmod{n} \\ &\equiv m \cdot (m^{(p-1)})^{b(q-1)} \pmod{n} \\ &\equiv m \pmod{p} \\ &\equiv m \pmod{q}. \end{aligned} \quad (5)$$

If

$$m^{p-1} \equiv 1 \pmod{p} \quad (6)$$

and

$$m^{q-1} \equiv 1 \pmod{q}, \quad (7)$$

then $(p, q) = 1$, therefore

$$c^d \equiv m \pmod{n}. \quad (8)$$

2). If $(m, n) \neq 1$, because $m < n$, to set $p|m$ and $q \nmid m$ (or $q|m$ and $p \nmid m$), then

$$c^d - m \equiv m^{ed} - m \equiv m(m^{b(p-1)(q-1)} - 1) \equiv m(1 - 1) \equiv 0 \pmod{n}. \quad (9)$$

Because $q|(m^{ed} - m)$, from $p|m$ obtained $p|(m^{ed} - m)$, let $n|(m^{ed} - m)$ based on $c^d \equiv m \pmod{n}$. Similarly proof if $q|m$ and $p \nmid m$ still has $c^d \equiv m \pmod{n}$.

2.3 The Zhang-Liu Scheme

In the traditional RSA encryption algorithm based on the four prime numbers for RSA encryption and decryption algorithm is still valid. The four prime numbers algorithm is describe as follow:

Notation:

p, q, r, s : denote prime numbers.

e and d : denote a public and provate key.

m : denote message or plain text.

c : denote cipher text.

$E(\cdot)$ and $D(\cdot)$: denote an encryption and a decryption function.

Step 1. Randomly choose four varieties prime numbers p, q, r, s to computer

$$n = p \cdot q \cdot r \cdot s \tag{10}$$

and

$$\phi(n) = (p - 1) \cdot (q - 1) \cdot (r - 1) \cdot (s - 1). \tag{11}$$

The $p, q, r, s, \phi(n)$ do not publish and n can be published.

Step 2. Compute the public key e where $0 < e < x$ and

$$\gcd(e, \phi(n)) = 1. \tag{12}$$

Step 3. Compute the private key d where

$$ed \equiv 1 \pmod{\phi(n)}. \tag{13}$$

The encryption and decryption process are the same with the traditional two-primes RSA encryption/decryption algorithms.

Encryption

$$c \equiv E(m) \equiv m^e \pmod{n}. \tag{14}$$

Decryption

$$m \equiv D(c) \equiv c^d \pmod{n}. \tag{15}$$

2.4 The Zhang-Liu’s Algorithm Proof

Assume plain text is M , ciphertext is c , private key pair (d, n) and public key pair (e, n) .

$$D(c) \equiv c^d \pmod{n} \equiv (m^e)^d \pmod{n} \equiv m^{ed} \pmod{n}. \tag{16}$$

However

$$ed \equiv 1 \pmod{\phi(n)}, \tag{17}$$

therefore $ed \equiv 1 + k\phi(n)$ where k is a positive integer constant. To imply equation (17) and obtains

$$D(c) \equiv m^{1+\phi(n)} \pmod{n} \rightarrow m^{1+\phi(n)} \equiv m \pmod{n}. \tag{18}$$

1) If $(m, n) = 1$, according to Euler theorem implies:

$$m^{\phi(n)} \equiv 1 \pmod{n}. \tag{19}$$

Thus,

$$m^{1+\phi(n)} \equiv m \pmod{n}. \tag{20}$$

2) If $(m, n) \neq 1$, but $n = pqrs$, so (m, n) contained some factors of (p, q, r, s) or (pq, pr, ps, qr) or (pqr, prs, qrs) . We will described above situation following.

3) If $(m, n) = p$, and $m = cp$, where $1 \leq c \leq qrs$, to obtain from Euler theorem:

$$m^{\phi(q)} \equiv 1 \pmod{q}, m^{\phi(r)} \equiv 1 \pmod{r}, m^{\phi(s)} \equiv 1 \pmod{s}. \tag{21}$$

For any k matches

$$m^{k(q-1)} \equiv m^{k(q-1)(r-1)(s-1)(p-1)} \equiv 1 \pmod{q}. \tag{22}$$

to get

$$m^{k(q-1)(r-1)(s-1)(p-1)} \equiv 1 + hq, \tag{23}$$

then

$$m^{k\phi(n)} \equiv 1 + hq, \tag{24}$$

because $m = cp$, it therefore

$$m^{k\phi(n)+1} \equiv m + cphq. \tag{25}$$

The parameter p, q, r, s are prime numbers, so

$$\gcd(p, qrs) = 1, \gcd(c, qrs) = 1, \gcd(m, qrs) = 1. \tag{26}$$

then

$$m^{\phi(qrs)} \equiv 1 \pmod{qrs}. \tag{27}$$

4) If $(m, n) = pq$, assume $m = cpq$, the parameters p, q, r, s are coprime each others, so $(pq, rs) = 1$, that exist

$$\gcd(c, rs) = 1, \gcd(cpq, rs) = 1, \gcd(m, rs) = 1. \tag{28}$$

to obtain from Euler theorem:

Proof.

$$m^{\phi(rs)} \equiv (m^{\phi(rs)})^{k\phi(p)\phi(q)} \equiv m^{k\phi(n)} \equiv 1 \pmod{rs}. \tag{29}$$

Because $m = cpq$, then $m^{k\phi(n)} + 1 \equiv m + hrscpq$, promptly $m^{k\phi(n)+1} \equiv m \pmod{n}$. Same reason, (pr, ps, qr, qs) are setting up from above situations.

5) If $(m, n) = pqr$, assume $m = cpqr$, then the parameters p, q, r, s are relatively prime, because $(pqr, s) = 1$, promptly $(c, s) = 1, (cpqr, s) = 1$. It means $(m, s) = 1$, to obtain from Euler theorem: $m^{k\phi(n)} \equiv 1 \pmod{s}, \forall$ integer k exist:

$$m^{k\phi(s)} \equiv m^{k\phi(s-1)} \equiv m^{k\phi(s-1)(r-1)(p-1)} \equiv 1 \pmod{s}. \tag{30}$$

It has $m^{k\phi(n)} \equiv 1 + hs$, where $m = cpr$, and $m^{k\phi(n)+1} \equiv m + chpqr$, therefore $m^{k\phi(n)+1} \equiv m \pmod{n}$. From the above inference reasons that the four prime-numbers theory based on RSA algorithm is established.

2.5 Four-Prime Using Chinese Remainder Theorem

Using Chinese remainder theorem, this scheme can be transferring follow operations.

Step 1. Compute

$$c_p \equiv c \pmod{p}, c_q \equiv c \pmod{q}, c_r \equiv c \pmod{r}, c_s \equiv c \pmod{s}. \tag{31}$$

Step 2. Compute

$$d_p \equiv d \pmod{p-1}, d_q \equiv d \pmod{q-1}, d_r \equiv d \pmod{r-1}, d_s \equiv d \pmod{s-1}, \quad (32)$$

it exist

$$M_1 \equiv c_p^{d_p} \pmod{p}, M_2 \equiv c_q^{d_q} \pmod{q}, M_3 \equiv c_r^{d_r} \pmod{r}, M_4 \equiv c_s^{d_s} \pmod{s}. \quad (33)$$

Step 3. Compute

$$M \equiv (M_1(qrs)^{p-1} + M_2(prs)^{q-1} + M_3(pqs)^{r-s} + M_4(pqr)^{s-1}) \pmod{n}. \quad (34)$$

we can obtain the plaintext M .

Step 4. Assume $n = pqrs$, and

$$a_1 = M_1, a_2 = M_2, a_3 = M_3, a_4 = M_4. \quad (35)$$

According to Chinese remainder theorem, because $(M_j, m_j) = 1$, so it exist an integer y_j , where n get from one solution as:

$$M \equiv (M_1M_p y_p + M_2M_q y_q + M_3M_r y_r + M_4M_s y_s) \pmod{n}. \quad (36)$$

However,

$$M_p = \frac{n}{p} = qrs, M_q = \frac{n}{q} = prs, M_r = \frac{n}{r} = pqs, M_s = \frac{n}{s} = pqr, \\ M_p y_p \equiv 1 \pmod{p}, M_q y_q \equiv 1 \pmod{q}, M_r y_r \equiv 1 \pmod{r}, M_s y_s \equiv 1 \pmod{s}. \quad (37)$$

According to Fermat's little theorem, let p be a prime, and $A \not\equiv 0 \pmod{p}$ where $A^{p-1} \equiv 1 \pmod{p}$, so $A^{-1} \equiv A^{p-2} \pmod{p}$. We can rewrite the equation (36) as follow:

$$M \equiv (M_1 qrs (qrs)^{-1} \pmod{p} + M_2 prs (prs)^{-1} \pmod{q} + \\ M_3 pqs (pqs)^{-1} \pmod{r} + M_4 pqr (pqr)^{-1} \pmod{s}) \\ \equiv (M_1 qrs (qrs)^{p-2} \pmod{p} + M_2 prs (prs)^{p-2} \pmod{q} + \\ M_3 pqs (pqs)^{p-2} \pmod{r} + M_4 pqr (pqr)^{p-2} \pmod{s}) \\ \equiv (M_1 (qrs)^{p-1} \pmod{p} + M_2 (prs)^{p-1} \pmod{q} + \\ M_3 (pqs)^{p-1} \pmod{r} + M_4 (pqr)^{p-1} \pmod{s}) \pmod{n} \quad (38)$$

As known:

$$M \equiv M_1 \pmod{p}, M \equiv M_2 \pmod{q}, M \equiv M_3 \pmod{r}, M \equiv M_4 \pmod{s}. \quad (39)$$

It has exist:

$$M_1 \equiv M \pmod{p}, M_2 \equiv M \pmod{q}, M_3 \equiv M \pmod{r}, M_4 \equiv M \pmod{s}. \quad (40)$$

And

$$M_1 \equiv M \equiv (c^d) \pmod n \equiv (c^d) \pmod{pqrs} \equiv c^d \pmod p. \tag{41}$$

If integer $A \nmid p$, and

$$n \equiv m \pmod{p-1}, \tag{42}$$

then

$$A^n \equiv A^m \pmod{p-1} \tag{43}$$

to get

$$c^d \pmod p \equiv c^{d \pmod{p-1}} \pmod p. \tag{44}$$

Let

$$d_p \equiv d \pmod p, \tag{45}$$

it exist

$$M_1^{d_p} \pmod p. \tag{46}$$

Assume

$$c_p \equiv c \pmod p, \tag{47}$$

then

$$M_1 \equiv c^d \equiv c^{d \pmod{p-1}} \equiv c^{d_p} \equiv c_p^{d_p} \pmod p. \tag{48}$$

The same reasons,

$$M_2 \equiv c_q^{d_q} \pmod q, M_3 \equiv c_r^{d_r} \pmod r, M_4 \equiv c_s^{d_s} \pmod s. \tag{49}$$

From above inference formula, we knew the message M could be partition for some blocks of M_1, M_2, M_3 and M_4 so on.

3 Our Comment

In this section, we will point out some incorrect formula which it based on group theory and give a simple practical example later.

3.1 Erratum I

From equation (36) and (38), we knew

$$M \equiv (M_1M_p y_p + M_2M_q y_q + M_3M_r y_r + M_4M_s y_s)$$

From equation (37), it re-express

$$\begin{aligned} M_p &\equiv qrs \pmod p, M_q \equiv prs \pmod q, M_r \equiv pqs \pmod r, M_s \equiv pqr \pmod s, \\ M_p y_p &\equiv 1 \pmod p, M_q y_q \equiv 1 \pmod q, M_r y_r \equiv 1 \pmod r, M_s y_s \equiv 1 \pmod s. \end{aligned} \tag{50}$$

Because

$$\begin{aligned} M_p y_p &\equiv qrs \pmod p \cdot (qrs)^{-1} \pmod p \equiv qrs \pmod p \cdot (qrs)^{p-1} \pmod p \\ &\neq qrs(qrs)^{p-2} \pmod p \end{aligned} \tag{51}$$

For example:

$$\begin{aligned} 15 &= 3 \cdot 5 \equiv (3 \pmod 7) \cdot (3^5 \pmod 7), \\ \implies 3 \cdot 3^5 \pmod 7 &= 1 \neq 15. \end{aligned}$$

3.2 Erratum II

On the other hand, there is a typo from equation (43). The correction is

$$A^n \equiv A^m \pmod{p}. \quad (52)$$

4 Conclusion

As the erratum in Zhang-Liu's scheme, their results derived an error in the section 4 of original citation. The wrong result, is an incorrect assumption. Now that Zhang-Liu's scheme becomes unnecessary.

Acknowledgment The authors would like to thank the anonymous reviewers for their comments that help improve the manuscript. This work was partially supported by the Department of Education of Fujian Province funding under the type B funding number JB06197 and JB11263. This work is also partially supported by the Project of Fuqing Branch of Fujian Normal University of China under the contract number KY2010030.

References

1. Boneh, D., Durfee, G., Howgrave-Graham, N.: Factoring $n = p^q r$ for Large r . In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 787–787. Springer, Heidelberg (1999)
2. Fujioka, A., Okamoto, T., Miyaguchi, S.: Esign: An Efficient Digital Signature Implementation for Smart Cards. In: Davies, D.W. (ed.) EUROCRYPT 1991. LNCS, vol. 547, pp. 446–457. Springer, Heidelberg (1991)
3. May, A.: Secret Exponent Attacks on RSA-Type Schemes with Moduli $n = p^r q$. In: Bao, F., Deng, R., Zhou, J. (eds.) PKC 2004. LNCS, vol. 2947, pp. 218–230. Springer, Heidelberg (2004)
4. Okamoto, T., Uchiyama, S.: A New Public-Key Cryptosystem as Secure as Factoring. In: Nyberg, K. (ed.) EUROCRYPT 1998. LNCS, vol. 1403, pp. 308–318. Springer, Heidelberg (1998)
5. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. *Communications of ACM* 21, 120–126 (1978)
6. Takagi, T.: Fast RSA-Type Cryptosystem Modulo $p^k q$. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 318–326. Springer, Heidelberg (1998)
7. Zhang, H., Liu, F.: A research and analysis for four-prime RSA. *Microcomputer Information* 26(5-3), 29–31 (2010)

An Improved FastIMM Algorithm Based on α - β and α - β - γ Filters

Junchen Sha, Jianping Xing^{*}, Zhenliang Ma, Liang Gao,
Can Sun, and Juan Sun

School of information science and engineering, Shandong University,
250100 Jinan, China
shajunchen@gmail.com, xingjp@sdu.edu.cn

Abstract. In this paper, an Improved FastIMM algorithm is presented by optimizing the acceleration gain factor γ and tracking indices of the α - β and α - β - γ filters. In this algorithm, a better acceleration gain factor γ is selected to improve the FastIMM filter's performance in target tracking. Besides, to get the appropriate ranges for tracking indices of the α - β and α - β - γ filters, numerical simulations are carried out and the Root Mean Square Error (RMSE) of the position and velocity are analyzed. Simulation results show that the Improved FastIMM has much higher tracking accuracy while keeping the same computational burden.

Keywords: α - β filter, α - β - γ filter, IMM, tracking index, gain factor.

1 Introduction

The Interacting Multiple Model (IMM) filter is one of the most cost-effective filters that can estimate the state of a dynamic system with several behavior modes which can “switch” from one to another [1]. The α - β and α - β - γ filters are second order and third order steady state kalman filters respectively, which could reduce the computational complexity greatly. Dahmani Mohammed proposed a FastIMM algorithm based on the α - β and α - β - γ filters, which decreases the computational burden greatly while having a high tracking accuracy [2].

Dahmani Mohammed has proved the good performance of FastIMM in tracking a target with constant velocity or constant turn rate [2]. However, such FastIMM renders low accuracy when tracking a constant-acceleration target. To circumvent this problem one could think of adjusting the gain factors of the α - β and α - β - γ filters. Sklansky described the performance of the α - β filter [3]. Then, Benedict and Bordner proposed the first optimal relation between filter coefficients of the α - β filter [4]. Following his work, Simpson modified the optimization condition when acceleration terms are

^{*} Corresponding author.

included and the tracker becomes an “ α - β - γ ” tracker [5]. Kalata presents a procedure and derivation of a parameter called the tracking index, which implicitly defines the optimal tracking parameters in closed forms [6]. Based on this tracking index, Ekstrand and Gray proposed the optimal gain factors of the α - β and the α - β - γ filters [7,8] respectively. And the FastIMM filter proposed by Dahmani Mohammed uses the same gain factors α - β with [7,8], while employing different acceleration gain factor with $\gamma = \beta^2 / (2\alpha)$ that presented by [8].

Aiming at improving the FastIMM filter’s tracking accuracy especially in tracking a target with constant acceleration, we use $\gamma = \beta^2 / (2\alpha)$ as the gain factor of acceleration estimate, then we analyze and compare the RMSE of FastIMM filters with different tracking indices to get appropriate ranges of the tracking indices based on simulation. Finally, an Improved FastIMM with a better acceleration gain factor and tracking indices is presented.

2 α - β Filter and α - β - γ Filter

2.1 α - β Filter

α - β filter is the steady state of the second order Kalman filter. According to [2] and [7], this filter is defined as follows:

$$\hat{x}(k) = x_p(k) + \alpha(x_0(k) - x_p(k)) \tag{1}$$

$$\hat{v}(k) = \hat{x}(k-1) + \frac{\beta}{T}(x_0(k) - x_p(k)) \tag{2}$$

$$x_p(k+1) = \hat{x}(k) + T\hat{v}(k) \tag{3}$$

where $\hat{x}(k)$, $x_0(k)$, $x_p(k)$, $\hat{v}(k)$ are the coordinate of the target’s estimated position, measured position, predicted position, and estimated velocity at the kth scan, respectively. T is the sample interval, and α, β are fixed coefficients filter gain factors that can be expressed in one parameter, the so called tracking index [2,7]:

$$\lambda = \frac{\sigma_v T^2}{\sigma_w} \tag{4}$$

where σ_w is the standard deviation of the measurement noise and σ_v is the standard deviation of the system noise. And the α, β gain factors can be calculated by the following equations[7]:

$$\alpha = -(\lambda^2 + 8\lambda - (\lambda + 4)\sqrt{\lambda^2 + 8\lambda}) / 8, \quad \beta = (\lambda^2 + 4\lambda - \lambda\sqrt{\lambda^2 + 6\lambda}) / 4 \tag{5}$$

2.2 α - β - γ Filter

α - β - γ filter is the steady state of the third order Kalman filter. According to [2], this filter is defined as follows:

$$\hat{x}(k) = x_p(k) + \alpha(x_0(k) - x_p(k)) \quad (6)$$

$$\hat{v}(k) = \hat{x}(k-1) + \frac{\beta}{T}(x_0(k) - x_p(k)) \quad (7)$$

$$\hat{a}(k) = \hat{a}(k-1) + \frac{\gamma}{2T}(x_0(k) - x_p(k)) \quad (8)$$

$$x_p(k+1) = \hat{x}(k) + T\hat{v}(k) + 0.5T^2\hat{a}(k) \quad (9)$$

where $\hat{a}(k)$ represents the target's estimated acceleration, and the other quantities have been defined previously. The α, β, γ gain factors are also function of the tracking index defined in [2],

$$\alpha = 1 - s^2, \quad \beta = 2(1 - s)^2, \quad \gamma = 2\lambda s \quad (10)$$

where

$$s = z - \left(\frac{p}{3z}\right) - \frac{b}{3}, \quad (11)$$

$$z = -\sqrt[3]{-q - \sqrt{(q^2 + \frac{4p^2}{27})/2}} \quad (12)$$

in which

$$b = \frac{\lambda}{2} - 3, \quad p = c - \frac{b^2}{3} \quad (13)$$

with

$$c = \frac{\lambda}{2} + 3, \quad q = \frac{2b^3}{27} - \frac{bc}{3} - 1 \quad (14)$$

When γ is defined as $\gamma = \beta^2 / (2\alpha)$ [8], while α, β keeping the same expression as Eq. 10, the α - β - γ filter performs much better in velocity tracking of a constant-acceleration target and keeping same tracking accuracy in position.

On this condition, we improve the FastIMM by using following α, β, γ expressions:

$$\alpha = 1 - s^2, \quad \beta = 2(1 - s)^2, \quad \gamma = \beta^2 / (2\alpha) \quad (15)$$

The comparison between the Standard FastIMM presented by Dahmani Mohammed and the Improved FastIMM will be discussed in chapter 4.

Further details about the α - β - γ filter are included in [8].

3 FastIMM Algorithm

FastIMM algorithm is composed of the α - β and α - β - γ filter, which decreases the computational burden greatly. And the FastIMM algorithm steps are as follows [2]:

Step 1: Calculation of the normalization constant:

$$\bar{c}_j = \sum_{i=1}^r p_{ij} \mu_i(k-1), \quad j = 1, 2 \tag{16}$$

Step 2: Calculation of the probabilities for models mixing.

$$\mu_{i/j}(k-1 | k-1) = \frac{1}{c_j} p_{ij} \mu_i(k-1), \quad i, j = 1, 2 \tag{17}$$

Step 3: Mixing of the model conditioned estimates.

$$x^{\wedge 0j}(k-1 | k-1) = \sum_{i=1}^r x^{\wedge i}(k-1 | k-1) \mu_{i/j}(k-1 | k-1) \tag{18}$$

Step 4: Mode conditioned state estimation. Using the α - β and α - β - γ filters to estimate the target's state, respectively.

Step 5: Likelihood function computation.

$$\begin{aligned} \Lambda^j(k) &= P[z(k) | M_j(k), Z^{k-1}] \\ &= \frac{1}{\sqrt{|2\pi S^j(k)|}} \exp(-0.5(v^j(k))^T (S^j(k))^{-1} v^j(k) \end{aligned} \tag{19}$$

where $v(k) = z(k) - H^j(k) x^{\wedge 0j}(k-1 | k-1)$, $S_j = \frac{\sigma_w^2}{(1-\alpha_j)}$ (20)

Step 6: updating the models probabilities.

$$\mu_j(k) = P(M_j(k) | Z^k) = \frac{1}{c} \Lambda^j(k) \bar{c}_j \quad j = 1, 2 \tag{21}$$

with $c = \sum_{j=1}^r \Lambda^j(k) \bar{c}_j$ (22)

Step 7: Yielding the overall state estimate by combining the estimates.

$$\hat{x}(k|k) = \sum_{j=1}^r \hat{x}^j(k|k) \mu_j(k) \tag{23}$$

Further details of the FastIMM algorithm are included in [2].

4 Performance Analysis

4.1 Simulation Settings

The target trajectory is generated as follows: The target moves in different state for three periods. Starting from [10,10m], the target moves with constant velocity $v_x = 100m/s, v_y = 100m/s$ in a straight line from 0s to 70s. Then it maneuvers and turns left with $\omega = 0.005 \text{ rad/s}$ from 71s to 140s. Finally, it moves with constant acceleration $a_x = -2m/s^2, a_y = -2m/s^2$ from 141s to 210s. And the sample interval is $T=3s$, while the standard deviation of the measurement noise and the system noise is chosen to be $\sigma_w = 10m$ and $\sigma_v = 0.1m/s^2$ respectively, which generate the tracking index. In addition, the model transition probability matrix is

$$\pi_{ij} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$$

The position trajectory and the velocity trajectory of the target are shown in Figure 1.a and Figure 1.b, respectively.

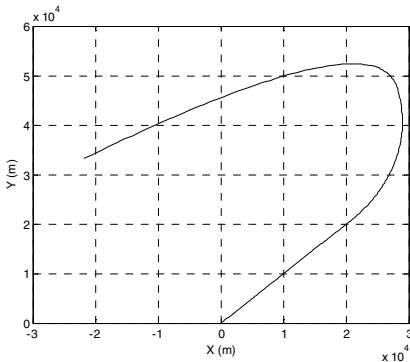


Fig. 1. a. True position trajectory

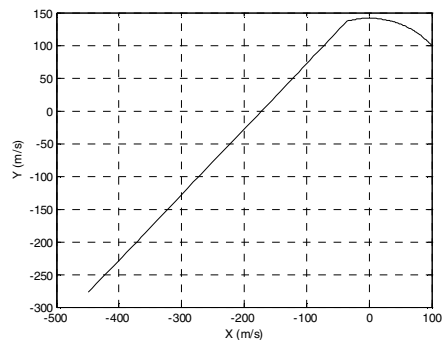


Fig. 1. b. True velocity trajectory

4.2 FastIMM with Different α - β Filter Tracking Indices

To compare and analyze the FastIMM performance of target tracking with different tracking indices, we propose a Compared FastIMM by changing the tracking index of its α - β - γ filter while keeping the α - β filter using the standard tracking index, which has been defined in Eq. 4. In the Compared FastIMM, λ_1 denotes the tracking index of the α - β filter whose gain factors have been defined in Eq. 5, while λ_2 denotes the tracking index of the α - β - γ filter whose gain factors are calculated by Eq. 14. Besides, the Standard FastIMM employs the standard tracking index λ as its sole tracking index.

Figure 2 and Figure 3 show the Comparison in position and velocity between the Standard FastIMM and Compared FastIMM respectively based on 200 Monte Carlo runs. And in the Compared FastIMM, $\lambda_2 = 0.5\lambda, 2\lambda, 5\lambda, 10\lambda$ respectively, while $\lambda_1 = \lambda$.

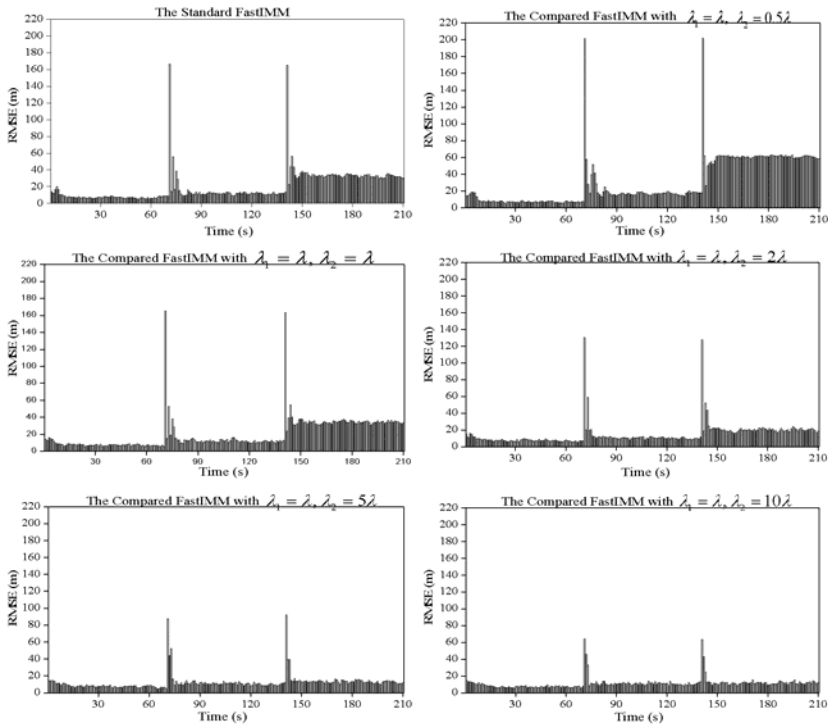


Fig. 2. RMSE Comparison in position between the Standard FastIMM and Compared FastIMM

According to Figure 2, when tracking a constant-acceleration target's position, the Compared FastIMM deteriorates badly when $\lambda_2 = 0.5\lambda$, but as the growing of λ_2 , both the average RMSE and the convergence rate of the Compared FastIMM improve greatly compared with the Standard FastIMM.

According to Figure 3, when tracking a constant-acceleration target's velocity, as the growing of λ_2 , the average RMSE of the Compared FastIMM increases slightly, while the Compared FastIMM having a bigger error when the mode of the target changes. Summing up the above, $5\sim 10\lambda$ is an appropriate range for the α - β - γ filter tracking index in the FastIMM algorithm.

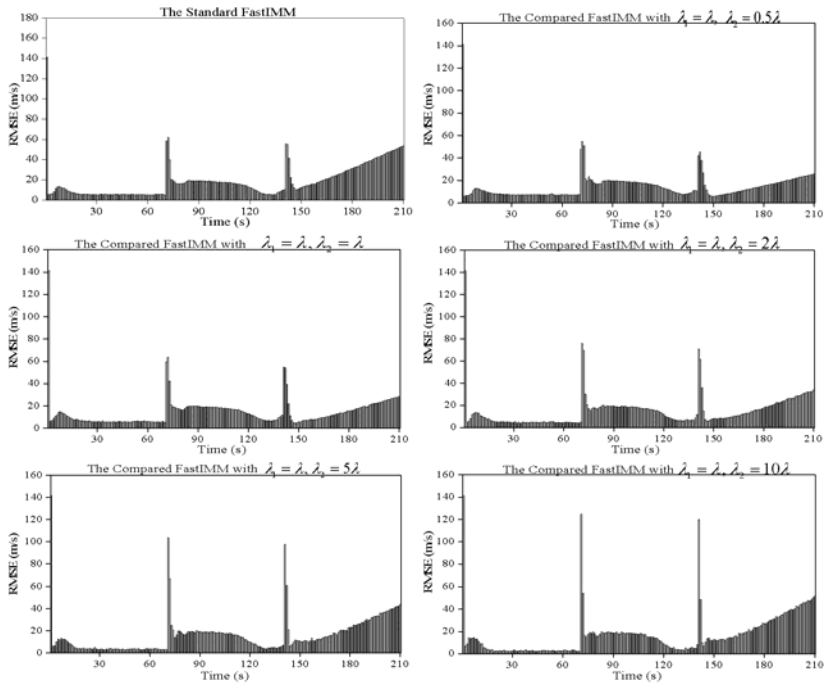


Fig. 3. RMSE Comparison in position between the Standard FastIMM and Compared FastIMM

4.3 FastIMM with Different α - β - γ Filter Tracking Index

Similarly with 4.2, we test another Compared FastIMM with different tracking index of its α - β filter while keep the α - β - γ filter using the standard tracking index.

Figure 4 and Figure 5 show the Comparison in position and velocity between the Standard FastIMM and Compared FastIMM respectively based on 200 Monte Carlo runs. And in the Compared FastIMM, $\lambda_1 = 0.5\lambda, 2\lambda, 5\lambda, 10\lambda$, respectively, while $\lambda_2 = \lambda$.

According to Figure 4, when tracking a constant-velocity target's position, the Compared FastIMM performs slightly better in convergence rate, which sacrifices some average tracking accuracy.

According to Figure 5, when tracking a constant-velocity target's velocity, the Compared FastIMM performs better and better in convergence rate while the average tracking accuracy becomes worse and worse as the growing of λ_1 ; when tracking a target with constant turn rate or constant accelerate, the bigger of λ_1 , the better the Compared FastIMM performs. The simulations show that $2\sim 5\lambda$ is an appropriate range for the α - β filter tracking index in the FastIMM algorithm.

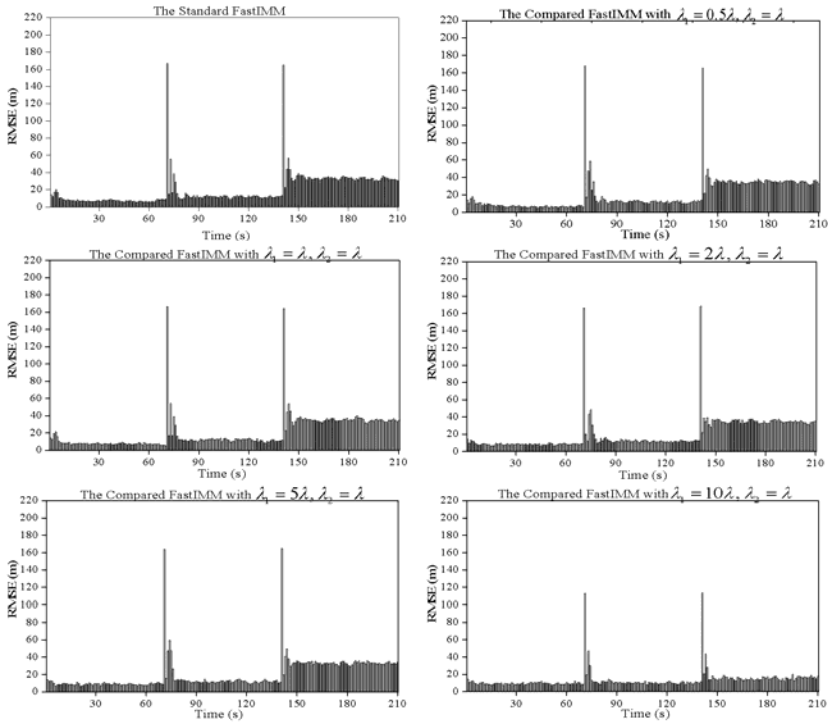


Fig. 4. RMSE Comparison in position between the Standard FastIMM and Compared FastIMM

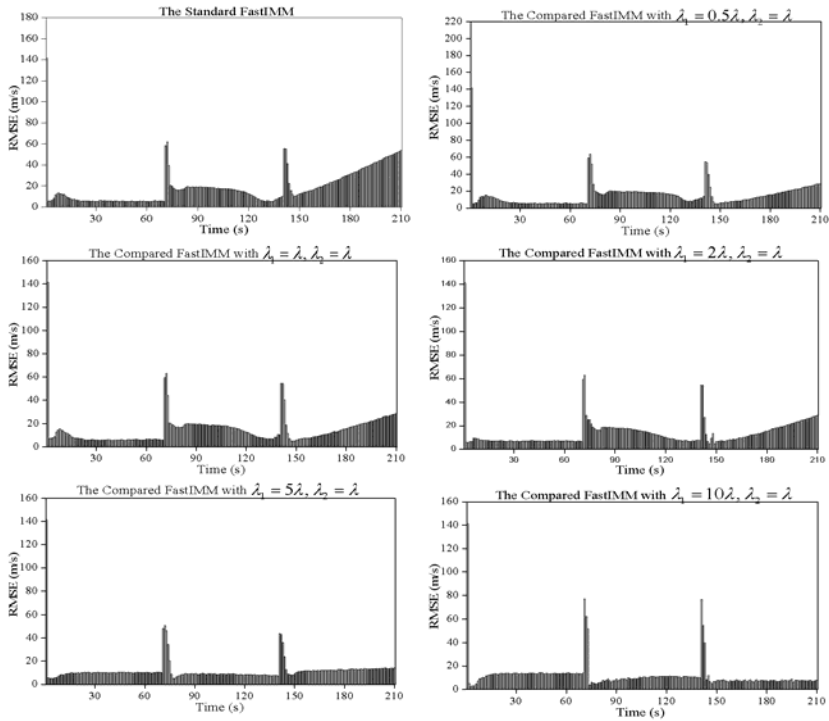


Fig. 5. RMSE Comparison in velocity between the Standard FastIMM and Compared FastIMM

4.4 An Improved FastIMM

Following section 4.1 and section 4.2, an Improved FastIMM is proposed with tracking indices $\lambda_1 = 5\lambda, \lambda_2 = 10\lambda$, while the α - β - γ filter gain factors α, β, γ are generated according to Eq. 14.

Figure 6 shows the Comparison in position and velocity between the Standard FastIMM and Improved FastIMM based on 200 Monte Carlo runs.

From Figure 6, it can be obviously seen that the Improved FastIMM performs much better both in position and velocity tracking compared with the Standard FastIMM, especially in tracking a constant-acceleration target's position, or tracking the velocity of a target with constant turn rate.

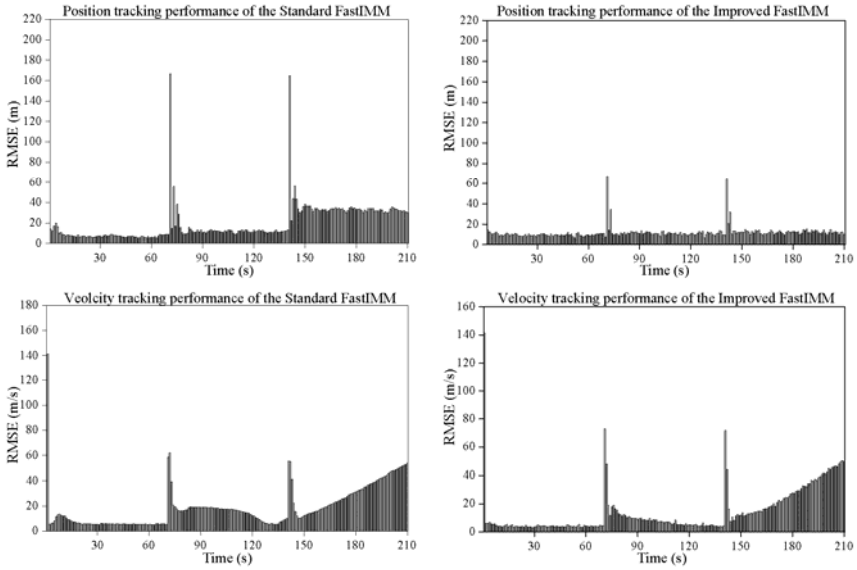


Fig. 6. RMSE Comparison in position and velocity between the Standard FastIMM and Improved FastIMM

5 Conclusion

In this paper, we proposed an Improved FastIMM that performs better in target tracking, especially in tracking a target with constant acceleration, which is promising in tracking maneuvering targets. In this algorithm, firstly, the performance of the α - β - γ filter is improved by selecting another expression of the acceleration gain factor γ ; secondly, we compared and analyzed the FastIMM with different tracking indices λ , and numerical simulations are carried out to get appropriate ranges for tracking indices of α - β and α - β - γ filter that compose the FastIMM filter; Finally we present an Improved FastIMM with a better acceleration gain factor and tracking indices. From the result of the simulation we can see that while keeping the same complexity, the Improved FastIMM improves the tracking accuracy about 62% in tracking a constant-acceleration target’s position and 42% in velocity tracking of a target with constant turn rate.

Acknowledgments. This work is funded by Independent Innovation Foundation of Shandong University (Grant No.2010JC015) and Shandong Province Natural Science Foundation (Grant No. ZR2011FM039).

References

1. Blom, H.A.P., Bar-Shalom, Y.: The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transactions on Automatic Control* 33(8), 780–783 (1988)
2. Dahmani, M., Keche, M., Ouamri, A., Meche, A.: A new IMM algorithm using fixed coefficients filters (FastIMM). *AEU - International Journal of Electronics and Communications* 64(12), 1123–1127 (2010)
3. Sklansky, J.: Optimizing the dynamic parameter of a track-while-scan system. RCA Laboratories, Princeton (1957)
4. Benedict, R.T., Bordner, G.W.: Synthesis of an optimal set of radar track-while scan smoothing equations. *IRE Transactions on Automatic Control* AC-7, 27–32 (1962)
5. Simpson, H.R.: Performance measures and optimization condition for a third-order sampled-data tracker. *IEEE Transactions on Automatic Control* AC-8, 182–183 (1963)
6. Kalata, P.R.: The tracking index: A generalized parameter for α - β and α - β - γ target trackers. *IEEE Transactions on Aerospace and Electronic Systems* AES-20, 174–182 (1984)
7. Ekstrand, B.: Poles and zeros of α - β and α - β - γ tracking filters. *IEE Proceedings: Control Theory and Applications* 148(5), 370–376 (2001)
8. Gray, J.E., Murray, W.: Rivation of an analytic expression for the tracking index for the alpha-beta-gamma filter. *IEEE Transactions on Aerospace and Electronic Systems* 29(3), 1064–1065 (1993)

An Approach on Feature Selection of Cascaded Support Vector Machines with Particle Swarm Optimization Algorithm

Ruihu Wang^{1,2}, Zhangping Hu³, Liang Chen³, and Jing Xiong³

¹ Department of Science and Technology,

Chongqing University of Arts and Sciences, 402160 Chongqing, China

² College of Computer Science, Chongqing University, 400040 Chongqing, China

³ School of Computer Science, Chongqing University of Arts and Sciences,
402160 Chongqing, China

ruihu.wang@yahoo.com.cn

Abstract. Support Vector Machine is one of the most frequently used new approaches for machine learning and pattern classification in many machine intelligent domains. Traditional SVM takes more computational run-time which impacts some real time application areas. Feature subset selection and parameters optimization are also two crucial issues to impact the classification and learning performance. In this paper cascaded SVMs are introduced to improve the classifier's ability. Moreover, Particle Swarm Optimization is a simple and easily to implemented swarm intelligent algorithm, which show promising performance in optimization. We proposed a novel approach which combines cascaded support vector machines optimized with PSO to enhance the classifier's performance.

Keywords: Cascaded, Support Vector Machines, Feature Selection, Parameters Optimization, Particle Swarm Optimization.

1 Introduction

Support Vector Machine was first proposed by Vapnik. It is build upon Structural Risk Minimisation (SRM), rather than Empirical Risk Minimisation (ERM). SVM was developed as an effective statistical learning method for pattern classification problem and has been used in a wide range of applications including biometric feature recognition (face recognition, facial expression recognition, and Iris data classification, etc.)

In the case where a linear discrimination boundary is unsuitable to solve the problem, the SVM can construct a map from the input vector into a high dimensional feature space by the use of reproducing kernel functions. Among alternative mapping are polynomials, radius basis functions and sigmoid functions [1].

Nonlinear SVMs utilize a set of support vectors to define a margin boundary between classes. SVM provides a generic capability that fits the hyperplane surface to the training data using a kernel function.

However, the computation is critically dependent upon the number of training patterns and to provide a good data distribution for a high dimensional problem will generally require a large training support vectors. Usually they are slower than neural networks as the run-time complexity is proportional to the number of support vectors. There has been a fair amount of research on methods for reducing the run-time complexity of SVMs. Many researchers have proposed a new architecture of SVMs, namely cascaded SVMs to solve this problem. Their methods achieved significant speed-ups over traditional SVM-based approaches without reducing detection rate too much and the hierarchical architecture of the detector also reduces the complexity of training of the nonlinear SVM classifier.

Generally, there are two problems we are faced to and have to solve when using SVM : the first is how to make the optimal input feature subset for SVM, and the other is to set the best kernel parameters [2].

To design a SVM, one must choose a kernel function, set the kernel parameters such as the gamma σ for the radius basis function kernel, and determine a soft margin constant C (penalty parameters). As suggested by Cheng Lung Huang, Chieh Jen Wang [2], feature subset choice influences the appropriate kernel parameters and vice versa, therefore, obtaining the optimal feature subset and SVM parameters should be dealt with at the same time. They proposed a Genetic Algorithm based for feature selection and parameters optimization simultaneously in an evolutionary way. Their experimental results showed that GA-based approach has better accuracy performance with fewer features than Grid algorithm.

Li-Juan Tang, Yan-Ping Zhou et.al [3] introduced Particle Swarm Optimization to make it possible for synergetic optimization of all parameters including kernel centers and kernel widths as well as linear model coefficients which are also essential for a quick convergence. Farid Melgani and Yakoub Bazi [4] proposed a novel classification system based on particle swarm optimization (PSO) to improve the generalization performance of the SVM classifier. Their PSO-SVM classification system has achieved substantial improvements in terms of classification accuracy. U Paquet and AP Engelbrecht suggested that training a support vector machine requires solving a constrained quadratic programming problem. Meanwhile linear particle swarm optimization is intuitive and simple to implement. They presented an alternative method to current numeric SVM for training. In their experiment, accurate and scalable training results were shown on the MNIST dataset, with the PSO algorithm finding fewer support vectors and better scalability than other approaches [5].

X.C.Guo and J.H.Yang, et.al. presented a novel hyper-parameter selection method for LS-SVMs based on the particle swarm optimization. The proposed method does not need any priori knowledge on the analytic property of the generalization performance measure and can be used to determine multiple hyper-parameters at the same time. Their experimental results show that the best or quasi-best test performance could be obtained by using the scaling radial basis kernel function and RBF kernel function, respectively [6].

In the literature mentioned above, some researchers have made satisfied classification performance with SVM classifier which work with parameters selected by PSO algorithm. However SVM has its intrinsic drawback as a classifier. It is difficult to be implemented as a constrained quadratic programming problem involved

in needs to be solved. The quadratic programming is rather slow and computationally expensive in order to converge to a solution [7].

Sami Romdhani and Philip Torr, et.al. [8] proposed a cascaded support vector machine expansion to detect face efficiently. The full support vectors expansion is only evaluated on the face-like parts of the image, while the largest part of typical images is classified using a single expansion vector. The cascaded evaluation offers a thirty-fold speed-up over an evaluation using the full set of reduced set vectors which itself already is thirty times faster than classification using all the support vectors. X. Ding trained cascaded SVM to detect head-shoulder using histograms of oriented gradients [9]. Seo Sang-Wook presented a reinforcement learning method with cascaded Support Vector Machine based on structural risk minimization and distributed genetic algorithms for behavior learning and evolution of collective autonomous mobile robots [10]. In this paper, we propose a novel approach on how to obtained the optimized parameter for cascaded SVMs using Particle Swarm Optimization.

The remaining parts of this paper are organized as follows: first some fundamental knowledge about support vector machines and particle swarm optimization are reviewed. Then we introduce a novel architecture of cascaded Support vector machines. Furthermore an approach of feature selection and parameters optimization for classifier training and classification based on cascaded SVMs and PSO is presented. We will implement the new combined method to facial expression recognition and red blood cell classification. Finally some discussion and conclusion are made in the end.

2 Support Vector Machine

At first, we will restrict the classification problem to the consideration of the two class problem in terms of simplicity. The classification problem can be regarded as separate two classes by a function which is induced from available examples. Consider the problem of separating the set of training vectors belonging to two separate classes,

$$T = \{x_1, y_1), \dots (x_l, y_l)\} \in (\mathcal{X} \times \{-1, 1\})^l \tag{1}$$

where

$$x_i \in \mathcal{X} = \mathfrak{R}^n, y_i \in \{-1, 1\}, i = 1, \dots, l. \tag{2}$$

If $\exists \omega \in \mathfrak{R}^n, b \in \mathfrak{R}, \varepsilon > 0$, such that for each label i and $y_i = 1$, $(\omega \cdot x_i) + b \geq \varepsilon$ holds true, whereas $(\omega \cdot x_i) + b \leq -\varepsilon$ for $y_i = -1$. The points x which lie on the hyperplane satisfy $\omega \cdot x + b = 0$,where ω is normal to the hyperplane, $|b| / \|\omega\|$ is the perpendicular distance from the hyperplane to the origin, and $\|\omega\|$ if the Euclidean norm of ω . The support vector algorithm aims to find the separating hyperplane with largest margin. This leads to the optimization problem of variables ω and b ,

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 \tag{3}$$

such that $y_i((\omega \cdot x_i) + b) \geq 1, \quad i = 1, \dots, l.$ (4)

The dual problem is

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j (x_i \cdot x_j) - \sum_{j=1}^l \alpha_j, \tag{5}$$

such that

$$\begin{aligned} \sum_{i=1}^l y_i \alpha_i &= 0, \\ \alpha_i &\geq 0, i = 1, \dots, l. \end{aligned} \tag{6}$$

If $\alpha^* = (\alpha_1^*, \dots, \alpha_l^*)^T$ is a solution of (2.5)~(2.6), then

$$\omega^* = \sum_{i=1}^l y_i \alpha_i^* x_i, b^* = y_j - \sum_{i=1}^l y_i \alpha_i^* (x_i \cdot x_j), \forall j \in \{j \mid \alpha_j^* > 0\}$$

(ω^*, b^*) is the optimized solution of (2.3)~(2.4).

2.1 Nonlinear Support Vector Machines

In the case where a linear boundary is inappropriate the SVM can map the input vector into a high dimensional feature space. Among acceptable mappings are polynomials, radial basis functions and certain sigmoid functions. Assume that there exists a mapping

$$x \in \mathfrak{X}^l \rightarrow z \in \mathfrak{X}^k \tag{7}$$

from the input feature space into a k -dimensional space, where the classes can be separately by a hyperplane satisfactorily.

According to **Mercer's** Theorem: given a mapping ϕ , the inner product has an equivalent representation:

$$\sum_r \phi_r(x) \phi_r(z) = K(x, z) \tag{8}$$

where $\phi_r(x)$ is the r -component of the mapping $\phi(x)$ of \mathbf{x} , and $K(x, z)$ is a symmetric function satisfying the following operation

$$\int K(x, z) g(x) g(z) dx dz \geq 0 \tag{9}$$

for any $g(x), \quad x \in \mathfrak{X}^l$ such that

$$\int g(x)^2 dx < +\infty \tag{10}$$

Now the optimization problem of Equation (2.5) becomes,

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(x_i \cdot x_j) - \sum_{j=1}^l \alpha_j, \quad (11)$$

subject to

$$\begin{aligned} \sum_{i=1}^l y_i \alpha_i &= 0, \\ C \geq \alpha_i \geq 0, i &= 1, \dots, l. \end{aligned} \quad (12)$$

where $K(x_i \cdot x_j)$ is the kernel function performing the non-linear mapping into feature space. One of the most common used kernel function is Gaussian Radius Basis Function which is with the form,

$$K(x, z) = \exp\left(-\frac{\|x - z\|^2}{\sigma^2}\right) \quad (13)$$

2.2 Feature Selection and Parameters Setting on SVM

According to Yang and Honvar, the choice of features used to represent patterns that are presented to a classifier affects several pattern recognition properties, including the accuracy of the learned classification algorithm, the time needed for learning a classification function, the number of examples needed for learning, and the cost associated with the features [11]. In addition to the feature selection, Cheng Lung Huang and Chieh Jen Wang suggested that proper parameters setting can also improve the SVM classification accuracy. The parameters that should be optimized include penalty parameter C and the kernel function parameters such as the sigma (σ) for the RBF kernel. In order to achieve optimal feature subset selection and SVM parameters, Hsu and Lin proposed a Grid algorithm to find the best C and sigma for RBF kernel function. However, this method's computational complexity is expensive and does not perform well [12].

Genetic algorithm is an another alternative tool, which has the potential to generate both the optimal feature subset and SVM parameters at the same time. Huang and Wang conducted some experiments on UCI database using GA-based approach. Their result has better accuracy performance with fewer features than grid algorithm [2].

Compared to Genetic Algorithm, Particle Swarm Optimization has no evolution operators such as crossover and mutation. There are few parameters to adjust. It works well in a wide variety of applications with slight variations. As introduced above, PSO algorithm can also be used in feature selection and parameters optimization.

3 Particle Swarm Optimization

Particle Swarm Optimization was introduced firstly by James Kennedy and Russel Eberhart [13]. It is a population-based evolutionary computation search technique. In

PSO, each potential solution is assigned a randomized velocity, and the potential solutions, called particles, fly through the problem space by following the current optimum particle. Each particle keeps track of its coordinates in hyperspace which are associated with the best solution (fitness) it has achieved so far. This value is called *pbest*. Another “best” value is also tracked. The “global” version of the PSO keeps track of the overall best value, and its location, which is called *gbest* [14]. Each particle is treated as a point in a D-dimension space. The original PSO algorithm is described below:

$$\begin{aligned} v_{id}^{k+1} &= v_{id}^k + c_1 r_1 (p_{id} - z_{id}^k) + c_2 r_2 (p_{gd} - z_{id}^k) \\ z_{id}^{k+1} &= z_{id}^k + v_{id}^{k+1} \end{aligned} \tag{14}$$

The *i*th particle’s location vector is represented as $z_i = (z_{i1}, z_{i1}, \dots, z_{iD})$; the velocity is denoted by $v_i = (v_{i1}, v_{i1}, \dots, v_{iD})$. $p_i = (p_{i1}, p_{i1}, \dots, p_{iD})$ and $p_g = (p_{g1}, p_{g1}, \dots, p_{gD})$ are *pbest* and *gbest* respectively. r_1 and r_2 are two random numbers in the range [0,1].

Equation (3.1) describes the flying trajectory of a population of particles, how the velocity and the location of a particle is dynamically updated. Equation (3.1.1) consists of three parts. The first part is the momentum part. The velocity is changed by current value. The second part is the cognitive part which represents the particle’s learning capability from its own experience. The third part is the social part which represents the collaboration among all particles [15].

In order to improve the convergence performance of PSO algorithm, Shi and Eberhart proposed a modified particle swarm optimizer. An inertial weight w is brought into the original PSO algorithm. This w plays the role of balancing the global search and local search. It can be a positive constant or even a positive linear or nonlinear function of time [16].

$$\begin{aligned} v_{id}^{k+1} &= w * v_{id}^k + c_1 r_1 (p_{id} - z_{id}^k) + c_2 r_2 (p_{gd} - z_{id}^k) \\ z_{id}^{k+1} &= z_{id}^k + v_{id}^{k+1} \end{aligned} \tag{15}$$

Simulations have been conducted on this modified PSO to illustrate the impact of this parameter introduced. It was concluded that the PSO with the inertial weight in the range [0.9,1.2] on average will have a better performance. A time decreasing inertial weight can also bring in a significant improvement on the PSO performance, as shown in Equation (3.3):

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{iter_{\max}} * k \tag{16}$$

4 Feature Selection and Parameters Optimization Using PSO

In this paper, PSO is utilized to search the optimal solution of a RBFN-SVM by minimizing the cost function Φ as the fitness function. Each particle is encoded as a real string representing the kernel centers (z) and widths (σ) as well as linear model coefficients w and b . With the movement of the particles in the solution space, the

optimal solution with a minimum value of the cost function Φ will be obtained. Optimizing the kernel centers and widths and the weights of the SVM model relating the feature variables synergistically keeps the model from getting trapped into a local optima and improves the model performance [3]. There are two key factors to determine the optimized hyperparameters when using PSO. One is how to represent the hyper-parameter as the particle’s position, namely how to encode. The other problem is how to define the fitness value function which evaluate the goodness of a particle.

In this section, we describe the proposed SVM system for classification task. The aim of this system is to optimize the SVM classifier accuracy by detecting the subset of the best discriminative feature and solving the SVM model selection.

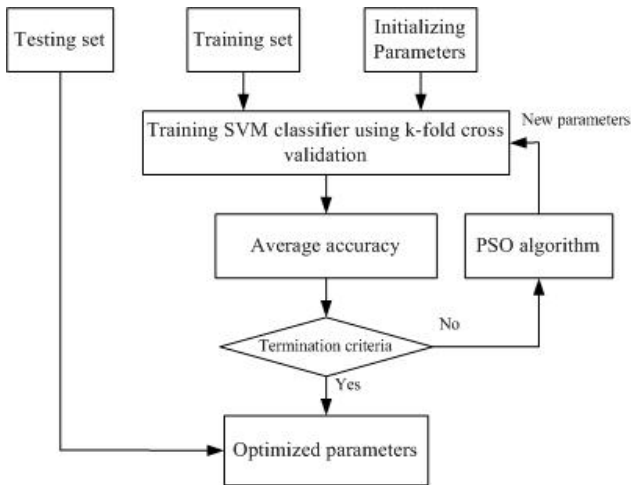


Fig. 1. PSO-SVM based feature selection and parameter optimization architecture scheme

4.1 PSO Setup [4]

The position $p_i \in \mathfrak{R}^{d+2}$ of each particle P_i from the swarm is regarded as a vector encoding.

- Feature subset f selection: f is a candidate subset of features, and F is a features set which consists of d available input features.

$$f = \{(x_1, \dots, x_d) \mid (x_1, \dots, x_d) \in F \subset \mathfrak{R}^d\}$$

- KBNF based SVM parameters optimization, includes C and σ .

The position vector of each particle can be represented as the other kind of form:

$$p_i = \{(x_{i1}, \dots, x_{id}, C_i, \sigma_i) \mid (x_{i1}, \dots, x_{id}, C_i, \sigma_i) \in \mathfrak{R}^{d+2}\}$$

Let $f(i)$ be the fitness function of the i th particle P_i . As suggested by Melgani, the choice of the fitness function is important on which PSO evaluates each candidate

solution p_i based for designing SVM classifier. They explored a simple SV count as a fitness criterion in the PSO optimization framework.

4.2 SVM Training and Classification with PSO

The pseudo code of the proposed method for SVM-PSO classification is given below.

Let N be the particle number of particle swarm.

1 Initialization

1.1 Initialize the population of N particles with random positions and velocities on D dimensions in the solution space.

1.2 Set the velocity vectors v_i ($i=1,2,\dots,N$) to **zero**.

1.3 For each position $p_i \in \mathcal{R}^{d+2}$ of the particle P_i ($i=1,2,\dots,N$) from the swarm, train the SVM classifier and compute the fitness function value.

2 Particle swarm search

2.1 Detect the best global position p_g in the swarm which showing the minimal value of the fitness function value over all explored trajectories.

2.2 Update the speed and position of each particle using Equation (3.2).

2.3 For each candidate particle p_i , train the SVM classifier and compute the fitness function $f(i)$.

2.4 Update the best position of each particle p_{bi} if its current position has a smaller fitness function.

3 Convergence

3.1 If the maximum iteration times have reached, then exit, else return 2.1.

4 SVM Training and Classification

4.1 Select the best global position p_g^* of the particle swarm and train the SVM with the detected feature subset mapped by p_g^* and modeled with the optimized parameters C and σ .

4.2 Make SVM classification based on the trained classifier.

5 PSO-Cascaded SVM Based Method

Support Vector Machine Classifier is remarkably robust machine learning method which is derived from statistical learning theory. It has gained much more popularity due to many attractive features, and promising empirical performance. Its intrinsic fundamental features

characterized by Structural Risk Minimization have been proved to be superior to traditional Empirical Risk Minimization principle. A set of support vectors are utilized to define a boundary between two classes by means of kernel function. However, SVM is usually slower than neural network. The reason is that the run-time complexity is proportional to the number of Support Vectors, i.e. to the number of training examples that the SVM algorithm utilizes in the expansion of the decision function. Many researchers have proposed cascaded SVMs architecture to solve the computational complexity problem. Among the cascaded structure, there are mainly two kinds of SVMs architecture. One is parallel SVMs, and the other is serial SVMs [17].

The parallel cascaded structure of SVM is more relied on hardware architecture of computer, which is developed based on decomposing the original complex problem into a number of independent simplified smaller problems, and in the end the partial results are combined into an ultimate output [18]. Compared to parallel SVMs, the serial structure cascaded SVMs is more feasible and easily to be implemented. Y. Ma and X. Ding proposed a cost-sensitive SVM classifier to detect face [19]. In many classification cases, the cost of False Negative is far more than False Positive. In their cost-sensitive SVMs, different cost are assigned to two types of misclassification to train the cascaded SVMs in different stage of the face detector. The optimization goal is given by

$$\min \left\{ \frac{1}{2} \|w\|^2 + C_p \sum_{y_i=1} \xi_i + C_n \sum_{y_i=-1} \xi_i \right\} \tag{17}$$

where C_p is the cost for face samples, and C_n is for non-face samples, usually $C_p > C_n$.

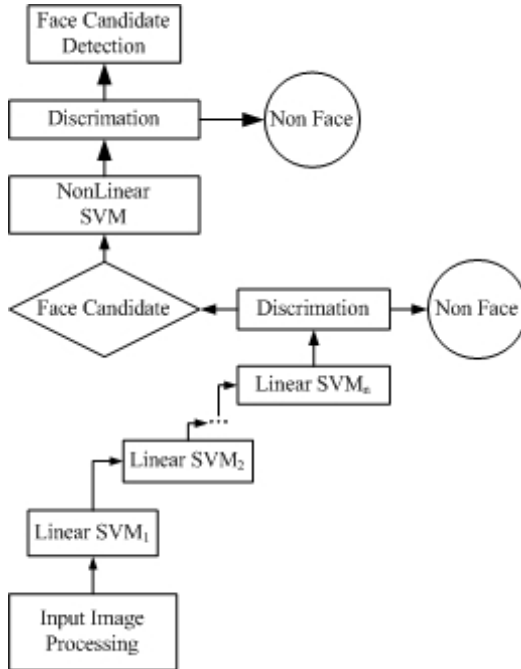


Fig. 2. Cascaded Architecture of Serial SVMs

We propose here a novel approach named PSO-Cascaded SVMs classification method to combine Cascaded SVMs with particle swarm optimization algorithm. As shown in Figure 2, there are some linear SVMs cascaded to form a serial structure at the front end and a nonlinear SVM at the end of this system.

We think of two optional strategies to introduce particle swarm optimization into the cascaded architecture of serial SVMs: one is integrate PSO into the serial linear SVMs and the other is integrate PSO into the end nonlinear SVM only. Regarding that the serial SVMs structure at the front end is more easily to construct and implement than nonlinear SVM, we take the later strategy. The nonlinear SVM classification is more complicated because it has to deal with feature subset selection and kernel function parameters optimization.

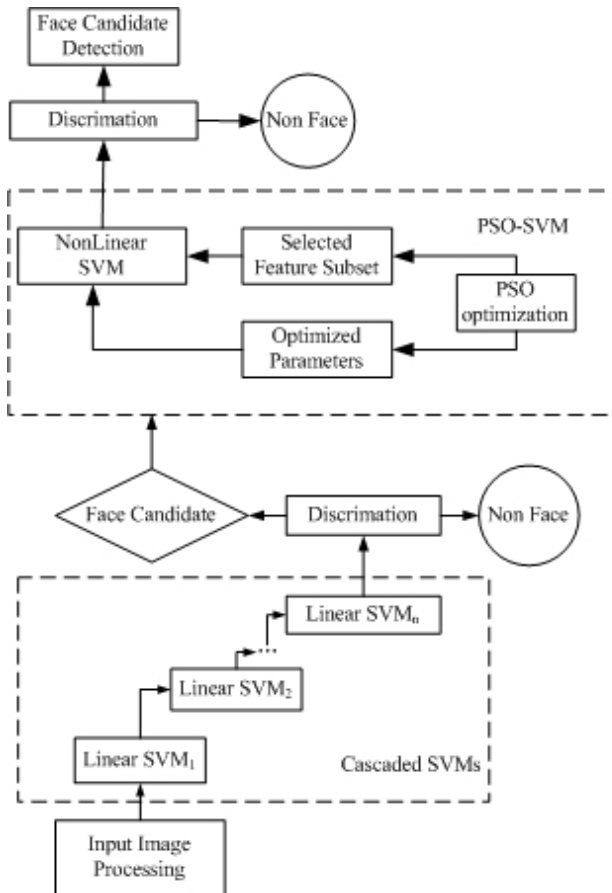


Fig. 3. PSO-Cascaded SVMs for Feature Selection and Parameters Optimization2.2 Formulas

6 Future Work and Discussion

In this paper we have introduced some fundamental knowledge of Support Vector Machine. As SVM is computational time expensive tool for classification, we have to deal with some optimization such as feature subset selection before training and classifying to improve its performance. Especially for nonlinear SVM, the selection of kernel function and parameters optimization are more important. Two major types of cascaded SVMs are briefly reviewed. The cascaded SVMs structure show up superior performance in pattern classification.

Further more, as a simple stochastic global optimization technique inspired by social behavior of bird flocking, PSO can be used into the cascaded SVMs to select feature subset and optimize parameter for kernel function. Two PSO-SVM strategies have been proposed in this paper for SVM to enhance its learning and classifying capability.

In our future work we would like to use this novel PSO-Cascaded SVMs approach to solve Red Blood Cell classification issue, which is a typical medical image computer aided diagnose. And also we prepare to test this approach for human emotion state recognition. In public surveillance system, affective computing reveals some potential action motivation. With the requirement of time sensitive and real time processing, our proposed combined method should be an ideal tool for spontaneous emotion recognition.

Acknowledgments. This work is sponsored by the Science and Technology Foundations of Chongqing Municipal Education Commission under Grant No. KJ091216, and Excellent Science and Technology Program for Overseas Studying Talents of Chongqing Municipal Human Resources and Social Security Bureau under Grant No. 09958023, and also by Key project of Science and Research Foundation of Chongqing University of Arts and Sciences under Grant No. Z2009JS07.

References

1. Gunn, S.R.: Support Vector Machines for Classification and Regression. Technical Report. School of Electronics and Computer Science, University of Southampton
2. Huang, C.L., Wang, C.J.: A GA-based feature selection and parameters optimization for support vector machines. *Expert Systems with Applications* 31, 231–240 (2006)
3. Tang, L.-J., Zhou, Y.-P., et al.: Radius Basis Function Network-Based Transform for a Nonlinear Support Vector Machine as Optimized by a Particle Swarm Optimization Algorithm with Application to QSAR Studies. *J. Chem. Inf. Model.* 47, 1438–1445 (2007)
4. Melgani, F., Bazi, Y.: Classification of Electrocardiogram Signals with Support Vector Machines and Particle Swarm Optimization. *IEEE Transaction on Information Technology in Biomedicine* 12(4) (September 2008)
5. Paquet, U., Englbrecht, A.P.: Training Support Vector Machines with Particle Swarms. In: *Proceedings of the International Joint Conference on Neural Network*, vol. 2, pp. 1593–1598 (2003)
6. Guo, X.C., Yang, J.H., et al.: A novel LS-SVMs hyper-parameter selection based on particle swarm optimization. *Neuro Computing* 71, 3211–3215 (2008)
7. Li, Z.: A Support Vector Machine training Algorithm based on Cascaded Structure. In: *Proceedings of the First International Conference on Innovative Computing, Information and Control* (2006)

8. Romdhani, S., Torr, P., et al.: Efficient Face Detection by a Cascaded Support Vector Machine Expansion. *Proceedings of the Royal Society* 460(2051), 3283–3297 (2004)
9. Ding, X., Xu, H., Cui, P., Sun, L.: A Cascaded SVM approach for head-shoulder detection using histogram of oriented gradient. In: *IEEE International Symposium of Circuits and Systems*, pp. 1791–1794 (2009)
10. Seo, S.-W., Yang, H.-C., Sim, K.-B.: Object tracking algorithm of Swarm Robot System for using Polygon based Q-learning and Cascade SVM. In: *IEEE International Symposium on Industrial Electronics*, 07 (2009)
11. Yang, J., Honavar, V.: Feature subset selection using a genetic algorithm. *IEEE Intelligent System*, 13(2), 44–49
12. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classification (2003), <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>
13. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. In: *Proceedings of IEEE International Conference on Neural Networks*, IEEE Service Center, Piscataway (1995)
14. Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. In: *The Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43 (1995)
15. Shi, Y.: Particle Swarm Optimization. *IEEE Neural Network Society* (February 2004)
16. Shi, Y., Eberhart, R.: A modified Particle Swarm Optimizer. In: *Proceedings of the IEEE Congress on Evolutionary Computation (CEC 1999)*, Piscataway NJ, pp. 69–73 (1999)
17. Wang, R., Fang, B., Hu, Z., Chen, L., Wang, W.: Cascaded SVMs in pattern Classification for Time-Sensitive Separating. In: *The International Symposium on Intelligent Information Technology and Security Informatics* (2010)
18. Yang, J.: An improved cascade SVM training algorithm with crossed feedbacks. In: *Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences* (2006)
19. Ma, Y., Ding, X.-Q.: Face Detection Based on Cost-Sensitive Support Vector Machines. In: Lee, S.-W., Verri, A. (eds.) *SVM 2002*. LNCS, vol. 2388, pp. 260–267. Springer, Heidelberg (2002)

Quantitative Analysis of Survivability Based on Intrusion Scenarios

Chengli Zhao^{1,2} and Zhiheng Yu³

¹ Department of Computer Science and Technology,
Jilin University, Jilin Changchun, China

² Department of Management Science and Information Engineering,
Jilin University of Finance and Economics, Jilin Changchun, China

³ Academy of Fineart, Northeast Normal University, Jilin Changchun, China
{zhaoc1069, yuzh496}@gmail.com

Abstract. Quantitative analysis of survivability is to use mathematical methods to evaluate the survival situation of the existing network quantitatively. It contributes to detect damage to the system in time and give the appropriate recommendations for improvement, so that the system can provide sustained and stable services. This paper presents a quantitative model based on intrusion scenarios to analysis the survivability of the information system. After constructing the attack trees, the intrusion scenarios are produced. Through calculating the degree of risk of the intrusion scenarios, the survivability of each critical service is determined. Furthermore, the survivability of the information system is measured effectively.

Keywords: Survivability, quantitative analysis, attack tree, intrusion scenario, information system.

1 Introduction

When the network systems are increasing scale and complexity, network attack techniques are also constantly developing and evolving. People have realized that the network information system can not be indestructible, and the system will inevitably be attacked. Traditional security technology has been unable to adapt to needs of the times, and the survivability of information systems has become the focus of the problem of security. The concept of the survivability was long, but unanimously accepted definition was given by Ellison in 1997. That is, the survivability is the capability of the system to provide critical services and timely recovery of all services under attack, failures and unexpected events [1].

Network survivability analysis provides a measure of overall survival situations to system administrators by real-time evaluation in the case of a variety of attacks and accidents. Survivability assessment analysis provides a scientific basis and criterion to use enhanced technology.

Currently, the quantitative assessment of the survivability is mainly in two lines of research. On the one hand, according to Ellison's definition of survivability, the overall

survivability of information systems can be reflected in a comprehensive assessment of all critical services, and the performance of each service can be evaluated through the key indicators [2,3]. On the other hand, because the survival information system should have the 3R features those are the resistance, the recognition and the recovery. Therefore, many researchers have studied the survivability from these three aspects. Starting from 3R features, Lin gave a quantitative formula [4]. Li proposed a general expansion and improvement from a mathematical point of view [5]. Zhang and Wang measured the system survivability from the attack graph model and the vulnerability grades respectively [6,7].

This paper proposes a computation model of survivability based on intrusion scenarios. Starting from the attack tree, determine the intrusion scenarios of each goal of the system by which the degree of risk of each goal can be evaluated. At last, we give the mathematic model to quantify the performance of the system's survivability.

2 Attack Tree and Intrusion Scenario

2.1 The Definition of Attack Tree

Attack tree model was proposed by Bruce in 1999 as a modeling approach to deal with the threat to the security of the system. The attack to the system was expressed in and/or trees. The root node of an attack tree represented a hazardous event to the system security. Such ways that the attackers can cause the hazardous events can be represented iteratively and incrementally in lower nodes. The leaf node represented a realization of the hazard events that may take atomic means. Typically, an information system has a collection or forest of relevant attack trees. A node of an attack trees can be decomposed as an and-decomposition if all the sub-goals must be achieved for the attack to succeed and it can be decomposed as an or-decomposition if any one of the sub-goals must be achieved to succeed for the attack. Attack trees can be represented in graph. Figure 1 and Figure 2 depict an and-decomposition and an or-decomposition separately.

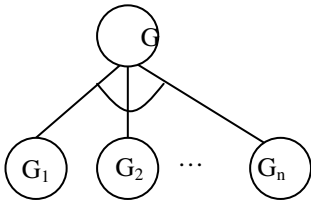


Fig. 1. A node of and-composition

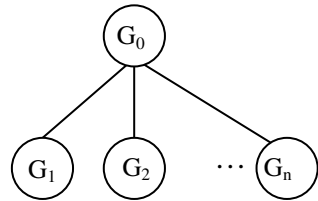


Fig. 2. A node of or-composition

2.2 Intrusion Scenario

Intrusion scenario is composed of the atom action set, by which the intruder can achieve his goal of the attack. Intrusion scenario can be obtained according to the corresponding attack tree. As the attack tree is broken down by any and/or combination, we can traverse a whole tree and generate some independent intrusion scenarios.

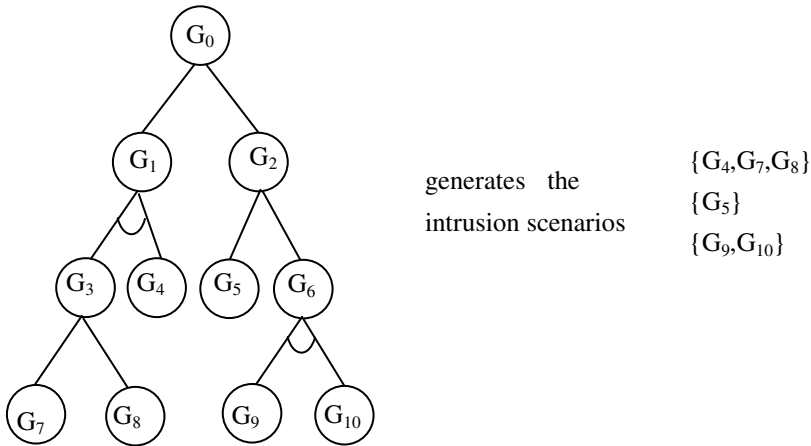


Fig. 3. An and/or attack tree

The attack tree showed in Figure 3 generates three independent intrusion scenarios, which are respectively $\{G_4, G_7, G_8\}$, $\{G_5\}$ and $\{G_9, G_{10}\}$. As long as all the atom targets in one of the independent intrusion scenarios are compromised, the ultimate goal will be successfully captured. That is, any set of $\{G_4, G_7, G_8\}$, $\{G_5\}$ and $\{G_9, G_{10}\}$ is achieved, you can succeed to attack the ultimate goal of G_0 .

3 The Probability of a Single Goal

In fact, attackers often attack the system by means of system vulnerabilities, and the degree of difficulty is different by use of different vulnerability to attack. Moreover, the system scanners are not 100% accurate, and there is certain credibility, so the leaf goal is a probability event influenced by a number of factors.

Suppose a intrusion scenario IS is composed with n atom leaf nodes, namely, $G_{IS} = \{G_1, G_2, \dots, G_n\}$, and the success probability of G_i is P_i , the succeed probability of the entire intrusion scenario IS can be calculated by the formula (1).

$$p_{IS} = \prod_{i=1}^n p_i \quad (1)$$

Suppose an attack tree AT is composed with m intrusion scenarios, each of which can succeed to attack the ultimate goal, the probability of an attack tree can be quantified by the formula (2).

$$p_{AT} = \min\left\{\sum_{j=1}^m p_m, 1\right\} \quad (2)$$

The easier the goal is compromised, the easier the system crashes and, the lower the survival capability of the system is. Therefore, the survivability of a goal can be defined as follows.

$$Sur_{GOAL} = 1 - p_{TT} \quad (3)$$

4 The Entire System Survivability

A system usually needs to provide a variety of services to users. These services can be divided into key services and secondary services. Even key services often have different weights. Each attack to a key service target may constitute an attack tree. A number of key service objectives form a forest. Thus, the survivability of the system is mainly determined by the performance of the forest's survival. According to the definition of survivability, each key service must be survival, so the entire survivability of the system can be calculated by the formula (4). Where w_i is the weight of the key service S_i , and δ is the threshold. Based on experience, $\delta = 0.4$.

$$Sur = \begin{cases} \sum_{i=1}^k w_i Sur_i & \forall i, Sur_i > \delta \\ 0 & else \end{cases} \quad (4)$$

Of course, when the overall survivability of the system is equivalent to 0, there must be some key services that can not survive. If so, based on the calculations in intrusion scenarios, the system administrator should identify the objectives that are the most likely to break and take prior security measures to enhance their survivability.

5 Conclusions

The research on survival system has been ongoing for many years, in which quantitative assessment has always been the focus of concern. Starting from the attack

tree, this paper proposed the use of attack scenarios for the survival assessment. By achieving the probability of various different attack scenarios, we calculated the survival probability of a single goal, and further, the overall survivability of the system can be quantified.

The probability of the leaf node in an attack tree was often carried out by the system vulnerabilities. Although the vulnerability was classified and graded simply in some existing articles, the probability of each level was lack of strong support. Therefore, how to quantify the degree of the system vulnerability in difficulty is the focus of future work. In addition, the simulation will be carried out and specific examples will be given to validate the proposed algorithm in the future.

Acknowledgement. This work was supported in part by Jilin Province Science and Technology Development Plan under Grant No. 20100173; Jilin Province Department of Education Social Science Research, "Eleventh Five-Year Plan" under Grant No. 2010407; Jilin University of Finance and Economics Science Research under Grant No. 2010018.

References

1. Ellison, R.J., Fisher, D.A., Linger, R.C., et al.: *Survivable Network Systems: an Emerging Discipline*. CMU/SEI-97-TR-013. Carnegie Mellon University, Pittsburgh (1997)
2. Zhao, G., Wang, H., Wang, J.: Study on Situation Evaluation for Network Survivability Based on Grey Relation Analysis. *Mini-Micro Systems* 27(10), 1861–1864 (2006)
3. Zhao, C., Liu, Y., Yu, Z.: Combining Grey Relation Analysis and TOPSIS for Evaluating the Survivability of Networked Information System. *Advanced Materials Research* 219–220, 210–213 (2011)
4. Lin, X., He, X.: Survivability Analysis and Implementation for Network Information System. *Com. Eng.* 31(24), 161–163 (2005)
5. Li, X.: Research on the Quantified Evaluation Method of Information System's Survivability. *J. Cha. Univ.* 24(5), 6–9 (2007)
6. Zhang, L., Wang, W., Guo, L., Yang, W., Yang, Y.: Survivability Computation of Information Systems Based on Intrusion Scenario. *Computer Engineering* 34(6), 137–139 (2008)
7. Wang, Y., Xian, M., Liu, J., Wang, G.: Study of network security evaluation based on attack graph model. *Journal on Communications* 28(3), 29–34 (2007)

Facial Expression Recognition Based on Local Binary Patterns and Least Squares Support Vector Machines

Xiaoming Zhao¹ and Shiqing Zhang²

¹ Department of Computer Science, Taizhou University,
318000 Taizhou, China

² School of Physics and Electronic Engineering, Taizhou University,
318000 Taizhou, China

{tzxyzxm, tzczsq}@163.com

Abstract. In this paper a new facial expression recognition method based on local binary patterns (LBP) and least squares support vector machines (LS-SVM) is proposed. LBP is adopted as facial representations for facial expression recognition since LBP tolerates against illumination changes and operates with its computational simplicity. After extracting LBP features, LS-SVM with radial basis function (RBF) kernel is employed for facial expression classification. The experimental results on the popular JAFFE facial expression database demonstrate that the recognition accuracy based on LBP and LS-SVM comes up to 78.57%.

Keywords: Facial expression recognition, Local binary patterns, Least squares support vector machines.

1 Introduction

Facial expression recognition plays a vital role in human emotion perception and social interaction and has increasingly attracted much interest in recent years [1]. Many methods have been employed for automatic facial expression recognition by means of various pattern recognition techniques. Generally, according to the facial feature representations used for facial expression analysis, these methods can be divided into two categories: geometry-based methods and appearance-based methods. Geometric features such as the shape and location of the eyes, mouth, nose, etc., are the mostly extracted features. Appearance-based methods directly work on the facial images to represent facial textures and are thus simple to implement. Two representative appearance-based methods are Gabor wavelets [2] and local binary patterns (LBP) [3], which have been successfully used to describe the local appearance of facial expressions. As shown in [4, 5], the simplicity of LBP histogram allows for very fast feature extraction, does not need complex analysis in extracting a large set of Gabor wavelet coefficients. Compared to high dimensionality of the Gabor representation, the LBP features with short vector length lies in a lower dimensional space, thus requires much less computational resource.

Recently, the support vector machine (SVM) [6] and its modified form, least squares support vector machine (LS-SVM) [7], have been successfully applied on

classification problems owing to their robustness and simplicity. Motivated by the lack of studying on the application of LS-SVM for facial expression recognition, in this paper we aim to illustrate the potential of LS-SVM in facial expression recognition. In order to testify the performance of LS-SVM for facial expression recognition, we perform facial expression recognition experiments on the popular JAFFE [8] facial expression database.

2 Local Binary Patterns

The original Local Binary Patterns (LBP) operator labels the pixels of an image by thresholding a 3×3 neighborhood of each pixel with the center value and considering the results as a binary code. The LBP code of the center pixel in the neighborhood is obtained by converting the binary code into a decimal one. Fig.1 gives an illustration for the basic LBP operator. Based on the operator, each pixel of an image is labeled with an LBP code. The 256-bin histogram of the labels contains the density of each label and can be used as a texture descriptor of the considered region.

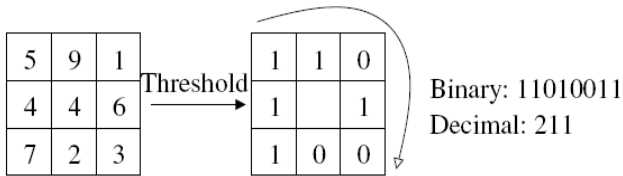


Fig. 1. The basic LBP operator

The procedure of LBP feature extraction is implemented as follows:

Firstly, a face image is divided into several non-overlapping blocks. Secondly, LBP histograms are computed for each block. Finally, the block LBP histograms are concatenated into a single vector. As a result, the face image is represented by the LBP code.

3 Least Squares Support Vector Machines

Given a training data points $\{x_i, y_i\}, i=1,2,\dots,l$ with input data x_i and output data y_i , in the feature space LS-SVM [7] takes the form

$$y = \text{sgn}(w^T \varphi(x) + b) \tag{1}$$

where the nonlinear mapping maps the input data into a high-dimensional feature space whose dimension can be infinite. To obtain a classifier, LS-SVM solves the following optimization problem:

$$\min \left\{ \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^l e_i^2 \right\} \tag{2}$$

subject to $y_i = w^T \varphi(x_i) + b + e_i$.

Its Wolfe dual problem is

$$\min \left\{ \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j \varphi(x_i)^T \varphi(x_j) + \sum_{i=1}^l \frac{\alpha_i^2}{2\gamma} - \sum_{i=1}^l \alpha_i y_i \right\} \tag{3}$$

subject to $\sum_{i=1}^l \alpha_i = 0$.

We can delete the equality constraint and obtain

$$\min \left\{ \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j \varphi(x_i)^T \varphi(x_j) - \sum_{i=1}^l \alpha_i y_i + \sum_{i=1}^l \frac{\alpha_i^2}{2\gamma} + b \sum_{i=1}^l \alpha_i \right\} \tag{4}$$

The form in (4) is often replaced with a so-called positive-definite kernel function $k(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$. There are several kernels that can be used in SVM models. These include linear, polynomial, radial basis function (RBF) and sigmoid function. Then we can get the resulting LS-SVM model for function estimation as follows:

$$f(x) = \text{sgn}(w^T \varphi(x) + b) = \text{sgn} \left(\sum_{i=1}^l a_i k(x, x_i) + b \right) \tag{5}$$

For multi-class problem, to construct a set of M binary classifiers, there are mainly four encoding strategies: one-versus-one, one-versus-all, minimal output coding (MOC) and error correcting output codes (ECOC). For the one-versus-one approach, classification is done by a max-wins voting strategy, in which every classifier assigns the instance to one of the two classes, then the vote for the assigned class is increased by one vote, and finally the class with most votes determines the instance classification. For the one-versus-all method, classification is done by a winner-takes-all strategy, in which the classifier with the highest output function assigns the class. Minimum output coding (MOC) uses M outputs to encode up to 2^M classes and thus have minimal M . Error correcting output codes (ECOC) use more bits than MOC, and allow for one or more misclassifications of the binary classifiers.

4 Facial Expression Database

The popular JAFFE facial expression database [8] used in this study contains 213 images of female facial expressions. Each image has a resolution of 256×256 pixels. The head is almost in frontal pose. The number of images corresponding to each of the 7 categories of expression (neutral, happiness, sadness, surprise, anger, disgust and fear) is almost the same. A few of them are shown in Fig.2.

As done in [4, 5], we normalized the faces to a fixed distance of 55 pixels between the two eyes. Automatic face registration can be achieved by a robust real-time face detector based on a set of rectangle haar-like features [9]. From the results of automatic face detection, such as face location, face width and face height, two square bounding boxes for left eye and right eye are created respectively. Then, two eyes location can be quickly worked out in terms of the centers of two square bounding boxes for left eye and right eye. Based on the two eyes location, facial images of 110×150 pixels were cropped from original frames. No further alignment of facial features such as alignment of mouth was performed in our work.

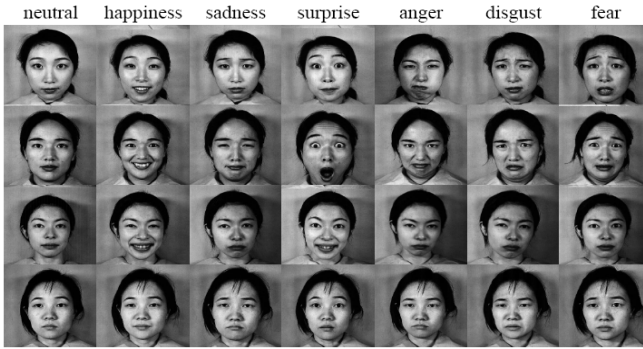


Fig. 2. Examples of facial images from JAFFE database

The cropped facial images of 110×150 pixels contain facial main components such as mouth, eyes, brows and noses. The LBP operator is applied to the whole region of the cropped facial images. For better uniform-LBP feature extraction, two parameters, i.e., the LBP operator and the number of regions divided, need to be optimized. Similar to the setting in [10], we selected the 59-bin operator, and divided the 110×150 pixels face images into 18×21 pixels regions, giving a good trade-off between recognition performance and feature vector length. Thus face images were divided into 42 (6×7) regions, and represented by the LBP histograms with the length of 2478 (59×42).

5 Experimental Results

For LS-SVM, the most popular radial basis function (RBF) kernel was used for its better performance compared with other kernels. The RBF kernel is defined as

$$K(x_i \cdot x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (3)$$

The kernel parameter γ of RBF kernel was optimized by using a grid search in the hyper-parameter space. For all facial expression recognition experiments, a 10-fold stratified cross validation scheme was performed and the average recognition results were reported. To begin with, we compared the different performance of four

encoding methods, i.e., one-versus-one, one-versus-all, MOC and ECOC. Table 1 presents the different performance of four encoding methods. From the results in Table 1, we can see that the one-versus-one method obtains the best performance of 78.57%, significantly outperforming the other encoding methods.

Table 1. Performance comparisons of four encoding methods for LS-SVM

Methods	one-versus-one	one-versus-all	MOC	ECOC
Accuracy (%)	78.57	40.34	58.57	54.76

Table 2. Confusion matrix of recognition results with one-versus-one for LS-SVM

	Anger (%)	Joy (%)	Sad (%)	Surprise (%)	Disgust (%)	Fear (%)	Neutral (%)
Anger	85.19	0	0	0	11.11	0	3.70
Joy	0	92.86	3.57	0	0	0	3.57
Sad	0	3.31	63.33	0	10	13.35	10
Surprise	0	3.45	0	93.10	0	0	3.45
Disgust	3.70	0	3.70	0	74.07	18.53	0
Fear	0	0	3.23	6.45	16.13	74.19	0
Neutral	0	0	3.45	0	0	3.45	93.10

To further investigate the recognition results of different facial expressions when using LS-SVM with one-versus-one method, the confusion matrix of facial expression recognition results is given in Table 2. The confusion matrix of the results in Table 2 shows that three facial expressions, i.e., neutral, surprise and joy could be discriminated well. In detail, the recognition rates are 93.10% for neutral, 93.10% for surprise, and 92.86% for joy, respectively. While the other three facial expressions, i.e., sad, disgust and fear, could be classified with much lower recognition rates, in detail 63.33% for sad, 74.07% for disgust and 74.19% for fear.

6 Conclusions

In this paper, we presented a new method of facial expression recognition incorporating LBP with LS-SVM. With LS-SVM, four different encoding methods such as one-versus-one, one-versus-all, MOC and ECOC are investigated. The experiment results on the popular JAFFE facial expression database indicate that LS-SVM with one-versus-one obtains the best performance of 78.57%. It's worth pointing out that in our work only 7 categories of expression (neutral, happiness,

sadness, surprise, anger, disgust and fear) from the JAFFE database could be classified, it's thus an interesting task to construct a large facial expression database including more than 7 facial expressions and further investigate the performance based on LBP and LS-SVM.

Acknowledgments. This work is supported by Zhejiang Provincial Natural Science Foundation of China under Grant No.Z1101048 and No. Y1111058.

References

1. Tian, Y., Kanade, T., Cohn, J.: Facial expression analysis. In: Handbook of Face Recognition, pp. 247–275 (2005)
2. Lyons, M., Akamatsu, S., Kamachi, M., Gyoba, J.: Coding facial expressions with Gabor wavelets. In: Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, pp. 200–205 (1998)
3. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray scale and rotation invariant texture analysis with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7), 971–987 (2002)
4. Shan, C., Gong, S., McOwan, P.: Robust facial expression recognition using local binary patterns. In: IEEE International Conference on Image Processing (ICIP), Genoa, pp. 370–373 (2005)
5. Shan, C., Gong, S., McOwan, P.: Facial expression recognition based on Local Binary Patterns: A comprehensive study. *Image and Vision Computing* 27(6), 803–816 (2009)
6. Vapnik, V.: The nature of statistical learning theory. Springer, New York (2000)
7. Suykens, J.A.K.: Least squares support vector machines. World Scientific Pub. Co. Inc. (2002)
8. Lyons, M.J., Budynek, J., Akamatsu, S.: Automatic classification of single facial images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(12), 1357–1362 (1999)
9. Viola, P., Jones, M.: Robust real-time face detection. *International Journal of Computer Vision* 57(2), 137–154 (2004)
10. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28, 2037–2041 (2006)

Facial Expression Recognition Using Kernel Locality Preserving Projections

Xiaoming Zhao¹ and Shiqing Zhang²

¹Department of Computer Science, Taizhou University,
318000 Taizhou, China

²School of Physics and Electronic Engineering, Taizhou University,
318000 Taizhou, China

{tzxyzxm, tzczsq}@163.com

Abstract. Considering the nonlinear manifold structure of facial expression images, a new facial expression recognition method based on kernel locality preserving projections (KLPP) is proposed in this paper. KLPP is used to extract the low-dimensional embedded data representations from the original extracted high-dimensional local binary patterns (LBP) facial features. The experimental results on the popular Cohn-Kanade facial expression database demonstrate that KLPP obtains the best accuracy of 86.94%, outperforming the other used methods including locality preserving projections (LPP) and kernel Isomap (KIsomap).

Keywords: Facial expression recognition, Local binary patterns, Kernel locality preserving projections.

1 Introduction

Facial expression provides a natural and immediate indication about the affective states of human beings. Automatic facial expression recognition aims to analyze one's emotions and intentions, and serves in various applications such as human-computer interaction, video indexing, data driven animation, etc. Facial expression recognition is a challenging task and has increasingly attracted much attention in the artificial intelligence, pattern recognition, computer vision, etc.

As shown in [1, 2], an N -dimensional facial representation can be considered a point in an N -dimensional face space, and the variability of facial expression can be represented as a low-dimensional nonlinear manifold embedded in this space. People change facial expressions continuously over time. Thus all images of an individual's facial expressions represent a smooth manifold in the N -dimensional face space with the 'neutral' face as the central reference point. Given the nonlinear manifold structure of facial expression images, some nonlinear dimensionality reduction (also called manifold learning) methods such as isometric feature mapping (Isomap) [3], locally linear embedding (LLE) [4] have been tried to apply for facial expression analysis. However, these two methods, i.e., ISOMAP and LLE, are defined only on the training data points and it is unclear how to evaluate the map for new test points. Therefore, they may not be very suitable for facial expression recognition. In recent years, a general nonlinear dimensionality reduction method called locality preserving

projections (LPP) [5] is developed. LPP is designed to obtain the optimal linear approximations to the eigenfunctions of the Laplace Beltrami operator on the manifold. LPP has two predominant characteristics. LPP can preserve the local structure of the data space as LLE done. More crucially, as a linear method, LPP can be defined everywhere in the ambient space rather than just on the training data. So, LPP has significant advantage over LLE in locating and explaining new test data in the reduced subspace. However, LPP can not explore higher order information of data, since LPP is not a kernel method. To tackle this problem, the kernel variant of LPP (KLPP) [5] is also developed.

In this paper, motivated by the deficiency of studying on KLPP for facial expression recognition, we investigate the performance of KLPP on facial expression recognition tasks. The facial expression recognition experiments are performed on the benchmark Cohn-Kanade [6] facial database.

2 Kernel Locality Preserving Projections

2.1 Locality Preserving Projections

Given data points $x_i, i=1,2,\dots,m$, let $y_i \in R^d$ be the low-dimensional data representation of a sample x_i , where d is the dimension of the embedding space. LPP finds a transformation matrix w by minimizing the following objective function

$$\min_w \sum_{i,j} (w^T x_i - w^T x_j)^2 S_{ij} \tag{1}$$

where S_{ij} evaluates the local structure of data space. If x_i and x_j are close,

$S_{ij} = e^{-\frac{\|x_i - x_j\|^2}{t}}$, where t is a parameter that can be determined empirically, otherwise, $S_{ij} = 0$. By simple algebra formulation, the objective function can be reduced to:

$$\begin{aligned} \frac{1}{2} \sum_{i,j} (w^T x_i - w^T x_j)^2 S_{ij} &= \sum_i w^T x_i D_{ii} x_i^T w - \sum_{ij} w^T x_i S_{ij} x_j^T w \\ &= w^T X(D - S)X^T w = w^T X L X^T w \end{aligned} \tag{2}$$

where $X = [x_1, x_2, \dots, x_m]$ and D is a diagonal matrix whose entries are column (or row, since S is symmetric) sums of S , $D_{ii} = \sum_j S_{ji}$. $L = D - S$ is the Laplacian matrix. The bigger the value D_{ii} (corresponding to y_i) is, the more important is y_i . Therefore, a constraint is imposed as follows:

$$y^T D y = 1 \Rightarrow w^T X D X^T w = 1 \tag{3}$$

Finally, the minimization problem reduces to finding:

$$\begin{aligned} \arg \min w^T X D X^T w \\ \text{s.t. } w^T X D X^T w = 1 \end{aligned} \tag{4}$$

The transformation vector w that minimizes the objective function is given by the minimum eigenvalue solution to the generalized eigenvalue problem:

$$XLX^T w = \lambda XDX^T w \tag{5}$$

It is easy to show that the matrices XLX^T and XDX^T are symmetric and positive semidefinite. The vectors w_i that minimize the objective function are given by minimum eigenvalues solutions to the generalized eigenvalues problem. Let the column vectors $w_0, w_1, w_2, \dots, w_{d-1}$ be the solutions of Eq. (5), ordered according to their eigenvalues, $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_{d-1}$. Thus, the embedding is as follows

$$x_i \rightarrow y_i = w^T x_i, w = [w_0, w_1, w_2, \dots, w_{d-1}] \tag{6}$$

where y_i is d -dimensional vector, and w is a $m \times d$ matrix.

2.2 Kernel Locality Preserving Projections

Let $\varphi(X) = [\varphi(x_1), \varphi(x_2), \dots, \varphi(x_m)]$ denote the data matrix in the Hilbert space, the eigenvector problem in the Hilbert space can be rewritten as follows:

$$[\varphi(X)L\varphi(X)^T]w = \lambda[\varphi(X)D\varphi(X)^T] \tag{7}$$

To generalize LPP to the nonlinear case, we formulate it in a way that uses dot product on the Hilbert space given by the following kernel function:

$$K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle = \varphi(x_i)^T \varphi(x_j) \tag{8}$$

By simple algebra formulation, we can finally obtain the following eigenvector problem:

$$K L K w = \lambda K D K w \tag{9}$$

For the original training points, the maps can be obtained by $y = K w$, therefore the eigenvector problem can be reduced to

$$L y = \lambda D y \tag{10}$$

3 Facial Expression Database

The benchmark Cohn-Kanade [6] database used in this study consists of 100 university students aged from 18 to 30 years, of which 65% were female, 15% were African-American and 3% were Asian or Latino. Subjects were instructed to perform a series of 23 facial displays, six of which were based on description of prototypic emotions. Image sequences from neutral to target display were digitized into 640*490 pixels with 8-bit precision for grayscale values. In our experiments, we selected 320 image sequences from the Cohn-Kanade database. The selected sequences, each of which could be labeled as one of the six basic emotions, come from 96 subjects, with 1 to 6 emotions per subject. For each sequence, the neutral face and three peak frames

were used for prototypic expression recognition, resulting in 1409 images (96 anger, 298 joy, 165 sad, 225 surprise, 141 fear, 135 disgust and 349 neutral). A few of them are shown in Fig.1.



Fig. 1. Examples of facial images from the Cohn-Kanade database

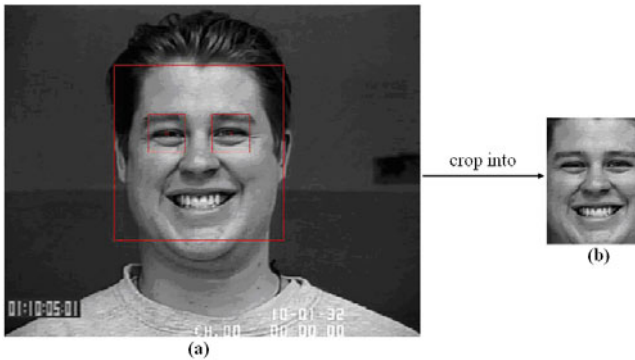


Fig. 2. (a) Two eyes location (b) the final cropped image of 110x150 pixels

As in [7], we normalized the faces to a fixed distance of 55 pixels between the two eyes. Automatic face registration can be achieved by a robust real-time face detector based on a set of rectangle haar-like features [8]. From the results of automatic face detection, such as face location, face width and face height, two square bounding boxes for left eye and right eye are created respectively. Then, two eyes location can be quickly worked out in terms of the centers of two square bounding boxes for left eye and right eye. Based on the two eyes location, facial images of 110x150 pixels were cropped from original frames. Fig.2 shows the process of two eyes location and the final cropped image. No further alignment of facial features such as alignment of mouth was performed in our work.

4 Experiment Study

Since the local binary patterns (LBP) [9] has a predominant characteristic, that is, LBP tolerates against illumination changes and operates with its computational simplicity, we adopted LBP for facial image representations for facial expression recognition. The cropped facial images of 110×150 pixels contain facial main components such as mouth, eyes, brows and noses. The LBP operator is applied to the whole region of the cropped facial images. For better uniform-LBP feature extraction, two parameters, i.e., the LBP operator and the number of regions divided, need to be optimized. Similar to the setting in [7], we selected the 59-bin operator, and divided the 110×150 pixels face images into 18×21 pixels regions, giving a good trade-off between recognition performance and feature vector length. Thus face images were divided into 42 (6×7) regions, and represented by the LBP histograms with the length of 2478 (59×42).

Table 1. The best accuracy for different methods with corresponding reduced dimension

Methods	LPP	KIsomap	KLPP
Dimension	17	20	18
Accuracy (%)	61.23	74.82	86.94

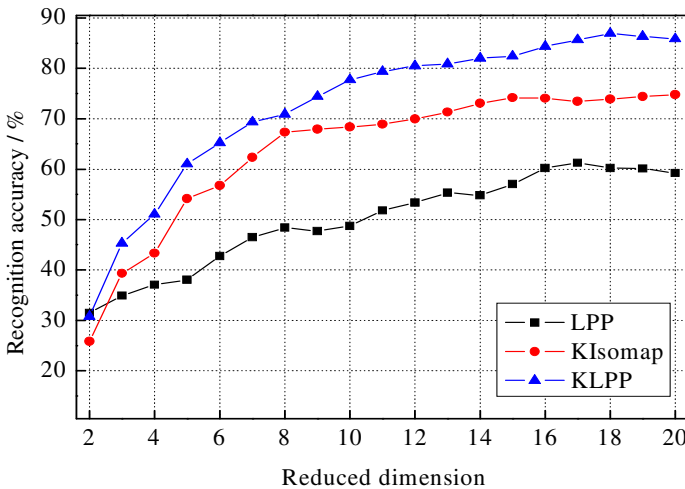


Fig. 3. Performance comparisons of different methods

The nearest neighbor classifier with the Euclidean distance was used to conduct facial expression recognition experiments owing to its simplicity. The performance of KLPP is compared with LPP and kernel Isomap (KIsomap) [10]. The reduced feature

dimension is limited to the range [2, 20]. For all facial expression recognition experiments, a 10-fold stratified cross validation scheme was performed and the average recognition results were reported.

The different recognition results of three used dimension reduction methods, i.e., LPP, KLPP and KIsomap, are given in Fig.3. The best accuracy for different methods with corresponding reduced dimension is presented in Table 1. As shown in Fig.3 and Table 1, we can see that KLPP performs best and obtains the highest accuracy of 86.94% with 18 embedded features. This is attributed to the fact that KLPP can explore and extract higher order nonlinear information as a kernel method.

5 Conclusions

In this paper, we presented a new method of facial expression recognition based on KLPP. The experiment results on the popular Cohn-Kanade facial expression database indicate that KLPP outperforms LPP and KIsomap. This demonstrates that KLPP has the better ability than LPP and KIsomap to extract the low-dimensional embedded data representations for facial expression recognition.

Acknowledgments. This work is supported by Zhejiang Provincial Natural Science Foundation of China under Grant No.Z1101048 and No. Y1111058.

References

1. Chang, Y., Hu, C., Feris, R., Turk, M.: Manifold based analysis of facial expression. *Image and Vision Computing* 24, 605–614 (2006)
2. Cheon, Y., Kim, D.: Natural facial expression recognition using differential-AAM and manifold learning. *Pattern Recognition* 42, 1340–1350 (2009)
3. Tenenbaum, J.B., de Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. *Science* 290, 2319–2323 (2000)
4. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. *Science* 290, 2323–2326 (2000)
5. He, X., Niyogi, P.: Locality preserving projections. In: *Advances in Neural Information Processing Systems (NIPS)*, vol. 16, pp. 153–160. MIT Press, Cambridge (2003)
6. Kanade, T., Tian, Y., Cohn, J.: Comprehensive database for facial expression analysis. In: *International Conference on Face and Gesture Recognition*, Grenoble, France, pp. 46–53 (2000)
7. Shan, C., Gong, S., McOwan, P.: Facial expression recognition based on Local Binary Patterns: A comprehensive study. *Image and Vision Computing* 27(6), 803–816 (2009)
8. Viola, P., Jones, M.: Robust real-time face detection. *International Journal of Computer Vision* 57(2), 137–154 (2004)
9. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray scale and rotation invariant texture analysis with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7), 971–987 (2002)
10. Choi, H., Choi, S.: Robust kernel isomap. *Pattern Recognition* 40, 853–862 (2007)

Author Index

- Ao, Youyun 95
- Bai, Fan 325
- Bai, Hui 333
- Bai, Jie 103, 109
- Bellafkih, Mostafa 465
- Cao, QingKui 543
- Cao, Zaihui 411
- Chang, Xiaoying 627
- Chen, Caiyuan 459
- Chen, Dyi-Cheng 279
- Chen, Gang 285
- Chen, Guochu 621
- Chen, Jia-Yu 575
- Chen, Kunchang 403
- Chen, Liang 601, 689
- Chen, ShaoChun 583, 589
- Chen, Wen-Jong 243, 279
- Chen, Ye 249
- Chen, ZhiWen 453
- Chen, Zhong 659
- Chen, Zhonghua 325
- Cheng, Changxian 291
- Chi, HuaShan 203
- Choi, In-O 389
- Chu, Jian 371
- Cong, Beihua 549
- Cui, Shigang 157, 163
- Dai, Yanyan 325
- Di, YongJiang 151
- Du, Shuxing 13
- Fekibelu, Abreham 157
- Feng, Ansong 627
- Feng, Jinchun 425
- Feng, Xiaoliang 297
- Fu, Liqin 45
- Fu, Zheng 7
- Gao, Guohong 1
- Gao, HuiSheng 237
- Gao, Ji 81
- Gao, Liang 677
- Gao, Qiang 211
- Gao, Rui 317
- Gao, Wei 595
- Gao, Ying 563
- Guo, JiWen 341
- Guo, Mingli 1
- Guo, YongNing 645, 659, 669
- Han, Xin 549
- Hao, Huan 601
- Hao, ZhanGang 377
- He, Shijun 325
- He, Xie 549
- Hu, Chunsheng 249
- Hu, Qiang 195
- Hu, Tianhao 403
- Hu, Wen-Song 317
- Hu, Xinru 145
- Hu, Zhangping 689
- Hu, Zhongyan 411
- Huang, Minxue 297
- Jeong, Yong Mu 63
- Jia, Bei 259
- Jia, Zhiqiang 67
- Jiang, Tong 1

- Jiao, Bin 633
 Jin, Dan 555
 Jin, Pan 175
 Jun, Chen 615

 Kaixiong, Su 615
 Kim, Young-Jun 389
 Kong, Haiyan 347
 Kong, Liu-an 19
 Kou, Lingyan 25

 Lee, Seung Eun 63
 Lei, DaoCheng 217
 Leng, Kaijun 517
 Li, Chaoliang 309
 Li, GaoJun 103
 Li, Huijuan 259
 Li, Jia-Hao 439
 Li, Kai 13
 Li, Kai-Yan 431
 Li, Kun 639
 Li, Kunlun 333
 Li, Lingling 145
 Li, LinKun 365
 Li, Qiong 115
 Li, Renwang 403
 Li, Rui 491
 Li, Shengwei 491
 Li, Wei 383
 Li, Xiang 145
 Li, XiaoRui 237
 Li, YiMing 365
 Li, Yue 621
 Lian, Ying 583, 589
 Liang, Guozhuang 137
 Liao, Wenjun 1
 Lim, Ki-Taek 63
 Lin, Hwei-Jyun 431
 Liu, Chenglian 669
 Liu, Donghan 115
 Liu, Guoqing 425
 Liu, Leiming 609
 Liu, Limin 353, 359, 447
 Liu, Maohong 297
 Liu, XiaoMei 583, 589
 Liu, Xuewu 309
 Liu, YanHui 267
 Lu, HuiShan 211
 Lu, Qing 371
 Luo, Weiping 67

 Ma, RunNian 217
 Ma, Zhenliang 677
 Meng, Fanming 481
 Miao, Rui 621
 Ming, Zhimao 39
 Mo, Canlin 403
 Mo, Mingzhong 181

 Nian, Fung-Ling 279

 Park, Sang-Hong 389
 Peng, Lingyi 309
 Peng, Zhang 175

 Ramdani, Mohammad 465
 Ren, Zhiqian 39
 Rui, Guosheng 7

 Sarem, Mohammad 465
 Sha, Junchen 677
 Shao, Xiaoyan 145
 Shi, Jinfa 411
 Shi, Lichun 473, 497
 Shi, Xiaochun 425
 Shu, KeLiang 109
 Song, XiaoQian 377
 Su, Wen-Cheng 279
 Sun, Aiqin 33
 Sun, Can 677
 Sun, Jinsheng 73
 Sun, Juan 677
 Sun, Jun 537
 Sun, Mingjian 481
 Sun, Tai-Ping 431, 439, 575

 Tan, XiaoDong 103, 109
 Tang, ShanShan 583, 589
 Tian, Wenbiao 7
 Tong, Chaonan 609
 Tong, Hao 137
 Tsau, Fu-Shiung 439
 Tu, Li 303

 Wang, Baoping 259
 Wang, Changjiang 45
 Wang, Chao 341
 Wang, Chia-Hung 431, 439, 575
 Wang, Feifei 157, 163
 Wang, Gang 217
 Wang, HongXing 203

- Wang, Jidai 33
 Wang, Jun 237
 Wang, Lifeng 653
 Wang, Mei 273
 Wang, Miao 333
 Wang, Minghu 89
 Wang, Qinghui 563, 569, 627
 Wang, Ruihu 689
 Wang, ShanHui 639
 Wang, Tong 377
 Wang, XiaoBo 267
 Wang, Xudong 191
 Wang, Yan 303
 Wang, Yang 653
 Wang, YaPing 525
 Wang, Ying 1
 Wang, Yuxia 517
 Wang, Zhen-Yu 537
 Wang, Zhongqing 195
 Wei, XiaoLing 453
 Wen, Fuan 417
 Wen, Zhiwei 555
 Wu, Bin 537
 Wu, Dichong 403
 Wu, Guian 563
 Wu, Ming-Huang 243
 Wu, Wen-jing 19
 Wu, Xiao-Hong 537
 Wu, Xinli 403

 Xiao, Haimei 473, 497
 Xie, DaLi 267
 Xin, Ruining 333
 Xing, Jianping 677
 Xiong, Jing 689
 Xiong, Wen 57
 Xiong, Youping 633
 Xu, Chengdong 249
 Xu, DongYang 231
 Xue, HuaPing 569

 Yan, HongWei 211
 Yan, Hua 639
 Yan, Jianhua 291
 Yan, Jun 555
 Yang, Chunqing 187
 Yang, Erlong 273
 Yang, Genghuang 157, 163
 Yang, JingJing 569
 Yao, Jie 595
 Ye, Chenyun 25

 Ye, Jae-Heung 389
 Ye, ZhongFu 231
 Yi, Zhenzhen 13
 Yin, MaoWei 231
 Yin, Shaojie 73
 Yongquan, Zeng 169, 225
 Young, Jieh-Shian 395
 Yu, Fan 89
 Yu, Kaichao 365, 531
 Yu, YongLin 217
 Yu, Zhiheng 701

 Zhang, Bo 601
 Zhang, Cheng-Qiu 317
 Zhang, Denghui 81
 Zhang, Dongyue 33
 Zhang, Fan 491
 Zhang, Jia 531
 Zhang, Jing 51
 Zhang, JingHua 543
 Zhang, Longjun 505, 511
 Zhang, Min 203
 Zhang, Pengfei 249
 Zhang, Qicong 121, 129
 Zhang, Shiqing 707, 713
 Zhang, Shutai 325
 Zhang, Xiaoning 459
 Zhang, Xiaowei 481
 Zhang, Xiaoying 555
 Zhang, Xinsheng 89
 Zhang, Yinguang 601
 Zhang, Yipeng 601
 Zhang, Yongmei 45
 Zhang, Yu 325
 Zhao, Chengli 701
 Zhao, JianMing 645
 Zhao, Ke 13
 Zhao, Li 157, 163
 Zhao, PeiHong 203
 Zhao, Shiting 325
 Zhao, Xiaoming 707, 713
 Zhao, Xiaopeng 137
 Zhao, Yue 417
 Zheng, Gang 491
 Zheng, Zewei 653
 Zhong, WeiSheng 525
 Zhou, Jianzhong 115
 Zhou, Wenjun 325
 Zhou, ZhenZhong 453
 Zhu, Jin 595